

|CURIEUX|

ACADEMIC JOURNAL

November Issue

Part 2 Issue 44

Editing Staff

Chief-In-Editor

Caroline Xue

George Hasnah

Chief of Operations

Anshal Vyas

Assisting Editors

Olivia Li

Luzia Thomas

Madalyn Ramirez

Shefali Awasthi

Soumya Rai

Sarah Strick

Table Of Contents

Page 6: Conventional Braces and Clear Aligners: An Evaluation of Outcome and Treatment By Ana Kublashvili

Page 17: Color Psychology in Marketing: Unveiling the Impact of Color on Consumer Behavior and Brand Perception By Olivia Lee

Page 26: Rock Climbing as a Therapeutic Approach: Enhancing Cognitive and Motor Skills in Neurological Patients By Aanika Tipirneni

Page 39: Female Empowerment in the Context of Rousseau's French Enlightenment By Avery Ertman

Page 53: Congestive Heart Failure in Humans and Dogs: A Review By Sakura Nakanishi

Page 79: Third Order Fibonacci Sequences By Stefan Maxim

Page 87: Segregated Economies: The Lingering Effects of African American Exclusion on Southern Prosperity By Cooper Martin

Page 104: The Impact of GenAI on International Relations, and Proposed Countermeasures to be Employed by the Foreign Service to Adapt in a New Geopolitical Paradigm By Sydney Garber

Page 110: G.P. Maffei: Shining a New Light on China By Matthew Sennelius

Page 130: The Effects Of Music-Induced Nostalgia On Healthy Aging By Eliana Sun

Page 151: Novel Applications of Elastin-like Polypeptides (ELPs) in the Treatment of Pathogenic Free-Living Amoeba By Avni Kalra

Page 167: Empirically-Derived Applications of Music Therapy Across Lifespans By Chloe Yoon

Page 180: Figure Skating Training and Balance: A Quantitative and Comparative Analysis By Maytal Zasler

Page 189: Formation of Children Consumer Behaviour in The Present and Future By Wenjing Zhou

Page 198: Takeaways for Chinese Policymakers: Promoting Social Connectedness of Autistic Children and Young Adults in China By Xiaoxi Jia

Page 213: MSCs-Derived Exosomes as a Biomaterial-Based Cell-Free Treatment for Spinal Cord Injury By Davidoula Georgopoulou

Page 231: The Effect of Student Loan Debate Forgiveness on Social Mobility By Pranav Pradeep

Page 239: A Review of Brain Stimulation as an Alternative Method of Treating OCDBy Aiden Chung

Page 268: The Interaction of Soil-Transmitted Helminth Infections, Geophagy, and Pregnancy Outcomes in Sub-Saharan Africa By Minseo Ryu

Page 282: The Ideal Ultimate Frisbee Throw: Grid search for the perfect pass By Anthony Ding

Page 293: The Cause of the Politicization of the Judiciary and How It Threatens Democracy in South Korea By Taeun Kim

Page 304: The Evolution of French Profession Titles and Feminism By Xinyi Li

Page 311: Understanding Global Population to Understand International Relations By Baron Jiang

Page 318: Getting Green: Roadmap to Reducing Costs for III-V Multijunction Solar Cells By Andrew Carson

Page 328: Sound Mind, Sound Memory: The Music-Stress Connection. A Mini Review Exploring the Comparative Effect of Stress and Music on Learning, Memory, and Brain Waves By Sienna Phillips

Page 340: CAR T-cell Therapy By Sayan Majumder

Page 349: What recent advances have been made in C-H methylation for late-stage functionalization? By Angela Wang

Page 369: Applying AI to Enhance Stock Price Prediction: Improving Investment Strategies and Risk-Management Through Data-Driven Insights By Lakshya Gupta and Odysseas Drosis

Page 379: The Effect of Augmented Reality and Haptics Technology on Student Engagement and Learning in STEM Education By Zachary Jones

Page 396: The Economic Aspects of Schizophrenia Treatments By Abdul Azis Ramzi

Page 400: Diagnosing and Treating Cancer: An Inspection to Long Non-Coding RNAs and Immunology Therapy By Zehra Desen Çelik & Eden Ng

Page 407: The Impact of Hormonal Fluctuations on Brain Structure and Function in Women Across Reproductive Stages By Misha Nasarpuri

Page 418: The Business of K-Pop: The Role of Angel Investments and Venture Capital in Global Success By Tung-Tung Yeh

Page 425: Responsible AI for Nonprofits: Ensuring Data Privacy and Mitigating Bias in LLMs By Keshav Rangan

Page 448: Nietzsche's Nihilism and How to Lead a Meaningful Life To What Extent Is Nietzsche's Theory on Finding Meaning in Life True and Applicable to Our Modern Lives? By Benjamin Ooi

Page 457: Coagulation Clarification of Microplastics: Affordable, Natural, and Available Way to Treat Water for Rural Communities By Thiyagesh Venkatesan

Page 469: Stomach Cancer in South Korea: How Diet, H. pylori and Alcohol Impact Incidence Rates By Lenna Yoon

Page 483: Predicting TED Talk Views Using Topic Tags By Ernest Lai

Page 488: Soft Tones, Strong Identities: Mandarin Language Dynamics in Taipei, Taiwan By Sindhu Talluri

Page 506: Etruscan Burial Methods By Shiv Sibal

Page 521: Scream, Typecasting, and African Americans: A Content Analysis of Typecasting in the Scream film Series By Hayden Yoest

Page 542: Host Cities' Labor Market Reactions To The Super Bowl By Jason Luo

Page 560: Evaluation of Injury Risk Factors in Youth Soccer By Situ Wang

Page 567: The Marriage Penalty: Myra Bradwell's Fight for Married Women in the Law By Julia Nathanson

Page 579: Reproduction and Fertility: From Basic Science to Clinical Treatments By Dhaya Sarathy

Page 594: Designing an Affordable Prosthetic Hand with Bipolex Functionality By Taneesh Matharasi

Page 613: Schrödinger's Equation: Variations, Potentials, and Modern Applications By Neev Jain

Page 623: The Optimization of Neural Networks By Zubin Gupta¹, Cyrus Ayubcha²

Page 630: Academic Bias in Liberal and Conservative Areas: Analysis of Gender and Sexual Orientation Bias in Papers from Democrat and Republican-Leaning Regions By Keilah Park

Conventional Braces and Clear Aligners: An Evaluation of Outcome and Treatment

By Ana Kublashvili

Abstract

Conventional orthodontic treatment to treat a variety of malocclusions using wires, brackets, or ligatures has been around since 1770. Development of new research and technology in the late 20th and early 21st century brought CAD/CAM manufacturing, which has allowed the development of plastic orthodontics including Invisalign. This systematic review focused on the following question: In young adults in need of orthodontic treatment, which orthodontic appliance, traditional braces or clear aligners, will have the best outcome and treatment in the most effective way possible? Orthodontic wires such as Nickel-Titanium and Stainless steel exert a great amount of force once applied. The time for the clear aligners was most likely shorter as it does not need a detailing or finishing phase, while traditional braces take up to 6 months for it. However, the American Board of Orthodontics objective grading system showed that the aligners did not correct malocclusions as well as braces did. Data gathered by Kaklamanos et al (29) showed that aligners for treatment may offer further advantages to the improved Oral Health-Related Quality of Life (OHRQoL) seen with orthodontic correction, compared to traditional treatment using conventional metal fixed appliances. These advantages could include enhanced comfort during sleep, eating, and social interactions, as well as increased self-esteem and overall satisfaction with oral health(29). Clear aligners and traditional braces aim to accomplish the same result through different methods. However, this study showed clear aligners improve quality of life in a more efficient, time reduced manner. It also corrects malocclusion and provides better outcomes for patients with periodontal disease.

Introduction

Conventional orthodontic treatment to treat a variety of malocclusions using wires, brackets, or ligatures has been around since 1770. Orthodontic materials were first developed in 1887 by Dr. Edward Angle, who used nickel alloy for its flexibility in wires. Different materials like copper, zinc, and 14-karat gold became more prominent and widely used for traditional orthodontic treatment. Stainless steel was introduced in 1929 and it gained popularity for its fracture resistant property under stress, and less costly than gold(1). In 1944, removable appliances started being used for mild and moderate tooth movements. Development of new research and technology in the late 20th and early 21st century brought CAD/CAM manufacturing, which has allowed the development of plastic orthodontics including Invisalign. ALIGN technologies, the proprietary company of Invisalign, was founded in 1997 and the FDA approval for Invisalign was presented in 1999(2).

There are many options of clear aligners available currently. Invisalign is the most popular clear aligner company and is unique due to their Smart track polymer system. This multi-layered polymer system replaces traditional aligner materials with a more elastic material, enhancing fit, comfort, efficiency. Invisalign aims to speed up tooth movement by 50% and

increase precision by 75% compared to other materials(3). Furthermore, Invisalign also claims to resolve rotations of 40 degrees in upper and lower central incisors, 45 degrees in canines and premolars, 30 degrees in lateral incisors, and 20 degrees in molars. SureSmile, another aligner brand, uses Essix ACE Plastic, a thinner, singled layered material (4). SureSmile has shown studies that illustrate a reduced treatment time than the traditional, wire approach with improved quality. Another aligner brand is ClearCorrect, made out of a tri-layer material, two layers of polymers with an elastomeric inner layer. ClearCorrect aligners have shown force retention overtime, have kept its shape throughout the wear, and have exhibited durability and demonstrated 4x better tear resistance than single layer polymers(5).

Traditional orthodontic treatment usually consists of NiTi (nickel titanium) and Stainless steel wires. Stainless steel alloys offer resistance to corrosion, greater rigidity, and reduced friction. Stainless steel alloys also come in multi-stranded variations, which exhibit lower rigidity that allows them to be used in the early stages of treatment. NiTi alloys differ from stainless steel in their flexibility and shape memory, which allows them to revert to their original shape even with significant force(6). Thermodynamic NiTi alloys become activated under a certain temperature so it becomes easier to insert NiTi wires into bracket slots during application. Many activations under a specific temperature provide different advantages. For example, alloys activated under 27°C exert greater loads, working well in the oral cavity, where the temperature is around 36-37°C. Wires activated at 40°C are well-fit for patients with high sensitivity, as the wires undergo minimal activation during warm salt water mouth rinses(7).

Braces and clear aligners, both effective in orthodontic treatment, use different approaches to get desired outcomes. Orthodontic treatment seeks to correct malocclusions or bite problems. There are 2 common types of malocclusion, overbite and underbite. Overbite is a clinical condition where the maxillary teeth are angled forward covering the mandibular teeth(8). In an underbite, the mandible is extended out covering the tip of the maxillary teeth(9). Intrusion and extrusion are additional malocclusion issues that can be present. Intrusion refers to the downward movement of the tooth into the alveolar bone, and extrusion displacement of a tooth out of its alveolar housing(10). Additionally, there are 3 different types of malocclusion classes. Class 1 is crowding around or spacing issues around the teeth. Class 2 is an overbite when the maxillary front teeth are positioned too forward protruding over the lower incisors. Class 3 is an underbite where the lower front teeth are positioned in front of the top front teeth(11). Figure 1 below demonstrates the three malocclusion classes discussed above. Overall gingival health is a big reason why orthodontic treatment is needed, as gingivitis and periodontal disease are common in patients before getting treatment (12). In extreme cases orthognathic jaw surgery is needed for malocclusions, facial asymmetry, and problems with jaw growth(13).

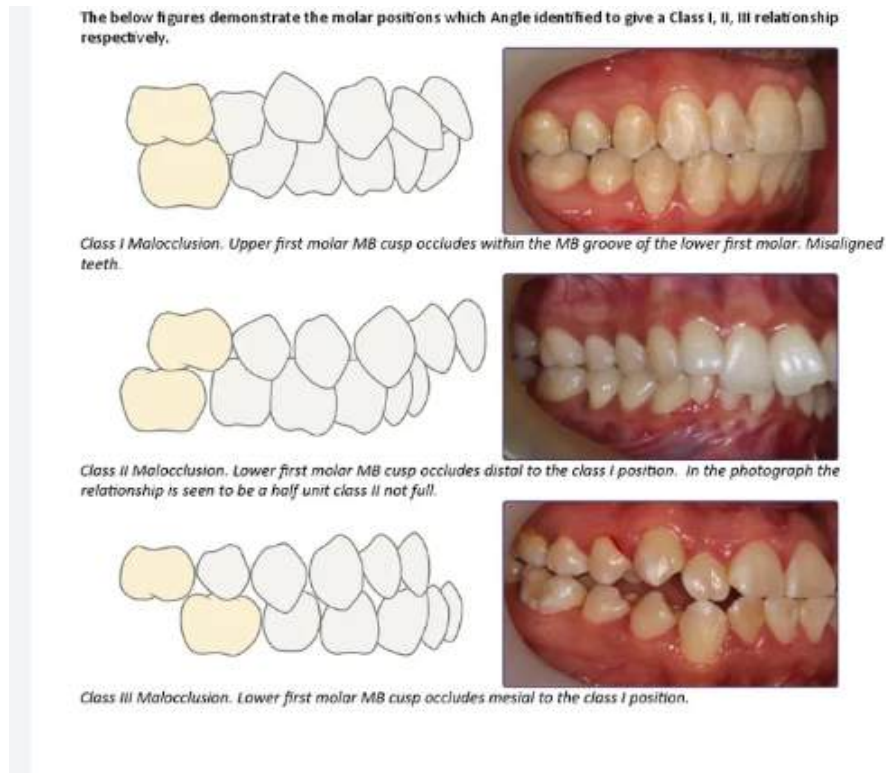


Figure 1- Illustration of class I, class II, and class III malocclusions. In Class II, maxillary teeth are positioned more anteriorly in comparison to mandibular teeth. In Class III, maxillary teeth are positioned more posteriorly in comparison to mandibular teeth.

During orthodontic movement the body goes through biological changes. During the initial phase, the teeth begin to shift, affecting the tissue around it and forcing the periodontal ligament surrounding the tooth to stretch and compress, which causes the blood vessels to leak, attract inflammatory cells, and recruit cells that help with bone remodeling(14). Throughout this process, gingival inflammation is present because of plaque build-up as wearing orthodontic devices makes it difficult to clean, especially in interproximal spaces. (15). Additionally, the dental pulp is affected during orthodontic treatment as pulpal tissue sits in hard tooth structure that is dependent on blood vessels passing through, any problem with blood flow and tissue pressure can affect dental pulp (16). As a result of orthodontic force, the periodontal ligament is affected as it undergoes a remodel with the breakdown of bone on the compressed side and the buildup of bone on the tension side (17). The metabolic state of the bone is influenced during orthodontic treatment, as bone turnover affects the rate of tooth movement(18).

MATERIALS/METHODS

This systematic review focused on the following question: In young adults in need of orthodontic treatment, which orthodontic appliance, traditional braces or clear aligners, will have the best outcome and treatment in the most effective way possible?

The definitions of population, intervention, comparison, and outcome (PICOs) were developed based on the focused question as follows:

Population: Young adults needing orthodontic treatment

Intervention: Traditional braces

Comparison: Clear aligners

Outcomes: Best outcome in the most effective way possible

Search strategies

An electronic search without time or language restrictions was conducted using Pubmed, Google Scholar, and other published articles. The reference lists of included studies and relevant reviews were also searched for other potential studies. The detailed search strategies were as follows: Orthodontics AND wires, wires AND brackets, clear aligners AND braces, Invisalign AND efficiency, periodontal status AND aligners, braces AND force, cost WITH clear aligners, oral hygiene AND traditional braces, mechanism AND clear aligners, 'Quality of life' with braces AND aligners, treatment AND braces,

Results/Discussion

Orthodontic treatments, including clear aligners and braces use distinct mechanisms to straighten teeth. A major difference in mechanism between clear aligners and traditional wire is the force, elasticity, and angulation used. The force used with traditional orthodontic wire and clear aligners differs significantly. Clear aligners have two different systems, the displacement driven system and the force-driven system, as shown in figure 2. The displacement driven system is a method orthodontics uses to create a custom set of aligners to apply gentle pressure and shift the teeth overtime. The system guides simple movements of the teeth such as minor rotations or tipping. The system is less effective in complex movements and root adjustments. The force-driven system designs the aligners so they are able to apply force in targeted areas. In certain cases, the aligners will be altered with pressure points for more difficult tooth movements, such as intrusion and uprighting. In some cases, power ridges, which are strategically placed bumps on the aligner to exert precise forces on specific teeth will be used to control root torque and to get desired force on the specific area (19) (20). Additionally, the force and magnitude is determined on the configuration of the aligners, with each tooth receiving a certain magnitude and a type of force. Aligners designed with specific intrusion patterns show force exerted on different types of teeth such as, incisors, canines, and premolars. This variety of force application is helpful in correcting deep bite issues, as it affects how effectively aligners achieve the desired outcome for the teeth. Most aligners are made from a polyurethane plastic material which severely influences their mechanical properties and force distribution(21). Comparing elasticity, clear aligner materials have viscoelastic properties which have viscous and elastic materials. Invisalign uses a polymer called SmartTrack, which gives the aligners elasticity and produces constant forces which improves overall efficiency.

Force-driven system vs. Displacement-driven system

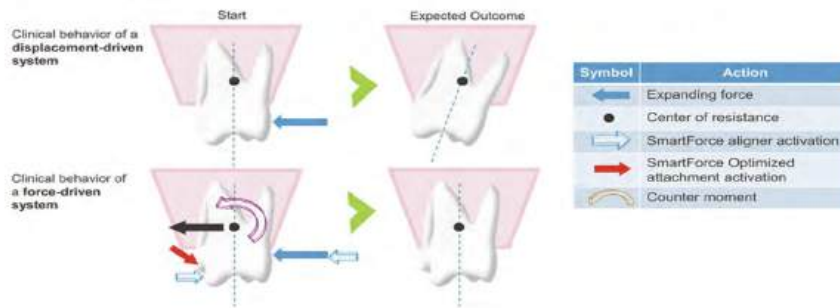


Figure 2- An illustration describing clear aligner force systems, the force driven system and displacement-driven system. The displacement driven system relies on physically moving teeth into new positions. The force driven system uses precise forces to move teeth in the desired position.

Orthodontic wires such as Nickel-Titanium and Stainless steel exert a great amount of force once applied. Friction is very influential on the force applied because it can determine the amount of force the wires exert. For example, stainless steel wires have lower friction rates which allows greater force to be released. Additionally, the low friction rates for stainless steel wires allows them to be less resistant to tooth movement compared to other alloys(22). Orthodontic treatment utilizes NiTi wires to achieve optimal results, considering factors like the degree of deflection, ligation techniques, and frictional forces to ensure the most effective force application for desired outcomes(21). A study done by Garner et al (22) observed that frictional forces between brackets and wires are greater with NiTi wires compared to stainless steel wires, but lower than those with beta-titanium wires, especially in zero torque or angulated brackets. Orthodontic wires like Niti and stainless steel have lower stiffness, making them flexible and more elastic. Due to the high yield strength and elasticity of stainless steel, stresses can severely impact the wire's elastic properties after bending. Therefore, heat treatment is used to relieve stress in stainless steel wires bent into arches, loops, or coils, thereby enhancing their elasticity. In NiTi the most beneficial characteristics are its springback and flexibility. NiTi's high springback is used in cases that need low forces but large elastic deflections. Also, it is commonly observed that NiTi wires exhibit more pronounced springback and higher revocable energy compared to stainless steel or beta-titanium wires(22). Zirconia brackets are a type of orthodontic bracket made from zirconium oxide and gained its popularity because of its durability and appearance. They are mainly used when strong movements are needed but patients would like a more aesthetic treatment option. Zirconia brackets are highly resistant to staining and discoloration, making them a good option for patients who want them to be more discreet. They also exhibit superior mechanical properties such as higher fracture toughness and better performance under stress compared to ceramic brackets. Although, one of the challenges with zirconia brackets is their higher friction with orthodontic wires which can affect the efficiency of tooth movement(41).

During orthodontic treatment, the quality of life and pain throughout is a key factor orthodontists need to look for. For patients who need orthognathic surgery, studies show that

clear aligners have a better quality of life outcome, and lower amount of pain after surgery than traditional braces. In the study by Patricia de Leyva (23), 2023, patients with dental deformities who underwent orthognathic surgery randomly received post-orthodontic treatment of either traditional braces or clear aligners. Quality of life was assessed through the Orthognathic Quality of Life Questionnaire and the Oral Health Impact Profile (23). Additionally, another study by Paula Coutinho Cardoso had a group of patients use invisalign and traditional braces(24). This study agrees with the most recent studies showing a greater incidence of pain when fixed appliances are used.

A key factor in orthodontics is the time the treatment takes to get the desired outcome. In many cases, aligners were effective in rapid orthodontic treatment. Rapid orthodontic treatment is an accelerated approach to straighten teeth and fix misalignments. A study by Dr. Edmund Khoo (25) shows a case of 3 patients who got micro-osteoperforation (MOPs) done, a technique to accelerate ortho movement. MOPs are created with tiny openings in the alveolar bone requiring movement typically under local anesthesia, without the necessity of lifting a tissue flap. These patients were treated with Invisalign. All three patients were satisfied with the outcome and stated that there was very little pain and felt mild gingival inflammation a day after the procedure(25). Three clinical studies showed that PAOO resulted in faster leveling and alignment compared to traditional orthodontic treatment, with accelerated percentages of 39%(246 days versus 402), 46% (171.9 versus 314 days), and 47% (74.5 versus 141.7 days) respectively for each study. Furthermore, two investigations demonstrated that PAOO accelerated the retraction of upper interior teeth, with acceleration rates of 44% (156 versus 441 days) and 61% (130.5 versus 234.1 days) (26) respectively. To achieve the desired outcome, it is very important to know which treatment with time efficiency is better. In cases not requiring rapid orthodontic movement, similar results are found in terms of timing for case completion. A study by Buschanga et al showed that during traditional treatment aligners took 11.5 months to achieve desired outcome while traditional braces took 17 months to fix malocclusions in the teeth (27). The factors included chair time, doctor time, and overall patient compliance. The time for the clear aligners was most likely shorter as it does not need a detailing or finishing phase, while traditional braces take up to 6 months for it. However, the American Board of Orthodontics objective grading system showed that the aligners did not correct malocclusions as well as braces did.

During orthodontic treatment it is very common for overall oral hygiene to be affected. In clear aligners compared with traditional braces, clear aligners facilitate better oral hygiene which improves periodontal status, decreases bleeding on probing and gingival inflammation(19). Traditional braces make it harder to clean those hard to reach spaces causing more problem areas(28). Clear aligners are significantly better at improving gingiva health, and more specifically periodontal disease, as patients have a greater ability to clean their teeth when the appliance is removed. However, traditional braces affect plaque removal, gingival health, and makes gingivitis more prominent. The bands, elastics, brackets and wires in traditional orthodontic treatment carry bacteria which can then cause periodontal disease.

Studies have shown it is best to perform orthodontic movement after active periodontal disease has been treated. Orthodontic movement can enhance bacteria and cause periodontal inflammation. In clear aligners, caries can be affected as well. The biomaterials in clear aligners can result in a growth of bacteria associated with caries and the emergence. In traditional braces as stated before braces make it harder to clean the teeth, which causes caries to become more prominent from the bacteria in the mouth. White spot lesions are common to show up during orthodontic treatment in both clear aligners and traditional braces. Again comparing the two, many cases show that patients who have traditional braces get white spot lesions because fixed appliances promote bacteria, plaque build up and limit the ability to clean the teeth. Compared to aligners they have a much better success rate with white spot lesions, since they are more accessible to clean and bacterial plaque doesn't get missed while brushing. A study by de Leyva et al (23) compared bleeding on probing and probing depth. Clear aligners had significantly lower probing depth compared to traditional braces, and comparing bleeding on probing clear aligners only had 1 where traditional braces had 8 (23). Another study by Luca Levrini et al (28) compares bleeding on probing between clear aligners and traditional braces, which concluded that clear aligners had less bleeding on probing compared the fixed appliances(28).

After orthodontic treatment, patients with malocclusions treated with aligners showed significantly better results in treating the malocclusions compared to traditional braces. However, in a 6-month post-treatment period for aligners, more relapse was shown compared to traditional braces. Data gathered by Kaklamanos et al (29) showed that aligners for treatment may offer further advantages to the improved Oral Health-Related Quality of Life (OHRQoL) seen with orthodontic correction, compared to traditional treatment using conventional metal fixed appliances. These advantages could include enhanced comfort during sleep, eating, and social interactions, as well as increased self-esteem and overall satisfaction with oral health(29). Another study was done by Qiuying Li et al, yielding similar results on the QoL of braces and aligners(30). A study done by Di Spirito et al (31) showed after a 6 month period, clear aligners had a better periodontal status compared to traditional braces(31). Studies about rebound percentage of clear aligners and traditional braces are very limited. However, a study done by Papadimitriou et al,(32) shows how teeth alignment with clear aligners deteriorate quicker than with braces. This implies that the rebound rate would be higher in aligners than in braces(32).

The total material cost for clear aligners is calculated differently. The price range for aligner systems in the United States range from \$2,650 and \$7,000. More specifically, in Invisalign, the cost varies depending on how long the treatment will be. If the patient's treatment plan is 6-12 months, the average cost would be \$2,650 – \$6,000 (33). Traditional braces on the other hand are calculated with more specifics. For example, a big part in the cost is what type of braces the patient needs, traditional metal braces average cost is \$3000-\$5000, for ceramic braces average cost is \$4000-\$7000, and for lingual braces the average cost is \$5000-\$8000. In addition, home care throughout orthodontic treatment is essential to the success of treatment. In this study by Peter H. Buschang et al (27), he compared the time efficiency of aligners and

braces. Aligners had 14 total appointments while braces had 19 (27). In the same study, the aligners group averaged one emergency visit while traditional braces averaged 3.5. In another study traditional braces took more chair time compared to aligners. And keeping in mind during this chair time the material costs also increase. On the other hand, total doctor time was more in aligners than traditional braces. As 25% of doctor time was over 33 minutes while traditional braces had 26 minutes. After finishing treatment, the corrections in the teeth were the patients who had aligners(19). Post Treatment after 3 years show how the alignment of teeth were worsening in the patients with clear aligner treatment compared to traditional braces(34).

Conclusion

Clear aligners and traditional braces yield many differences in outcomes. In orthognathic surgery clear aligners demonstrated superior quality of life (Qol) outcome and reduced pain levels compared to braces(23). Regarding time efficiency, clear aligners had quicker results than braces, although after post treatment the alignment of teeth tended to deteriorate in patients who underwent clear aligner treatment (27)(34). In overall hygiene clear aligners had better outcomes with periodontal status, bleeding on probing, and inflammation(19). Traditional braces also promoted white lesions during treatment(23). In treating malocclusions clear aligners tend to yield better results than braces(29). Periodontal status was achieved in both braces and clear aligners, but clear aligners yielded better results. Caries were more prominent in traditional braces compared to clear aligners. Both braces and clear aligners offer effective solutions for orthodontic treatment, although clear aligners have several advantages. Clear aligners provide greater comfort, are less visible and are easier to maintain. Additionally, they result in quicker treatment times with better periodontal status compared to traditional braces. Also complying with patients' satisfied outcome. Therefore, those seeking a convenient and aesthetic orthodontic solution, clear aligners are the preferable choice.

A new method for making clear aligners by 3D printing them offers many benefits.3D printing allows for highly precise aligners, better fit, higher efficacy, and reproducibility.The process includes high-process 3D models creating aligners with a smooth surface finish, which is important for transparency and patient comfort. Companies like Modern Clear utilize advanced 3d printing technology to maintain quality and consistency. To add on, the ability to control the thickness and design of aligners allow for more customized treatment plans, which leads to better orthodontic outcomes(35). A study by Gianluca M. Tartaglia et al shows that 3D printed aligners offer superior accuracy, load resistance, and reduced deformation compared to traditional thermoformed aligners(37). A study by James Grant et al measures the amount of force and movement with 3D printed aligners(38). The study proposes one of the key benefits of 3D printed aligners is increased precision in manufacturing, allowing for better predictability of movement. Also, agrees with the previous study that 3D printed aligners achieved greater accuracy and load resistance(37). Despite the advantages of 3D printing aligners, there is not enough technical and clinical data regarding these aligners(38).

Works Cited

- Hepdarcan, S. S., Yılmaz, R. B. N., & Nalbantgil, D. (2016). Which Orthodontic Wire and Working Sequence Should be Preferred for Alignment Phase? A Review. *Turkish journal of orthodontics*, 29(2), 47–50. <https://doi.org/10.5152/TurkJOrthod.2016.160009>
- AlMogbel A. (2023). Clear Aligner Therapy: Up to date review article. *Journal of orthodontic science*, 12, 37. https://doi.org/10.4103/jos.jos_30_23.
- SmartTrack™ material: The difference is clear. (n.d.). Retrieved from <https://www.diamondbraces.com/invisalign/smarttrack-aligners>
- Suresmile®. Orthodontic Associates. (2022, August 10). <https://orthodonticassoc.com/treatments/suresmile/#:~:text=SureSmile%20creates%20a%203D%20computer,your%20teeth%20and%20their%20roots>
- ClearCorrect: Clear. simple. friendly. ClearCorrect | A lifetime of smiles starts today. (2024, March). <https://www.straumann.com/clearcorrect/us/en/patients.html>
- Archambault, A., Major, T. W., Carey, J. P., Heo, G., Badawi, H., & Major, P. W. (2010). A comparison of torque expression between stainless steel, titanium molybdenum alloy, and copper nickel titanium wires in metallic self-ligating brackets. *The Angle orthodontist*, 80(5), 884–889. <https://doi.org/10.2319/102809-604.1>
- Hepdarcan, S. S., Yılmaz, R. B. N., & Nalbantgil, D. (2016). Which Orthodontic Wire and Working Sequence Should be Preferred for Alignment Phase? A Review. *Turkish journal of orthodontics*, 29(2), 47–50. <https://doi.org/10.5152/TurkJOrthod.2016.160009>
- Beddis, H. P., Durey, K., Alhilou, A., & Chan, M. F. (2014). The restorative management of the deep overbite. *British dental journal*, 217(9), 509–515. <https://doi.org/10.1038/sj.bdj.2014.953>
- Watkinson, S., Harrison, J. E., Furness, S., & Worthington, H. V. (2013). Orthodontic treatment for prominent lower front teeth (Class III malocclusion) in children. *The Cochrane database of systematic reviews*, (9), CD003451. <https://doi.org/10.1002/14651858.CD003451.pub2>.
- Belmonte, F. M., Macedo, C. R., Day, P. F., Saconato, H., & Fernandes Moça Trevisani, V. (2013). Interventions for treating traumatised permanent front teeth: luxated (dislodged) teeth. *The Cochrane database of systematic reviews*, 2013(4), CD006203. <https://doi.org/10.1002/14651858.CD006203.pub2>
- Kanas, R. J., Carapezza, L., & Kanas, S. J. (2008). Treatment classification of Class III malocclusion. *The Journal of clinical pediatric dentistry*, 33(2), 175–185. <https://doi.org/10.17796/jcpd.33.2.431877341u182416>
- Willmot D. Orthodontic Treatment and the Compromised Periodontal Patient. *Eur. J. Dent.* 2008;2:1–2. doi: 10.1055/s-0039-1697352.
- Weiss, R. O., 2nd, Ong, A. A., Reddy, L. V., Bahmanyar, S., Vincent, A. G., & Ducic, Y. (2021). Orthognathic Surgery-LeFort I Osteotomy. *Facial plastic surgery : FPS*, 37(6), 703–708. <https://doi.org/10.1055/s-0041-1735308>
- Asiry M. A. (2018). Biological aspects of orthodontic tooth movement: A review of literature. *Saudi journal of biological sciences*, 25(6), 1027–1032. <https://doi.org/10.1016/j.sjbs.2018.03.008>.

- Liu, Y., Li, C. X., Nie, J., Mi, C. B., & Li, Y. M. (2023). Interactions between Orthodontic Treatment and Gingival Tissue. *The Chinese journal of dental research*, 26(1), 11–18. <https://doi.org/10.3290/j.cjdr.b3978667>
- Golež, A., Ovsenik, M., & Cankar, K. (2023). The effect of orthodontic tooth movement on the sensitivity of dental pulp: A systematic review and meta-analysis. *Heliyon*, 9(4), e14621. <https://doi.org/10.1016/j.heliyon.2023.e14621>
- Meeran N. A. (2013). Cellular response within the periodontal ligament on application of orthodontic forces. *Journal of Indian Society of Periodontology*, 17(1), 16–20. <https://doi.org/10.4103/0972-124X.107468>
- Verna, C., Dalstra, M., & Melsen, B. (2000). The rate and the type of orthodontic tooth movement is influenced by bone turnover in a rat model. *European journal of orthodontics*, 22(4), 343–352. <https://doi.org/10.1093/ejo/22.4.343>
- Tamer, İ., Öztaş, E., & Marşan, G. (2019). Orthodontic Treatment with Clear Aligners and The Scientific Reality Behind Their Marketing: A Literature Review. *Turkish journal of orthodontics*, 32(4), 241–246. <https://doi.org/10.5152/TurkJOrthod.2019.18083>
- Higa, R. H., Semenara, N. T., Henriques, J. F., Janson, G., Sathler, R., & Fernandes, T. M. (2016). Evaluation of force released by deflection of orthodontic wires in conventional and self-ligating brackets. *Dental press journal of orthodontics*, 21(6), 91–97. <https://doi.org/10.1590/2177-6709.21.6.091-097.oar>
- Katib, H. S., Hakami, A. M., Albalawei, M., Alhajri, S. A., Alruwaily, M. S., Almusallam, M. I., & Alqahtani, G. H. (2024). Stability and Success of Clear Aligners in Orthodontics: A Narrative Review. *Cureus*, 16(1), e52038. <https://doi.org/10.7759/cureus.52038>
- Kapila, S., & Sachdeva, R. (1989). Mechanical properties and clinical applications of orthodontic wires. *American journal of orthodontics and dentofacial orthopedics : official publication of the American Association of Orthodontists, its constituent societies, and the American Board of Orthodontics*, 96(2), 100–109. [https://doi.org/10.1016/0889-5406\(89\)90251-5](https://doi.org/10.1016/0889-5406(89)90251-5)
- de Leyva, P., Eslava, J. M., Hernández-Alfaro, F., & Acero, J. (2023). Orthognathic surgery and aligners. A comparative assessment of periodontal health and quality of life in postsurgical orthodontic treatment with aligners versus traditional fixed appliances: a randomized controlled trial. *Medicina oral, patologia oral y cirugia bucal*, 28(3), e208–e216. <https://doi.org/10.4317/medoral.25555/>
- Cardoso, P. C., Espinosa, D. G., Mecnas, P., Flores-Mir, C., & Normando, D. (2020). Pain level between clear aligners and fixed appliances: a systematic review. *Progress in orthodontics*, 21(1), 3. <https://doi.org/10.1186/s40510-019-0303-z>
- Khoo, E. (2015, December). Accelerated orthodontic tooth movement with clear-aligner therapy . *Accelerated Orthodontic Tooth Movement with Clear Aligner Therapy*
- Alsino, H. I., Hajeer, M. Y., Burhan, A. S., Alkhouri, I., & Darwich, K. (2022). The Effectiveness of Periodontally Accelerated Osteogenic Orthodontics (PAOO) in Accelerating Tooth Movement and Supporting Alveolar Bone Thickness During Orthodontic Treatment: A Systematic Review. *Cureus*, 14(5), e24985. <https://doi.org/10.7759/cureus.24985>

- Buschang, P. H., Shaw, S. G., Ross, M., Crosby, D., & Campbell, P. M. (2013). Comparative time efficiency of aligner therapy and conventional edgewise braces. *The Angle orthodontist*, Advance online publication. <https://doi.org/10.2319/062113-466.1>
- Levrini, L., Mangano, A., Montanari, P., Margherini, S., Caprioglio, A., & Abbate, G. M. (2015). Periodontal health status in patients treated with the Invisalign® system and fixed orthodontic appliances: A 3 months clinical and microbiological evaluation. *European journal of dentistry*, 9(3), 404–410. <https://doi.org/10.4103/1305-7456.163218>
- Kaklamanos, E. G., Makrygiannakis, M. A., & Athanasiou, A. E. (2023). Oral Health-Related Quality of Life throughout Treatment with Clear Aligners in Comparison to Conventional Metal Fixed Orthodontic Appliances: A Systematic Review. *International journal of environmental research and public health*, 20(4), 3537. <https://doi.org/10.3390/ijerph20043537/>
- Li, Q., Du, Y., & Yang, K. (2023). Comparison of pain intensity and impacts on oral health-related quality of life between orthodontic patients treated with clear aligners and fixed appliances: a systematic review and meta-analysis. *BMC oral health*, 23(1), 920. <https://doi.org/10.1186/s12903-023-03681-w>
- Di Spirito, F., D'Ambrosio, F., Cannatà, D., D'Antò, V., Giordano, F., & Martina, S. (2023). Impact of Clear Aligners versus Fixed Appliances on Periodontal Status of Patients Undergoing Orthodontic Treatment: A Systematic Review of Systematic Reviews. *Healthcare (Basel, Switzerland)*, 11(9), 1340. <https://doi.org/10.3390/healthcare11091340>.
- Papadimitriou, A., Mousoulea, S., Gkantidis, N., & Kloukos, D. (2018). Clinical effectiveness of Invisalign® orthodontic treatment: a systematic review. *Progress in orthodontics*, 19(1), 37. <https://doi.org/10.1186/s40510-018-0235-z>
- The truth about invisalign cost - how to save more money. *diamondbraces*. (n.d.-b). <https://www.diamondbraces.com/invisalign/invisalign-cost/>
- Average cost of braces in Washington - Spain orthodontics. *Spain Orthodontics | Shoreline, WA*. (2023, January 31). <https://spainortho.com/blog/average-cost-of-braces-in-washington-state/>
- How modern clear manufactures millions of clear aligners with 3D printing. *Formlabs*. (2023, November). <https://dental.formlabs.com/blog/clear-aligner-production-modern-clear/>
- Tartaglia, G. M., Mapelli, A., Maspero, C., Santaniello, T., Serafin, M., Farronato, M., & Caprioglio, A. (2021). Direct 3D Printing of Clear Orthodontic Aligners: Current State and Future Possibilities. *Materials (Basel, Switzerland)*, 14(7), 1799. <https://doi.org/10.3390/ma14071799>
- Grant, J., Foley, P., Bankhead, B., Miranda, G., Adel, S. M., & Kim, K. B. (2023). Forces and moments generated by 3D direct printed clear aligners of varying labial and lingual thicknesses during lingual movement of maxillary central incisor: an in vitro study. *Progress in orthodontics*, 24(1), 23. <https://doi.org/10.1186/s40510-023-00475-2>
- Park, C., Giap, H. V., Kwon, J. S., Kim, K. H., Choi, S. H., Lee, J. S., & Lee, K. J. (2023). Dimensional accuracy, mechanical property, and optical stability of zirconia orthodontic bracket according to yttria proportions. *Scientific reports*, 13(1), 20418. <https://doi.org/10.1038/s41598-023-47827-w>

Color Psychology in Marketing: Unveiling the Impact of Color on Consumer Behavior and Brand Perception By Olivia Lee

Abstract

This paper aims to investigate the role of color in marketing and consumer behavior by focusing on three key areas: brand logos, product packaging, and consumers' shopping experiences. As color is one of the most prominent visual cues essential in sensory marketing, previous research has delved into its application in the marketing field. A comprehensive review of prior studies illustrates that colors evoke psychological associations that shape the overall satisfaction of the shopping experience, influence product appeal and brand image, and ultimately affect product purchase. This study thus emphasizes the significance of color in optimizing customer experience and suggests future research on the implications of customer age in sensory marketing, the interplay between color and other sensory inputs, as well as the effects of color based on consumers' varying involvement levels.

Introduction

Increased exposure to media and the development of technology has rendered sensory marketing, "marketing that engages consumers' senses and affects their perception, judgment, and behavior," an effective tool for marketers (Krishna, 2012, p. 333). One of the most widely studied sensory inputs for marketing is visual cues, namely colors, as early psychological research indicates that color activates certain concepts and ideas that shape people's behaviors and psychological associations they form with objects.

In the realm of marketing, various companies are effectively utilizing color to express their identity because color is one of the most important elements of branding that shapes customer perceptions and behaviors (Labrecque & Milne, 2012). For instance, different brands utilize color to differentiate themselves from other brands and portray a certain personality: Verizon uses red, T-Mobile uses pink, and AT&T uses blue as their primary brand color.

Furthermore, the use of color in marketing is not restricted to branding: color is widely used to enhance product packaging and shape the shopping experience itself. Product color is often used to enhance product appeal, delineate product categories, and signal the benefits of the products (Scott & Vargas, 2007; Doyle & Bottomley, 2006). Furthermore, digital and physical retail environments can utilize color to shape customer mood, perceived shopping time, and eventually, purchase intentions (Bellizzi & Hite, 1992; Crowley, 1993).

Therefore, this paper delves into the multifaceted impact of color in branding, product packaging, and shaping customers' shopping experience. Through a close scrutiny of research in this field, this paper aims to highlight how color can become a powerful tool for companies to elevate product perception, enhance the shopping experience, and foster a deeper connection with consumers.

The Role of Color in Brand Logos

Psychological research on color and the different associations it evokes laid the foundation for studies regarding the marketing implications of color. According to McClelland (1998), our memory consists of a web of nodes and links: the nodes symbolize ideas or concepts and the links serve as the pathways that connect the nodes. When an external stimulus triggers one of the nodes, other nodes connected to the triggered node through links are also activated, creating an association between the two nodes (McClelland, 1988). For instance, the sight of “fire” would activate the corresponding node and also stimulate the connected node “danger”, leading an individual to avoid the fire. The links may be strengthened and updated by repetitive associations formed by external feedback and stimuli (Janiszewski & Van Osselaer, 2000).

This network of memory is used to explore the associations between various psychological traits and colors. Early studies that delved into color symbolism indicate color serves as an external stimulus that activates other nodes representing ideals or even personality traits. For instance, the color black evokes feelings of power, sophistication, and death; red is associated with love and passion; and blue is linked with trust (Jacob et al., 1991). A comprehensive list of the psychological implications of color is presented in Table 1. More recent studies in color psychology have divided color into two categories: warm colors (e.g. red, orange) and cool colors (e.g. blue, green) (Hynes, 2009). Research has shown that cool colors are associated with calmness and relaxation while warm colors stimulate the viewers, as they are associated with excitement and anger (Elliot & Maier, 2012).

Table 1. Color and its Psychological Associations

Color	Psychological Implication	Literature
Red	Love, Passion, Active, Stimulating, Warm, Lively	(Elliot & Maier, 2014; Bellizi & Hite, 1992)
Orange	Sociable, Energetic, Exciting, Arousing, Extroverted	(Mahnke, 1996)
Yellow	Happy, Optimistic, Cheerful, Friendly, Exciting	(Clarke & Costall, 2008; Jacobs et al., 1991)
Green	Calming, Healthy, Clean, Fresh, Natural	(Elliot & Maier, 2014; Clarke & Costall, 2008)
Blue	Trust, Cold, Wise, Dependable, Competent	(Bellizi & Hite, 1992; Clarke & Costall, 2008; Mahnke, 1996; Kaya & Epps, 2004)
Purple	Luxurious, Elegant, Dignity, Power	(Jacobs et al., 1991; Labrecque & Milne, 2012)
Black	Sophisticated, Strong, Majestic, Death, Power	(Jacobs et al., 1991; Mahnke, 1996)

Through associative priming, the psychological process in which two nodes become linked, a brand becomes primed with a color along with the identity that the color embodies

(Baxter, 2010). Eventually, brand colors act as effective visual cues that signify brand personality and allow customers to differentiate brands (Kapferer, 1995; Zaichkowsky, 2010). Color is therefore not merely a tool that provides aesthetic pleasure, but also a powerful element that helps customers form emotional bonds with the brands (Goldman, 2005). For instance, IBM utilizes the color blue to communicate competence and trust, indicating that color can be used to convey a brand's personality and build brand trust, and Coca-Cola uses its iconic red color to convey energy and enthusiasm.

In fact, Suriadi et al. (2022) highlight that colors evoke affective and unconscious reactions, which subsequently affect product perception and purchase intent. Empirical research shows that aligning the perceived benefits of products with the brand color amplifies the perceived benefits for both functional and sensory-social products (Doyle & Bottomley, 2006). For example, using blue – a color commonly associated with functionality – for brands selling utilitarian products such as kitchen rolls, and red – a color associated with sensory and social appeal – for chocolates or ice cream brands, significantly augmented the perceived benefits promoted by the brand (Doyle & Bottomley, 2006). Likewise, Labrecque and Milne (2012), state that when a brand personality profile matches the color of a product's package, customers' intentions to purchase increases (Labrecque & Milne, 2012). Furthermore, customers who view a colorful brand logo were likely to infer that the brand offers more diverse products compared to less colorful logos (Song et al., 2022).

In summary, brand color is critical in promoting a coherent brand image through associative priming. Therefore, strategic usage of brand colors will enable marketers to shape customer behavior, enhance product perception, and render their brands more memorable to customers.

The Role of Color in Product Packaging

Color and Product Identification

Color is a visual trait that is widely used by customers to differentiate product categories. Foundational research conducted by Schiller (1935) indicates that female customers associated the colors green and yellow with utilitarian goods such as soap and breakfast goods, as these colors stimulated feelings of functionality and cleanliness. On the other hand, the colors silver and black evoked feelings of dignity and classiness and therefore were associated with luxury products (Schiller, 1935). Furthermore, the color white, specifically white bags, was used to identify the product category of flour and vivid colors were associated with children's products (Garber et al., 2000; Scott & Vargas, 2007). Product color also influences the degree to which a certain product grabs the consumer's attention in a shopping experience. Specifically, products that utilized colors that contrasted with typical product-category-color expectations were more likely to grab the attention of the consumers and eventually increase their shopping time (Garber, Hyatt, & Starr, 2003).

The Influence of Color on Product Perception

Beyond product category identification, the packaging color of a product has a significant influence on the perceived benefits, country of origin, and target segment of products. For instance, research by Garber and Hyatt (2003) indicates that the color of dishwashing liquid packaging affects perceived benefits: yellow suggests a lemon scent, green is associated with being unscented, and blue suggests that the product is effective at removing grease. Beyond functional goods, colors used in product packaging also influence the perception of drugs. Specifically, medicine that is packaged in a cover using “warm colors” such as red or yellow was perceived to contain drugs that were potent compared to those that were packaged in “cool colors” such as blue or green (Roullet & Droulers, 2005).

The saturation of the color also significantly impacts the perceived benefits of a particular product. For instance, Products that are vivid and colorful (colors with high saturation) are associated with children’s products and therefore, are thought to be of low quality. On the other hand, products that are packaged using a muted palette (low saturation) are considered to be of higher quality due to their association with classiness (Scott & Vargas, 2007).

Furthermore, cultural beliefs and ideals constructed around specific colors also render the color scheme of product packages a powerful tool that affects people’s product preferences. According to Block & Kramer (2009), a survey conducted on Taiwanese customers indicated that Taiwanese consumers showed a greater affinity towards red, as it is considered to be a lucky color in Taiwanese culture. Specifically, participants showed greater product expectations and purchasing intentions for red rice cookers compared to green rice cookers (Block & Kramer, 2009).

As the literature suggests that the color of a product’s package taps into the cultural associations and psychological norms associated with the color, utilizing colors that are closely related to the benefits of the product in the product development or the packaging stage will be helpful for companies to increase customer’s willingness to pay.

The Role of Color in the Shopping Experience

The role of color in shaping customers’ experience has been investigated as a crucial part of atmospherics, which refers to “the effort to design buying environments to produce specific emotional effects in the buyer that enhance his purchase probability” (Kotler, 1973, p. 50). Through the study of atmospherics, scholars have investigated various ways in which tactile, olfactory, and visual elements of the shopping environment could affect purchasing behavior.

Various studies have been conducted on the role of the background color of physical stores in shaping consumer behavior. Early studies in this field have shown that warm vs. cool colors influence consumers’ psychological state. Specifically, warm colors were found to arouse customers while customers exposed to cool colors reported feeling more relaxed and pleasant (Crowley, 1993; Bellizzi & Hite, 1992). The psychological implications of color led to differences in consumer behavior as well. Blue retail environments led to longer shopping time and more purchases (Bellizzi & Hite, 1992). In fact, this preference is also evident in the digital

space, as people showed fewer postponed purchases when they saw a product on television with a blue background compared to a red background (Bellizzi & Hite, 1992). Later studies, however, revealed that the behavioral differences instigated by color were dependent on consumers' cultural backgrounds: French-Canadians gave higher quality ratings to products in malls with a warm color décor while Anglo-Canadians gave higher quality ratings to products exhibited in a mall with cool color décor (Chebat & Morrin, 2007).

The interaction between color and other visual elements of the shopping atmosphere such as lighting has also been investigated. For instance, a study conducted by Babin et al. (2003) indicates that the effect of the color blue and orange walls in shopping-oriented retail stores is moderated by the effects of lighting (soft vs. bright). The results indicate that overall, blue led to better affective evaluations, higher levels of excitement, higher rates of perceived fairness, store patronage, and purchase intentions when it was shown with bright lights (Babin et al., 2003). However, when soft lighting was employed, participants showed greater affective evaluations and perceived fairness for orange walls. In fact, under bright lighting, there was no significant difference in excitement and purchase intentions between orange and blue wall conditions (Babin et al., 2003).

Conclusion & Future Studies

This paper investigated the use of color psychology in the field of marketing and consumer behavior. Previous literature indicates that color psychology has been applied to primarily three fields of marketing: brand logos, atmospherics, and product packaging. A close scrutiny of previous studies indicates that color psychology shapes brand personality, fosters product and brand recognition, and signals product category, eventually influencing customers' purchasing intent (Labrecque & Milne, 2012; Doyle & Bottomley, 2006; Song et al., 2022). Furthermore, the colors of the background in the digital space or the physical retail environments also impact customers' overall shopping experience (Chebat & Morrin, 2007; Babin et al., 2003). Overall, previous research has shown that color, one of the most prominent visual cues, has a profound effect on sensory marketing.

However, many topics require further research. For instance, the psychological influence of color may differ depending on the generation of the customers. With the development of technology and media, younger generations have greater exposure to diverse sensory experiences (Dusek, 2021). Therefore, the associations younger generations form with various colors can be different from those of the older generations. In fact, research indicates that the threshold sensitivity and temporal acuity differ based on age, further hinting that a single color could have multiple psychological and behavioral implications depending on one's stage of life (Humes et al., 2009). Furthermore, previous research has also delved into the interaction between other sensory inputs such as temperature and color (Porter & Mikellides, 1976). Therefore, the ways in which various sensory inputs such as music, touch, and smell interact with the perception of color to influence customers' information processing and consumption behaviors is also a promising field that requires further research.

Finally, future research should also shed light on the differing effects of engagement levels on the effectiveness of color marketing. Previous research on the Elaboration Likelihood Model (ELM) suggests that people process information through two routes: the central route and the peripheral route (Zander, 2006). The central route is used when people process information carefully in a high-motivation situation (Zander, 2006). On the other hand, people's peripheral route processing is activated when they are not deeply engaged in the situation and processing the information without much cognitive effort (Zander, 2006). As previous research indicates that people are more likely to be influenced by heuristic cues and sensory inputs when they are processing information through the peripheral route, the color of the retail environment or the brand may have a greater influence on customers when they are purchasing inexpensive products that require less cognitive effort compared to expensive products. Therefore, the ways in which the marketing implications of colors vary according to a customer's level of engagement should be investigated to shed light on how color can be used in real-life marketing situations (Meyers-Levy & Peracchio, 1995).

Works Cited

- Babin, B. J., Hardesty, D. M., & Suter, T. A. (2003). Color and shopping intentions: The intervening effect of price fairness and perceived affect. *Journal of Business Research*, 56(7), 541-551. [https://doi.org/10.1016/S0148-2963\(01\)00246-6](https://doi.org/10.1016/S0148-2963(01)00246-6)
- Baxter, S., Ilicic, J., & Kulczynski, A. (2017). You see Froot, you think fruit: Examining the effectiveness of pseudohomophone priming. *European Journal of Marketing*, 51(5/6), 885-902. <https://doi.org/10.1108/EJM-01-2016-0038>
- Bellizzi, J. A., & Hite, R. E. (1992). Environmental color, consumer feelings, and purchase likelihood. *Psychology & Marketing*, 9(5), 347-363. <https://doi.org/10.1002/mar.4220090502>
- Block, L., & Kramer, T. (2009). The effect of superstitious beliefs on performance expectations. *Journal of the Academy of Marketing Science*, 37(2), 161-169. <https://doi.org/10.1007/s11747-008-0116-y>
- Chebat, J. C., & Morrin, M. (2007). Colors and cultures: Exploring the effects of mall décor on consumer perceptions. *Journal of Business Research*, 60(3), 189-196. <https://doi.org/10.1016/j.jbusres.2006.11.003>
- Clarke, T., & Costall, A. (2008). The emotional connotations of color: A qualitative investigation. *Color Research & Application*, 33(5), 406-410. <https://doi.org/10.1002/col.20435>
- Crowley, A. E. (1993). The two-dimensional impact of color on shopping. *Marketing Letters*, 4(1), 59-69. <https://doi.org/10.1007/BF00994188>
- Doyle, H., & Bottomley, P. A. (2006). The interactive effects of colors and products on perceptions of brand logo appropriateness. *Marketing Theory*, 6(1), 63-83. <https://doi.org/10.1177/147059310606126>
- Dusek, B. (2021, July 30). Generational media consumption trends. Colorado State University. <https://social.colostate.edu/trends/generational-media-consumption-trends/>
- Elliot, A. J., Maier, M. A., Binser, M. J., Friedman, R., & Pekrun, R. (2009). The effect of red on avoidance behavior in achievement contexts. *Personality and Social Psychology Bulletin*, 35(3), 365-375. <https://doi.org/10.1177/0146167208328330>
- Elliot, A. J., & Maier, M. A. (2014). Color psychology: Effects of perceiving color on psychological functioning in humans. *Annual Review of Psychology*, 65(1), 95-120. <https://doi.org/10.1146/annurev-psych-010213-115035>
- Garber, L. L., Jr., Burke, R. R., & Jones, J. M. (2000). The role of package color in consumer purchase consideration and choice. *Marketing Science Working Paper*, 104, 1-46.
- Garber, L. L., Jr., & Hyatt, E. M. (2003). Color as a tool for visual persuasion. In R. Batra & L. M. Scott (Eds.), *Persuasive imagery: A consumer response perspective* (pp. 313-336). Lawrence Erlbaum Associates.
- Goldman, A. (2005). The aesthetic. In Gaut, B., Lopes, D., Gaut, B., & McIver Lopes, D. (Eds.), *The Routledge companion to aesthetics* (2nd ed., pp. 275-286). Routledge.
- Humes, L.E., Busey, T. A., & Craig, J. C. (2009). The effects of age on sensory thresholds

- and temporal gap detection in hearing, vision, and touch. *Attention, Perception, & Psychophysics*, 71(4), 860-871. <https://doi.org/10.3758/APP.71.4.860>
- Hynes, N. (2009). Colour and meaning in corporate logos: An empirical study. *Journal of Brand Management*, 16, 545-555. <https://doi.org/10.1057/bm.2008.5>
- Jacobs, L., Keown, C., Worthley, R., & Ghymn, K. I. (1991). Cross-cultural color comparisons: Global marketers beware! *International Marketing Review*, 8(3). <https://doi.org/10.1108/02651339110137279>
- Kapferer, J. N. (1995). Brand confusion: Empirical study of a legal concept. *Psychology & Marketing*, 12(6), 551-568. <https://doi.org/10.1002/mar.4220120607>
- Kaya, N., & Epps, H. H. (2004). Relationship between color and emotion: A study of college students. *College Student Journal*, 38(3), 396-405.
- Kotler, P. (1974). Atmospherics as a marketing tool. *Journal of Retailing*, 49(4), 48-64.
- Krishna, A. (2012). An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. *Journal of Consumer Psychology*. 22(3), 332-35. <https://doi.org/10.1016/j.jcps.2011.08.003>
- Krishnan, V., Kellaris, J. J., & Aurand, T. W. (2012). Sonic logos: can sound influence willingness to pay? *Journal of Product & Brand Management/Journal of Product & Brand Management*, 21(4), 275-284. <https://doi.org/10.1108/10610421211246685>
- Labrecque, L. I., & Milne, G. R. (2012). Exciting red and competent blue: The importance of color in marketing. *Journal of the Academy of Marketing Science*, 40(5), 711-727. <https://doi.org/10.1007/s11747-010-0245-y>
- Mahnke, F. (1996). *Color, environment, and human response: An interdisciplinary understanding of color and its use as a beneficial element in the design of the architectural environment*. John Wiley & Sons.
- McClelland, D. C. (1998). Identifying competencies with behavioral-event interviews. *Psychological Science*, 9(5), 331-339. <https://doi.org/10.1111/1467-9280.00065>
- Meyers-Levy, J., & Peracchio, L. (1995). Understanding the effects of color: How the correspondence between available and required resources affect attitudes. *Journal of Consumer Research*, 22(2), 121-138. <https://doi.org/10.1086/209440>
- Porter, T., & Mikellides, B. (1976). *Color for architecture*. Van Nostrand Reinhold
- Roth, M. S., & Romeo, J. B. (1992). Matching product category and country image perceptions: A framework for managing country-of-origin effects. *Journal of International Business Studies*, 23(3), 477-497. <https://doi.org/10.1057/palgrave.jibs.8490276>
- Roulet, B., & Droulers, O. (2005). Pharmaceutical packaging color and drug expectancy. In G. Menon & A. R. Rao (Eds.), *Advances in consumer research* (Vol. 32, pp. 164-171). Association for Consumer Research.
- Schiller, G. (1935). An experimental study of the appropriateness of color and type in advertising. *Journal of Applied Psychology*, 19(6), 652-664. <https://doi.org/10.1037/h0056090>

- Scott, L. M., & Vargas, P. (2007). Writing with pictures: Toward a unifying theory of consumer response to images. *Journal of Consumer Research*, 34(3), 341-356.
<https://doi.org/10.1086/519145>
- Song, J., Xu, F., & Jiang, Y. (2022). The colorful company: Effects of brand logo colorfulness on consumer judgments. *Psychology & Marketing*, 39(8), 1610-1620.
<https://doi.org/10.1002/mar.21674>
- Suriadi, J., Mardiyana, M., & Reza, B. (2022). The concept of color psychology and logos to strengthen brand personality of local products. *Linguistics and Culture Review*, 6(S1), 839-856. <https://doi.org/10.21744/lingcure.v6nS1.2168>
- Van Osselaer, S. M., & Janiszewski, C. (2001). Two ways of learning brand associations. *Journal of Consumer Research*, 28(2), 202-223. <https://doi.org/10.1086/322898>
- Zaichkowsky, J. L. (2010). Strategies for distinctive brands. *Journal of Brand Management*, 17, 548-560. <https://doi.org/10.1057/bm.2010.12>
- Zander, M. F. (2006). Musical influences in advertising: how music modifies first impressions of product endorsers and brands. *Psychology of Music*, 34(4), 465-480.
<https://doi.org/10.1177/0305735606067158>

Rock Climbing as a Therapeutic Approach: Enhancing Cognitive and Motor Skills in Neurological Patients By Aanika Tipirneni

Abstract

Over 1 billion people are affected by neurological diseases, without cures. Due to this, therapeutics become necessary, yet many have limitations when dealing with the many symptoms of neurodegenerative diseases. An individual with these diseases not only faces cognitive disabilities but also contends with motor impairments and the onslaught of negative emotions that come with such a debilitating disease. This is where rock climbing therapy comes into play. Rock climbing engages both the mind and the body, to effectively target and strengthen areas that are affected by neurodegenerative diseases such as Parkinsons, Alzhiemers, and Cerebral Palsy. Case studies of individuals with these diseases who participated in rock climbing interventions, were analyzed to find a positive correlation between the symptoms of neurological diseases and rock climbing. It was shown that rock climbing therapy was able to strengthen cognitive functions, improve motor functions specific to neurological diseases, and reduce depression.

Keywords: Behavioral and Social Sciences, Neuroscience, Neurological Diseases, Rock Climbing, Cognitive Functions, Motor System, Limbic System

Introduction

Rock climbing, a unique and multifaceted sport, requires physical prowess and high-level cognitive skills. This sport has been rapidly gaining popularity since it became an official Olympic sport in the 2020 Olympics and has piqued the interest of researchers and enthusiasts alike. In essence, rock climbing involves scaling rock formations, whether it's outdoors on a towering cliff or indoors in a carefully designed gym. A climber's objective is to vertically navigate through an unknown layout of rocks, or synthetic holds, using their entire body to pull themselves up to the finishing hold. As a result of the way rock climbing is practiced, there is an inherent coordination of the motor, cognitive, and emotional systems. This is important because rock climbing can be a useful tool therapeutically since it can target parts of the brain affected by neurological diseases, such as Parkinson's and Alzheimer's while also exercising the body. A few case studies that explore this topic concluded that there is a positive therapeutic effect of rock climbing for neurodegenerative diseases. This review paper will outline the feasibility of rock climbing as a neurological therapy by focusing on the cerebrum, motor, and limbic systems.

Discussion

Cerebrum

A. Components of the Cerebrum

One of the aspects behind rock climbing therapy is how it improves cognitive functioning. Cognitive functioning essentially means using your brain, and is a broad concept

that encompasses many mental processes, including attention, problem-solving, decision-making, memory, and motor function. These complex processes are orchestrated by different regions in the cerebrum working in union. The cerebrum is split into 4 distinct subsections: the frontal lobe, parietal, temporal, and occipital lobe (de Schotten et al.). The frontal lobe (Fig. 1a), located behind the forehead, is mainly in charge of executive functions such as attention, memory, and mood, making it the key mediator in cognitive functioning (Collins and Koechlin). On the other hand, the parietal lobe (Fig. 1b), located in the middle and back of the brain, focuses on reacting to one's environment, such as interpreting inputs received from the five senses (Jawabri and Sharma). Memory processes, both long and short-term are linked to the temporal lobe (Fig. 1d). It sits on both sides of the brain, near one's temples, and includes the hippocampus, which is an integral part of many memory-related actions. Located in the posterior of the brain, the occipital lobe (Fig. 1c) primarily focuses on vision, by interpreting and processing visual information. It controls cognitive functions such as visuospatial processing and depth perception (Queensland Brain Institute).

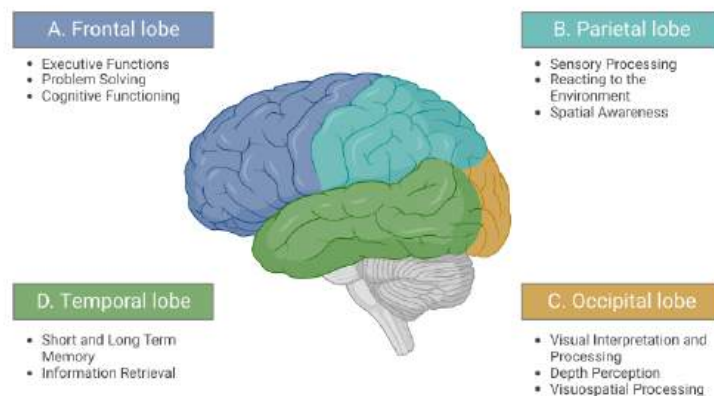


Figure 1: Mapping and Functionality Overview of the Four Lobes of the Brain

- (A) Frontal lobe functions are responsible for higher level cognitive functions.
- (B) Parietal Lobe is in control of interpreting sensory information
- (C) Occipital Lobe in charge of vision processing
- (D) Temporal Lobe in charge of memory

B. Effects of Rock Climbing on the Cerebrum

As mentioned above, rock climbing requires an incredible cognitive demand that is less discussed. The nature of rock climbing necessitates the intricate coordination of planning, problem-solving, visual-spatial skills, and muscle memory (Marczak et al.). For example, when encountering a new route, a rock climber will first visualize and plan out their moves while also using analytical problem-solving skills to determine the most efficient moves needed to

successfully reach the summit. While ascending the wall, a climber would engage their visual-spatial skills to figure out how to grab a certain hold and how much force they need to use to get to the next hold. A climber must also use muscle memory to recall past skills to prevent falling. In the event that they do fall, they will also use muscle memory to remember how to redo the climb and get to the summit. Given how rock climbing has a profound reliance on cognitive thinking, there is a big potential for future applications of it, and similar exercises, in the rehabilitation and mitigation of symptoms.

C. Case Study on Cognitive Functions

To further the claim that rock climbing strengthens cognitive functions, a 2018 study investigated the beneficial effects of rock climbing by comparing 30 climbers versus 30 age and gender-matched non-climbers (Marczak et al.). The study used the Tactual Performance Test, a common test to measure neurocognitive functions, by using three neurocognitive indicators: Location, Time, and Memory. In the three experiments conducted, rock climbers achieved higher memory and location ratios while also having lower time ratios by completing the task faster. These differences suggest that individuals who climb have enhanced cognitive abilities than their counterparts. The higher memory performance illustrates that rock climbing promotes better retention and recollection of information from an individual's surroundings, which is a crucial cognitive skill. It also has positive implications for people with neurocognitive diseases, as memory loss is a common symptom of diseases like Alzheimer's. The lower time ratios due to superior speed in processing tactile input could be accredited to enhanced cognitive processing and problem solving, both crucial to daily living. Additionally, climbers had greater location ratios, indicating higher-level spatial perception and executive functions. This is particularly beneficial for individuals with neurological diseases such as Parkinson's which severely affects orientation. In summary, this study provides compelling evidence that rock climbing improves cognitive functioning across a plethora of areas including, memory, processing speed, and spatial perception. These cognitive improvements have a noteworthy implication for individuals who struggle with cognitive deficiencies due to neurodegenerative diseases.

Supporting this idea, a 2019 study examined the relationship between rock climbing expertise and cognitive functions, focusing on visual and sequence memory (Whitaker et al.). The study tracked 20 climbers of varying experience levels between 20 to 45 years old, 8 of whom were female. In one experiment, climbers were asked to choose 3 climbing routes and preview the climbs for 1 minute each. Using a grayscale image, the climbers then visually marked all the holds that were a part of that climb to assess their visual memory. Next, the climbers were asked to sequence the route by predicting their movements on the ground. They then later recalled their climbing sequences while actually climbing the route. The results indicated that skilled climbers had better visual memory of the climbing holds when marking them on the grayscale image. Furthermore, the experienced climbers predicted and remembered the sequences of their movements with a higher accuracy than the novice climbers. The study also found that these heightened cognitive advantages were domain specific, meaning that they

were a direct result of rock climbing. This underscores the potential of rock climbing as a customized therapy, as it targets specific cognitive areas that the patients need help with. Additionally, the enhanced visual memory observed can directly benefit patients who grapple with memory deficiencies due to their neurological conditions. The improved sequence memory skills can assist in restoring patients' ability to plan and execute movements, a key aspect in many rehabilitation processes. By enhancing cognitive functions, this sport promises improved quality of life and cognitive abilities to those who suffer from neurological diseases.

Drawing a parallel, an experiment asked 16 climbers to replicate a climbing route on a scale model after looking at the climb for 2.5 minutes (Boschker, Bakker, and Michaels). This was repeated with 2 experienced and 2 novice climbers who verbally said what they were thinking during the experiment. Both of the experiments revealed that the experienced climbers were able to remember more information with higher accuracy than the novice and non-climbers. The expert climbers achieved this by focusing on remembering the functional aspects of climbing, meaning they remembered the holds due to their action-related traits in a sequence of moves. On the contrary, inexperienced climbers recalled holds based on their basic structural characteristics such as color or shape, indicating a very surface-level thought process. The observed cognitive advantages, where a climber was able to use a more mindful thought process to remember the holds with more relevance, ease, and accuracy, suggest yet again that climbing stimulates cognitive functions. However, it is important to note that further research is needed due to the relatively small sample size. Although this experiment emphasizes that these advantages are domain-specific, the positive outcomes of rock climbing can extend to individuals with neurocognitive impairments.

Motor System

A. Functionality of the Motor System

Another aspect behind rock climbing therapy is how it improves movement. The motor system is split into two components, broadly categorized as the central and peripheral nervous systems. The central nervous system is comprised of the motor cortex and brainstem, while the peripheral nervous system consists of the spinal cord and lower motor systems (de Carvalho and Swash; see Figure 2.). In the primary motor cortex, we have upper motor neurons which pass information using signals down through the brainstem to the lower motor neurons. The lower motor neurons synapse onto muscle groups that then lead to contractions, causing movement (Zayia and Tadi; Fig. 2b). The motor cortex is located in the frontal lobe and encompasses the primary motor cortex, which sends a majority of the signals that control body movements (Yip and Lui). The brainstem, found in the middle of the brain, connects the cerebrum to the spinal cord which acts as a gateway to the rest of the body (Fig. 2c). It starts at the medulla and extends through an opening in the skull to pass down signals (Basinger and Hogg; Fig. 2d). The lower motor neurons are responsible for transmitting signals to glands or muscles and are first located in the spinal cord before branching out to the whole body (Zayia and Tadi, 2023). Understanding the intricacies of the motor system is relevant in the context of neurological diseases. For

instance in Parkinson's nerve cells start to die, causing important signals that control body movements to not be passed on (Mazzoni, 2012). This leads to a variety of motor symptoms such as delayed movements, decreased dexterity, and rigidity. In contrast, Amyotrophic Lateral Sclerosis (ALS) causes the lower motor neurons located in the spinal cord and brain to degenerate. This causes signals that contain information about movement to never reach the muscles, eventually leading to the brain not being able to create or control movements. Common symptoms of this disease include spasticity, cramps, and twitches in muscles. Additionally, disorders like Cerebral Palsy occur when the motor system doesn't develop correctly in infants causing weakness, lack of coordination, and difficulty with movements in muscles.

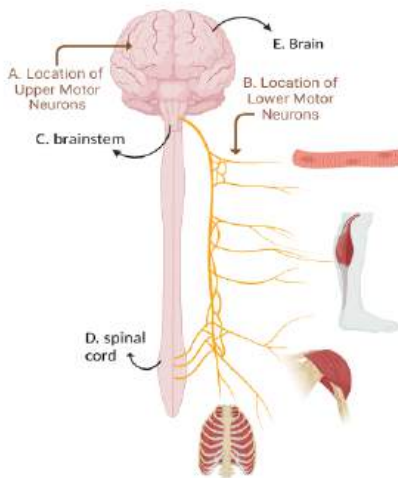


Figure 2: Overview of Motor System and Neural Pathways of Motor Neurons

This figure depicts the interaction between upper motor neurons, in the brain, and lower motor neurons in the body. They pass signals between each other through the brain stem, to cause contractions in the muscles, creating voluntary movement and motor function throughout the body.

B. Effects of Rock Climbing on the Motor System

Not only is rock climbing a mentally engaging sport, it is also physically demanding, therefore improving motor skills and overall physical wellbeing. The dynamic nature of climbing fosters muscular development, through the complex routes that demand muscle endurance as well as strength. Due to the constant pulling and stabilizing one's body needs to do on the wall, abdominal muscles and upper body muscles are mostly affected (Spaeth, 2021). Moreover, the intricacy of movements in rock climbing improves balance, flexibility, and coordination (Li et al, 2018). This physically demanding sport increases $VO_2\text{max}$ (the maximum intake of oxygen), lowers Body Fat Percentages and elevates overall aerobic fitness (Siegel and Fryer, 2017). In essence, the multifaceted nature of rock climbing not only fosters mental intellect but also a well-rounded improvement of physical capabilities, further illustrating it as a sport with broad benefits that could be used as an avenue of exercise or therapy for people with neurological diseases.

C. Case Studies on Motor System Functionality in Cerebral Palsy and Parkinson's

To further the claim that rock climbing strengthens the motor system, a study examined the effects of a 3-week indoor climbing program for children, ages 11-13, with Cerebral Palsy (CP) alongside their typically developing (TD) peers (Christensen et al, 2017). The primary objective of the study was to see if climbing interested the children so they could evaluate functional benefits and see if they could keep pace with their TD peers. The findings revealed that adolescents with CP were able to complete the three week program and remained active for the majority of the time (16 hours in total), similar to the TD children. Both of the groups climbing abilities significantly improved as well, with the TD groups climbing with increased speed and the CP group being able to climb more volume. Additionally, children with CP improved in specific motor skills such as the Sit-to-Stand test, rate of force development in the least affected hand during a pinch test, and muscular-muscular coherence during a pinch precision test. This study underscores the potential of climbing as an engaging and effective form of physical activity and rehabilitation for children with CP. Being such a diverse sport, climbing offers opportunities for increased peer socialization, motivation for training, and improved motor skills, likely due to increased synchronization between the cerebral cortex and muscles. The results emphasized the clinical significance of climbing as a form of exercise and contributed to our understanding of the potential physiological, psychological, and social benefits of climbing, paving the way for further exploration of the field.

Building upon these findings, a subsequent investigation involving 48 Parkinson's patients (PD) between the ages of 56 and 72 aimed to delve deeper into the therapeutic potential of rock climbing (Langer et al, 2021). In this research endeavor, the patients split up into 2 groups, where one underwent a 12-week climbing course and the other acted as a control and had unsupervised physical training. The study unveiled that rock climbing significantly improved motor symptoms in PD patients with striking improvements in bradykinesia (slowness of movement), rigidity, and tremor. The climbers also experienced an average improvement of 12.9 points on the Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale part III (MDS-UPDRS-III), which was shockingly high compared to the control group improvement of 3.5 points. Climbing also has a unique combination of motor skills such as flexibility, resistance training, balance, and coordination, which offers a comprehensive workout while also contributing to its success in addressing the motor symptoms mentioned. Furthermore, the physical training control group had a stable score on the MDS-UPDRS-III, but didn't experience the same improvements and benefits as the climbing group. Additionally, this study highlights the safety and feasibility of rock climbing as a therapy for PD patients, with many participants interested in continuing climbing after the study. These findings provide robust support for the integration of climbing as a tool in therapeutics, especially for those with Parkinson's disease, to mitigate symptoms, promote motor-skill development, and overall well-being.

In addition to this study, a pre-planned secondary study was conducted to determine if rock climbing could improve the posture of patients diagnosed with Parkinson's (Langer et al,

2023). The study was specifically aimed to investigate the forward-stooped posture, which can contribute to increased discomfort and risk of falls. The results revealed that the climbing group had predicted improvements in the biomechanical marker of axial posture, a specific measurement that assesses the alignment of the spine and upper body. Specifically, the sport climbers demonstrated a decrease in the horizontal distance from the wall to their seventh cervical vertebra after the 12-week study. However, it's important to note that the improved axial posture didn't enhance the patient's overall quality of life and that the unsupervised training group did not undergo the same improvement. The results suggest that climbing's blend of physical skills of muscle stretching, flexibility, strengthening and mental skills, spatial body awareness, and movement planning, was able to cause the improvement in posture. Although the forward-stooped posture is a challenging symptom to treat, this study provides valuable insight that can add to the limited options for managing this aspect of Parkinsons, further supporting the claim that rock climbing should be used as a therapy.

Limbic System

A. Physiology of the Limbic System

The limbic system is a complex network of structures within the brain that regulate emotions, behavior, and memory. These systems are located underneath the cerebral cortex and above the brainstem and include the amygdala, hippocampus, hypothalamus, and thalamus (Queensland Brain Institute ;Torrico and Abdijadid, 2023) (Fig. 3). This system acts as a bridge between higher cognitive functions and primitive behaviors, integrating sensory information with emotional responses to form and consolidate memories. One key component of the limbic system is the amygdala, which plays an important role in our emotional responses (Fig. 3B). The amygdala sends out electrical signals to regulate feelings like stress and anger. When the amygdala is miswired, a lack of feelings can occur, resulting in a sense of fearlessness. The amygdala also attaches emotions to our memory, allowing us to remember our fear of heights during a vacation 5 years ago (AbuHasan et al, 2023). The second key structure in this system is the hippocampus which is the key player in memory processes (Fig. 3A). It plays a critical role in forming new memories, particularly episodic and declarative memories. Furthermore, the hippocampus also contributes to spatial navigation and the regulation of hypothalamic functions (Anand and Dhikav, 2012; Fogwe et al, 2023). A common mental health disorder that affects the limbic system is depression. People with depression often have high amygdala activity in response to negative triggers and lower activity to positive stimuli (Redlich et al, 2018). Whether it's reminiscing about the joys of past birthdays, or the onset of nerve-racking anxiety when public speaking, the limbic system is a system of paramount importance, through which we can comprehend and learn from the experiences of life.

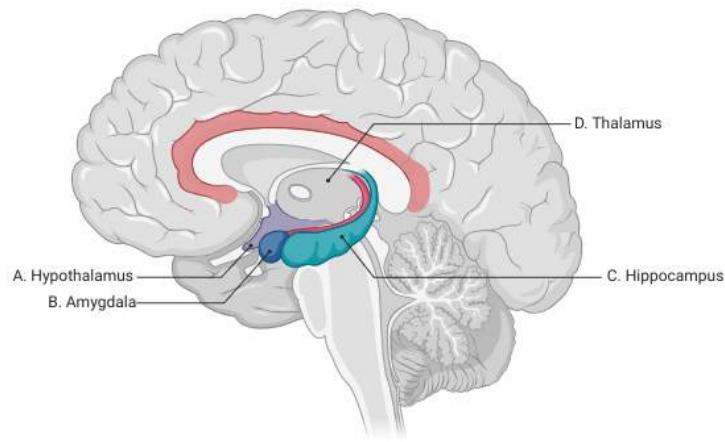


Figure 3: Physiology of the Limbic System

- (A) The Hypothalamus is responsible for basic instincts such as hunger and body temperature.
- (B) The Amygdala is in control of emotion, primarily fear.
- (C) The Hippocampus is responsible for memory and learning
- (D) The Thalamus acts like a relay station. All incoming sensory stimuli is processed by the thalamus, then sent to the brain.

B. Effect of Rock Climbing on the Limbic System

In recent years, there has been a growing amount of research highlighting the beneficial effects of physical activity on mental and emotional health (Gillian 2013). Among the plethora of exercises available, rock climbing stands out as a distinct and dynamic sport that not only engages the body but intricately connects to the complex interplay of emotions found within the limbic system. The limbic system is the brain's emotions control center and is crucial for controlling emotions, memories, and stress reactions. Because of its complexity, rock climbing necessitates a great level of mental focus and garners strong feelings such as pride, fear, and exhilaration (Luttenberger et al, 2015). Recent research on the therapeutic benefits of rock climbing, especially bouldering, a form of rock climbing without a rope, has shown that it can effectively reduce depressive symptoms (Schwarz et al, 2019). Being a sport that seamlessly incorporates cerebral engagement, rock climbing is well-positioned to provide a comprehensive approach to improve mental health.

C. Case Studies on Depression and Parkinsons

As previously mentioned, rock climbing has a significant impact on one's emotions. To further this claim, a 2015 study was done to investigate the therapeutic potential of bouldering for people with clinical depression through a pilot intervention (Luttenberger, 2015). 47 patients with depression were split into 2 groups, with one group acting as the control and the other participating in an eight-week rock climbing program. The eight-week bouldering sessions, resulted in a significant decrease in depressive symptoms (an average reduction of 6 points)

using the Beck Depression Inventory-II Scale (BDI-II). Additionally, this study aligns with pre-existing literature that underlines the effect of physical activity on emotions, with some publications even comparing the benefits to antidepressants (Knapen et al, 2015; Wegner et al, 2014; Adamson, 2015). This investigation proved that Bouldering Therapy, characterized by its emphasis on mental aspects and social interaction, could emerge as a valuable addition for treatments against depression, potentially extending its benefits to individuals with neurological diseases. Individuals faced with neurological diseases, such as Alzheimer's disease, often have many adjunct symptoms such as depression and negative emotions. Sometimes the early stages of these diseases are even presented as depression and about 80% of Alzheimer's disease develop symptoms related to depression on top of having the neurological disease (Baquero and Martín, 2015). Given the intricate correlation between neurological disorders and mental health, exploring innovative therapies such as climbing becomes crucial. Its successful approach to addressing depression, as seen in the study above, and holistic approach to providing physical and mental exercise, make rock climbing a very noteworthy therapy.

In addition, a similar 2022 study measured the differences in physical and mental health through self-reports with individuals diagnosed with Parkinson's (Gassner et al, 2022). Overall, the participants reported a remarkable average of 65% improvement in physical condition, a 59% enhancement in psychological aspects, and a 39% improvement in social aspects. Moreover, the participants expressed a 42% rise in admiration in a social setting and an average of 40% reported that they felt more sociable and outgoing. The participants' adherence to treatment was 100%, their motivation to continue therapeutic rock climbing was 70%, and 96% of the participants would recommend rock climbing to others. Statistically, there were significant improvements in the motor system, measured by the 2 and 10 minute walking test, the functional gait assessment, and the Nine-Hole-Peg tests, with 96% reporting improvements in strength, and 88% reporting improvements in balance and range of motion. Although there were exceptional improvements in physical processes and mental areas such as sociability, a limitation of this study was the relatively small sample size which could slightly inhibit the study's success. That being said, the heightened feelings of outgoingness, sociability, and admiration in social settings, reinforce the therapeutic value of rock climbing beyond its physical benefits. Social engagement, especially for people facing debilitating diseases, can improve their way of life and is shown to combat depression (Son and Sung, 2022). By fostering a positive and encouraging environment, this form of therapy is able to address the challenging emotional and social side effects present in neurological diseases.

Conclusion

In a world with no cures for life-threatening neurological diseases, rock climbing becomes a crucial therapeutic approach. Its impact on the cerebrum, motor system, and limbic system illustrates that this sport can provide the needed exercise for one's body while also engaging cognitive skills and boosting emotional well-being. The cognitive demands of climbing shown through the memory, spatial awareness, and problem solving needed in this illustrate that

it strengthens cognitive functions and can thus mitigate symptoms in neurodegenerative diseases. In the realm of the motor system, rock climbing greatly enhanced motor functionality such as balance and coordination in children with Cerebral Palsy and Parkinson's, while also improving symptoms specific to Parkinson's such as bradykinesia and axial posture. Yet it is not only the physicality that makes rock climbing so unique, it also has transformative effects on one's emotional health. Rock climbing can reduce depression, a common symptom of neurological diseases, and increase sociability and outgoingness in individuals with Parkinson's. Through fostering social engagement, providing a unique avenue of exercise, improving cognitive functioning, and addressing emotional challenges, rock climbing stands as an all-encompassing therapy for the different spectrums of neurological diseases. The research presented not only proves the feasibility of rock climbing as a therapy but also demonstrates its groundbreaking capabilities that could redefine neurological rehabilitation.

Acknowledgements

I would like to thank my mentor Nalini Rao for her guidance throughout this project. I would also like to thank Biorender for allowing me to use their platform. Finally, I would like to express my sincere gratitude to my family for all their support.

Works Cited

- “Amyotrophic Lateral Sclerosis (ALS) | National Institute of Neurological Disorders and Stroke.” National Institute of Neurological Disorders and Stroke, <https://www.ninds.nih.gov/health-information/disorders/amyotrophic-lateral-sclerosis-als>. Accessed 3 December 2023.
- Arroyo, Martin. “Effects of aerobic exercise on depression: a controlled study.” *British Medical Journal*, vol. 291, 2021, p. 109, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1416251/pdf/bmjered00456-0023.pdf>. Accessed 4 January 2024.
- Baquero, Miquel, and Nuria Martín. “Depressive symptoms in neurodegenerative diseases - PMC.” NCBI, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4539408/>. Accessed 4 January 2024.
- Basinger, Hayden. “Neuroanatomy, Brainstem - StatPearls.” NCBI, 4 July 2023, <https://www.ncbi.nlm.nih.gov/books/NBK544297/>. Accessed 3 December 2023.
- de Carvalho, Mamede. “Chapter 2 - Upper and lower motor neuron neurophysiology and motor control.” *Handbook of Clinical Neurology*, vol. 195, Elsevier B.V, 2023, pp. 17-19, <https://www.sciencedirect.com/science/article/pii/B9780323988186000182>. Accessed 3 December 2023.
- de Schotten, Thiebaut. “Physiology, Cerebral Cortex Functions - StatPearls.” NCBI, <https://www.ncbi.nlm.nih.gov/books/NBK538496/>. Accessed 3 December 2023.
- Dorscht, Lisa, et al. “A German climbing study on depression: a bouldering psychotherapeutic group intervention in outpatients compared with state-of-the-art cognitive behavioural group therapy and physical activation – study protocol for a multicentre randomised controlled ...” *BMC Psychiatry*, 17 May 2019, <https://bmcp psychiatry.biomedcentral.com/articles/10.1186/s12888-019-2140-5>. Accessed 4 January 2024.
- “Effect of Exercise on Depressive Symptoms in Adults With Neurologic Disorders: A Systematic Review and Meta-Analysis.” *Archives of Physical Medicine and Rehabilitation*, <https://dpl6hyzg28thp.cloudfront.net/media/1-s2.0-S000399931500009X-main.pdf>. Accessed 4 January 2024.
- “Effects of exercise on anxiety and depression disorders: review of meta- analyses and neurobiological mechanisms.” PubMed, <https://pubmed.ncbi.nlm.nih.gov/24923346/>. Accessed 4 January 2024.
- “Effects of Rock Climbing Exercise on Physical Fitness among College Students: A Review Article and Meta-analysis.” NCBI, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6277736/>. Accessed 3 December 2023.
- “Expertise effects on the perceptual and cognitive tasks of indoor rock climbing.” *2024 Springer Nature*, 30 October 2019, <https://link.springer.com/article/10.3758/s13421-019-00985-7>. Accessed 3 December 2023.
- “Hippocampus in health and disease: An overview.” NCBI, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3548359/>. Accessed 4 January 2024.
- Humphries, Katherine. “Neuroanatomy, Motor Neuron - StatPearls.” NCBI, 24 July 2023, <https://www.ncbi.nlm.nih.gov/books/NBK554616/>. Accessed 3 December 2023.
- “Indoor rock climbing (bouldering) as a new treatment for depression: study design of a waitlist-controlled randomized group pilot study and the first results.” NCBI, 25 August 2015, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4548691/>. Accessed 4 January 2024.
- Jensen, Thor. “To be active through indoor-climbing: an exploratory feasibility study in a group of children with cerebral palsy and typically developing children - BMC Neurology.”

- BMC Neurology, 15 June 2017,
<https://bmcneurol.biomedcentral.com/articles/10.1186/s12883-017-0889-z>. Accessed 3 December 2023.
- Langer, Agnes. “Climb up! Head up! Climbing improves posture in Parkinson's disease. A secondary analysis from a randomized controlled trial.” *Clinical Rehabilitation*, SAGE Publications, 8 May 2023, <https://journals.sagepub.com/doi/full/10.1177/02692155231174990>. Accessed 3 December 2023.
- Langer, Agnes. “A randomized controlled trial on effectiveness and feasibility of sport climbing in Parkinson’s disease.” *npj Parkinsons Dis*, 2024 Springer Nature Limited, 16 June 2021, <https://www.nature.com/articles/s41531-021-00193-8>. Accessed 3 December 2023.
- “The Limbic System in Youth Depression: Brain Structural and Functional Alterations in Adolescent In-patients with Severe Depression.” NCBI, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5770774/>. Accessed 4 January 2024.
- “The limbic system - Queensland Brain Institute - University of Queensland.” Queensland Brain Institute, <https://qbi.uq.edu.au/brain/brain-anatomy/limbic-system>. Accessed 4 January 2024.
- “Lobes of the brain - Queensland Brain Institute - University of Queensland.” Queensland Brain Institute, <https://qbi.uq.edu.au/brain/brain-anatomy/lobes-brain>. Accessed 3 December 2023.
- Marchal, Yannick. “Exercise therapy improves both mental and physical health in patients with major depression.” PubMed, <https://pubmed.ncbi.nlm.nih.gov/25342564/>. Accessed 4 January 2024.
- Mead, Gillian E. “Exercise for depression.” PubMed, 12 September 2013, <https://pubmed.ncbi.nlm.nih.gov/24026850/>. Accessed 4 January 2024.
- Michaels, Claire F. “(PDF) Memory for the Functional Characteristics of Climbing Walls: Perceiving Affordances.” ResearchGate, https://www.researchgate.net/publication/11483105_Memory_for_the_Functional_Characteristics_of_Climbing_Walls_Perceiving_Affordances. Accessed 3 December 2023.
- “Motor Control Abnormalities in Parkinson's Disease - PMC.” NCBI, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3367543/>. Accessed 3 December 2023.
- “Neuroanatomy, Amygdala - StatPearls.” NCBI, 17 July 2023, <https://www.ncbi.nlm.nih.gov/books/NBK537102/>. Accessed 4 January 2024.
- “Neuroanatomy, Hippocampus - StatPearls.” NCBI, 20 July 2023, <https://www.ncbi.nlm.nih.gov/books/NBK482171/>. Accessed 4 January 2024.
- “Neuroanatomy, Limbic System - StatPearls.” NCBI, 17 July 2023, <https://www.ncbi.nlm.nih.gov/books/NBK538491/>. Accessed 4 January 2024.
- “Neurocognitive Functioning of Sport Climbers - PMC.” NCBI, 31 December 2018, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6341964/>. Accessed 3 December 2023.
- “Physiology, Motor Cortical - StatPearls.” NCBI, <https://www.ncbi.nlm.nih.gov/books/NBK542188/>. Accessed 3 December 2023.
- “Reasoning, Learning, and Creativity: Frontal Lobe Function and Human Decision-Making.” NCBI, 27 March 2012, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3313946/>. Accessed 3 December 2023.
- Schwarz, Laura. “Long-term effects of bouldering psychotherapy on depression: benefits can be maintained across a 12-month follow-up.” NCBI, 7 December 2019, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6911955/>. Accessed 4 January 2024.

Siegel, Shannon R., and Simon M. Fryer. "Rock Climbing for Promoting Physical Activity in Youth." NCBI, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6125082/>. Accessed 3 December 2023.

Spaeth, Jonathan. "How do bouldering and rock climbing change your body?" Elevated Adventurer, 17 August 2021, <https://elevatedadventurer.com/how-do-bouldering-and-rock-climbing-change-your-body/>. Accessed 3 December 2023.

Female Empowerment in the Context of Rousseau's French Enlightenment **By Avery Ertman**

As the concept of the modern woman has evolved throughout history, how have societal forces influenced the metrics that define her? While art gives us a window into the material culture and fashion trends that have influenced a woman's status over time, it can also offer deeper insight into the changing landscape of social and cultural movements that have affected women's personal ideologies throughout history. The painting titled, *Marie Joséphine Charlotte du Val d'Ognes* (1801) painted by Marie Denise Villers reveals cultural shifts impacting the role of women in French society at the end of the 18th century. The painting features a portrait of du Val d'Ognes, sketching in a front-facing pose at a studio, portraying her with the authority of an accepted artist. Before the French Enlightenment, it was rare for a woman to have a career in the arts, and even more unusual for a woman to paint another woman in the act of being an artist (Auricchio & Getty). This painting brings forth the emerging concept of self-determination in contemporary French culture. Villers's choice of painting a female artist was revolutionary, yet it was embraced, indicating the changing values and gender roles in France. This progressive new image of a woman forces the viewer to reconcile the possibility that women could have a life and career of self-determination, offering us insight into the impact the Enlightenment had on French social structures.



Villers, Marie Denise. *Marie Joséphine Charlotte du Val d'Ognes*, 1801, Metropolitan Museum of Art, New York

The Enlightenment was a period of cultural expansion in 17th and 18th-century Europe where thinkers questioned the notions of human rights, class structure, and liberalism. (Fermon) Through debates and conversations on these topics, philosophers sought to find non-religious rationales to reform society. One of the most prominent philosophers to influence culture during this time was Jean-Jacques Rousseau. Scholars collectively have noted Rousseau for his attention to the concept of equality. One notable scholar, J. Clark Murray, emphasizes that Rousseau believed men must return to the state of simplicity and freedom, the natural state, and emancipate themselves from the artificial creation of a hierarchical structure (Murray). Rousseau believed

that a hierarchical class structure causes unnatural and unnecessary divisions in society that prevent progress. However, despite Rousseau's reputation for promoting societal equality, he was notorious for his contradictory beliefs on gender equality. Rousseau believed that the natural state for a woman is in subordination to men (Fermon). In the modern age, most scholars condemn this belief and recognize that Rousseau's theories result in the disenfranchisement of women, but in the 18th century, his ideas were widely supported. However, even though Rousseau's influence over culture suppressed the voices of women in the 18th century, some women still found ways to empower themselves in the role of caretakers and homemakers by using the power of art.

Throughout history, visual culture has consistently been a fundamental element of French society. Portraiture has been used as a political tool to build the superiority of elites and the French royal bloodline. Starting in the Renaissance period, the French royal family began contracting artists to paint portraits to secure their legacies in history. Monarchs such as Louis XIV used portraiture to glorify their reigns (Getty Museum). This tradition continued into the Enlightenment when the French were developing some of the most sought-after portrait painters of the 18th century in the Royal Academy of Painting and Sculpture. These French artists painted and sculpted royal families across Europe and established themselves as the most sought-after and prominent artists of the time (Rosenfeld). Almost all of these prominent artists were men. Women were not recognized as artists during this time, partly because of the shared societal belief of their innate subordination to men, but also because women were barred from attending the Royal Academy of Painting and Sculpture.

Women were not allowed to attend the Academy until the late 18th century and even then there was a strict four-woman quota on attendees (Auricchio & Getty). Women therefore lacked the education and support given to male artists, placing them at a disadvantage in an arts-related career. To make it even more difficult, women were not allowed to paint nudes, even if they wanted to self-study. It was impossible for a woman to break this precedent, as models would not accept a position to pose for a woman painter. Talent alone was not enough for a woman to become an artist, she also had to have the courage to reject the status quo. However, the choice for a woman to commission artwork did not require the same level of societal disobedience. With substantial funds and support from their families, women could commission art for personal spaces in their estate, or curate art for a Salon. French Salons were established in the 17th century as a place for patrons to discuss art, literature, and education. The walls were furnished with art for enthusiasts to discuss, and the Salons helped up-and-coming artists display their work (Rosenfeld). Historians differ in describing the level of influence women had over the happenings in French Salons, but it is clear that elite women played a role in commissioning art for Salons during the Enlightenment, indicating that women were beginning to enter the art space in France. This brings me to ask the question, how did Rousseau's Enlightenment ideals shape the culture of elite women's participation in the arts for self-empowerment?

Women patrons

With the start of the Enlightenment and the end of the Old Regime, society witnessed a massive shift towards focusing on the individual rather than the collective. (Blacker) This idea of self-determination brought forth an opportunity for elite women to commission art for the sake of personal gain. In the 18th century, the ability to commission work relied on the ability to produce funds, hence the reason why men were the primary patrons of the arts. For men, the idea of commissioning artwork to improve social status was not a new concept. One key example of this is the portrait, *King Louis XIV (1701)* commissioned by the King himself and painted by Hyacinthe Rigaud. The painting embellishes the royalty and power of the King by placing him in the surroundings of objects of power such as a gold scepter, crown, and sword. The color palette also indicates wealth by highlighting deep red surroundings and blue garb, dyes, and fabrics in a quantity that no regular French citizen could afford. Additionally, Rigaud positions the King in an authoritative stance, threatening the viewer of the piece. Through this portrait, the Monarchy is cultivating a powerful image of the King in order to prove authority over French citizens. This is just one example of portraiture with a motive to cultivate a certain character. With the movement towards individualism, away from a biblical-centered society, a space began to emerge for women to take part in the art world. With that notion, elite aristocratic women used portraiture as a method to portray themselves as powerful within the domestic sphere.



Rigaud, Hyacinthe. Portrait of King Louis XIV, 1701, Musée de Louvre, Paris

One of the most prominent examples of an aristocratic woman who used portraiture to cultivate her image is Marie Antoinette. Marie Antoinette was infamously known for her poor reputation and was the scapegoat for much of France's financial instability before the French Revolution. She was accused of promiscuity through public slander. Pamphlets with degrading images were passed out around France, depicting her having sexual relations with both men and women in her court, much of which is false (Kavanagh 61–78). Antoinette's image was dragged down in a period of financial hardship for the country because of her expenditures on fashion (Kavanagh 61–78). According to Lauren De Angelis, Antoinette found solace through fashion and frivolous pursuits as a method for coping with living with an unaffectionate husband and being separated

from her family. Antoinette spent large sums of money on ornate dresses in a time of financial hardship for the country, and the citizens expressed their disapproval through continued public slander (De Angelis). It would be hard to attribute Antoinette's affection for fashion to the country's financial ruin, but the public was so distressed by her opulent lifestyle that she became a scapegoat for the country's problems.

Antoinette turned to art to challenge the public slander and turmoil of late 18th-century France. Between the years of 1783 and 1787 Antoinette commissioned multiple portraits in response to the criticism of her sexuality and lavish lifestyle. The first of the two paintings, both painted by Élisabeth Louise Vigée LeBrun, is titled *Marie Antoinette in a Chemise Dress* (1783, *Kunsthistorisches Museum, Vienna, Gemäldegalerie*). The portrait is simplistic in comparison to earlier works patronaged by Antoinette. One of these earlier works, *Marie Antoinette in a Court Dress* (1778, *Kunsthistorisches Museum, Vienna, Gemäldegalerie*) is a full-body portrait that portrays Antoinette in a large, luxurious layered silk dress. The dress Antoinette wears in this portrait is cream and adorned with layers of fabric, bows, gold tassels, and gold trim. The dress is outrageous and impractical, representing Antoinette's lavish and fashion-oriented persona during this time (Wang). The painting she commissioned five years later, *Marie Antoinette en Chemise* (1783, Hessische Hausstiftung, Kronberg), is a stark divergence from the previous portrait. This latter portrait is also painted by Le Brun but features Antoinette in a simple loose-fitting white muslin gown. According to Heidi Strobel, the garment is a casual robe meant for strolling within the private gardens of Versailles, specifically le Petit Trianon. Strobel argues that this artistic choice forces the viewer to reconcile with the feminine side of Antoinette. It brings the viewer into the private sphere of Antoinette's life, which is in contrast to her lavish public sphere identity (Strobel). This portrait was displayed in the Salon, a very public location, which implies that the purpose of the commission was to influence public opinion about her persona.



Vigée LeBrun, Élisabeth Louise. Marie Antoinette in a Chemise dress, 1783, Kunsthistorisches Museum, Vienna, Gemäldegalerie

This choice to bring the public into Antoinette's private sphere is a direct call to Rousseau's belief that women in society are the icons of the private sphere. In a

non-religious-centered society, Rousseau believed that passion and sexuality would be the downfall of the peaceful institutions of society, so he employed the idea of separating the public and private spheres. The private sphere, which he deemed women should inhabit, would be a place of sanctuary uninterrupted by the unholiness of men. In his vision the private sphere equates to the home and the public sphere is the rest of society. When translating this idea to values, the private sphere would be a place of femininity, and the public sphere would be a place of passion and power (Fermon).

In a time when Antoinette was trying to reconcile with public disarray and distrust, she used art to appear as if she embodied traditional values of femininity. In cultivating an image where she appeared to comply with Rousseau's ideals of social structure, she used portraiture to convince the French citizens that she was not only in power but also in touch with her feminine side and therefore a moral woman. Marie Antoinette attempted to reject some of the power that she held as a leader and reclaim the power she held as a woman to better align her image with the societal conventions and beliefs of this time. To further her attempted retreat into a more feminine space, Antoinette commissioned a second portrait titled Marie Antoinette with her Children (1787, Musée National des Châteaux de Versailles et de Trianon). This portrait is once again a full-body portrait however, the focus of the portrait rests on Antoinette's three children. According to Junyan Wang, The latter portrait has more of a somber and serious undertone than the former, Marie Antoinette in a Chemise Dress (1783, Hessische Hausstiftung, Kronberg). Wang claims that through this portrait Antoinette aimed to demonstrate her vulnerability and relatability to the other mothers of France. (Wang) Antoinette used art to show that even though she was a ruler, she was also an emotional and endearing mother.

Both paintings highlight the notion that Antoinette embraced the traditional feminine role in an attempt to soften the public's perception of her. (Hyde et al.) Ironically, although it appears that Antoinette was letting go of power, the act of cultivating her public image was in fact empowering. She chose how she wanted to be perceived by using portraiture rather than letting the public dictate her reputation.

Different from Marie Antoinette in many respects, Marie Adélaïde lived in the shadows of the royal family due to the nature of her being a princess, which contrasted with Antoinette's very public career as the Queen of France. Marie Adélaïde, house of Bourbon, was born in 1732 as the fourth daughter of King Louis XV and was a contemporary of Marie Antoinette. From the time Adélaïde was a child, Marie Adélaïde's parents used art to depict her as desirable for marriage. (Hyde et al.) One of these portraits in particular was the painting by John Nattier, *Marie Adélaïde of France as Diana (1745, Galleria degli Uffizi, Florence)*. The portrait depicts thirteen-year-old Marie Adélaïde in furs and silks to highlight her wealth and elegance. Nattier then paints Adélaïde with a bare shoulder and breast to imply to the viewer a level of seductiveness. A male looking for a wife would look for these features: wealth and beauty. Depicting Adélaïde in such a manner attests to the innate powerlessness of the female body during this period. As Adélaïde grew older, she continued to shape her image through art,

choosing to slightly alter her image as she aged based on self-determination of her personal values.

Adélaïde never married, but nonetheless played an important role in her family upon her older sisters' deaths. She grew into the role of the Madame of the household and used portraiture to portray this idea to the public. According to Jennifer Milam, when Adélaïde obtained the status of a Madame upon the death of Princess Henriette, a full-body portrait of Adélaïde was commissioned by the family. This portrait, *Portrait of Madame Adélaïde Holding a Book of Music* (1758, *Château de Versailles, France*) by John Nattier paints Adélaïde as an authoritative and powerful figure in the context of her home. Adélaïde faces a direct angle to the viewer, sitting on a pedestal with a book of music in her lap and holding up her hand, implying conductor status with an instrument at her side. Her stature implies a proficiency in music and a power in the subject. Milam additionally points out that Adélaïde wears a similar dress as her mother in her mother's commissioned portrait by Nattier a few years prior. This choice also brings a sense of power to the piece as Adélaïde embraces an authoritative role in the painting. These artistic choices all point to Adélaïde embracing her newfound status as the Madame of the household and the power that comes with it. Given that Adélaïde was not married and had no children, she embraced the power of her house instead of focusing on family building. Adélaïde used portraiture to curate an image of leadership and influence within her private sphere, which would not necessarily have been obvious to the average citizen without viewing the paintings.



Nattier, John. *Portrait of Madame Adélaïde Holding a Book of Music*, 1758, *Château de Versailles, France*

Now looking at Adélaïde's patronage of this painting in the context of Rousseau's France and expanding upon Milam's ideas, it is important to mention the two other paintings of Adélaïde's sisters. As Adélaïde commands the viewer she also commands her elder sisters in their adjacent portraits. In the two additional portraits, both sisters face the viewer at an offset angle, looking to the left, whereas Adélaïde looks direct and forward. Although Adélaïde's portrait was not commissioned first, she is the center of the series of portraits. Besides being the sole sister facing the viewer, Adélaïde is painted to appear as if she is conducting her sister

Henriette, who is playing a string instrument. Although Adélaïde is poised and strong, the fact that her sisters fit into this portrait series changes the portrait's messaging. Together the paintings turn Adélaïde's demanding stature into power in the context of her household and her family, rather than a general theme of power. (Hyde et. al). So, while Adélaïde holds power through her procured image, she is also staying within the lanes of Rousseau's idea of how women should act. When combined with her sister's portrait, she appears to be perceived as a matriarch. Rousseau believed that a woman's sole job is to maintain a household. For this reason, Rousseau never directly discusses female empowerment, but his idea of confinement of women to the household allows for an entry point in French culture for women to collect power. Rousseau states "*The women's entire education should be planned in relation to men. To please men, to be useful to them, to win their love and respect, to raise them as children, care for them as adults... these are women's duties in all ages and these are what they should be taught from childhood.*" (Rousseau, *Emile*) In this quote, Rousseau degrades women by placing them in subservience to men, but at the same time, he creates a contradiction. If women are the caretakers and educators of men, then doesn't that mean that women hold a power that men do not? Adélaïde embraces this caveat to Rousseau's theory, using portraiture to create a powerful image of herself within the home, while still accepting the roles society prescribed.

Women artists

For a woman, entering the art world was no easy feat even into the Enlightenment. Even in the time of disestablishing class hierarchy in society, there remained a strict hierarchy when it came to art (Rosenfeld). Into the early 19th century there was a cap that disenfranchised women artists from attending the Academy of Royal Arts, both in France and England. Very few women were allowed into the arts, most of which had to be recommended by royalty. It became even more difficult for women chosen to attend the Academy when the royal family fell during the French Revolution since most women who were offered enrollment to the Academy had royal support (Aurichio & Getty). Those with ties to the monarchy were forced to flee to foreign countries for their safety during the revolution. Despite this hardship, a few phenomenal female artists emerged in this period. With the rise of the middle class during the Enlightenment, however, commissioning portraiture became a more attainable opportunity for the average French citizen. This directly benefited women painters because portraiture did not require the study of anatomy. Women were not allowed to study the nude because it was deemed improper, even when they were offered a spot at the Academy. When their male counterparts were taking classes with a live model, the women were required to paint still-lives. Women painters therefore excelled in the subject of portraiture, since it is typically painted from life. One such middle-class woman artist who found success during the Enlightenment was Élisabeth Louise Vigée LeBrun.

Vigée LeBrun was born in 1755 to artist Louis Vigée. Her father inspired her to pursue art from an early age. There is not a completely clear picture of her formative years, however, it is believed that Vigée LeBrun's father was the first master she ever learned from. Louis Vigée

practiced with oil pastels, which became the first medium Vigée LeBrun practiced. Unfortunately, he passed away when LeBrun was only 12 years old. But before the passing of Louis Vigée, he sent Vigée LeBrun to a fellow artist to learn the basics of oil painting. These talents proved to be useful when Vigée-LeBrun entered the art scene to make money for her family shortly after the death of her father. (Montfort) This was an unprecedented choice because at the time the art scene, specifically the painting scene, was typically occupied by older men. Nonetheless, Vigée LeBrun's talent spoke for itself and she was able to commission portraits to support her family's economic distress after her father's death. So even from an early age, Vigée LeBrun was already breaking normalities placed forth by society. Vigée LeBrun was unique because of her status as a female painter and a breadwinner in a male-dominated space.

The combination of the early start to her professional career and the lack of ability to attend the Academy, Vigée LeBrun lived an independent lifestyle. She was not tied down by the lengthy study requirements of the Academy and had already begun to establish herself. This allowed Vigée LeBrun the freedom to travel and make money on her art. It is important to note that besides her innate talent, Vigée LeBrun's status of being an established artist's daughter afforded her many opportunities that other women painters did not have. Even though her father passed away, Vigée LeBrun was able to visit many private Parisian collections with valuable Old Master paintings for her to copy and learn from. (Karvouni) This may have given her a leg up in comparison to her female counterparts, but in contrast to male painters, the opportunities she was given required her to develop skills all on her own. She did not have Masters to teach her. She taught herself how to paint at a high level through determination and diligent practice. Vigée LeBrun witnessed the biggest leap in her career in 1778 when she was called upon by the Queen, Marie Antoinette, who commissioned a portrait of herself. At this time in her career, Vigée LeBrun was a flourishing artist. She married an art dealer who helped provide her with broader opportunities and was featured as an artist at the Academie Saint-Luc. (Montfort) The first painting commissioned by the Queen is titled, *Portrait en robe à paniers* (1778). As described previously, this portrait was the first of 20 portraits Vigée LeBrun painted of Antoinette over 11 years until the fall of the French monarchy in 1789. Vigée LeBrun's choice to work with Antoinette despite Antoinette's poor reputation attests to Vigée LeBrun's character at its core. Vigée LeBrun was a woman who wanted to make herself known by breaking away from the status quo. As Antoinette's reputation continued to spiral downward, Vigée LeBrun continued to paint for her. Vigée LeBrun was even accused of having sexual relations with the Queen, but nonetheless Vigée LeBrun continued to paint the Queen in attempts to further her career. (De Angelis) In continuation of her disregard for the status quo set by Rousseau, when the French Revolution struck, Vigée LeBrun abandoned her husband and children in France and moved across different European countries for ten years to pursue her artistry and career as an individual. Her husband eventually divorced her, and she finished her career as a successful freelance portrait artist. Through such choices, it is evident that Vigée LeBrun was not interested in embracing the role of a mother in the same way as her female counterparts, such as Marie Antoinette. From becoming a breadwinner and artist for her family at an early age to leaving her

family for the sake of her career in adulthood, the visions of Rousseau did not dictate the career choices made by Vigée LeBrun.



Vigée LeBrun, Élisabeth Louise. Marie Antoinette in a Court Dress, 1778, Kunsthistorisches Museum, Vienna, Gemäldegalerie

Another artist who is considered to have a similar cultural impact as Vigée LeBrun is Adélaïde Labille Guiard. Labille Guiard comes from a different upbringing, born into a non-artistic middle-class family in 1749, Labille Guiard started her art career by being an apprentice to artists of miniatures, pastel painting, and oil painting. Similar to Vigée LeBrun, in 1774 Labille Guiard also displayed work at the Academie Saint-Luc, which was her first major exhibition. Labille-Guiard built her career from scratch. (Montfort). Unlike most other women artists at the time, Labille-Guiard was not born into a family in the arts. However, through studying the Old Masters and apprenticing, she was able to get voted into attending the Academy due to her exceptional talent. (Auricchio, Eighteenth-Century Women...) It was difficult for Labille-Guiard to establish herself in the art scene given her status as a woman, but she was able to build a career for herself despite the harsh conditions. Labille-Guiard stood out from her male counterparts through her “feminist” portraits. In her portrait work, Labille-Guiard typically painted the female muse, painting the aunts of King Louis XVI, or other esteemed women who demanded her services. However, Labille-Guiard also painted self-portraits to be displayed in the Salon. One of the most shocking portraits Labille-Guiard painted was her portrait titled, Self Portrait with Two Pupils (1785, Metropolitan Museum of Art, New York). This portrait is a lifesize painting that features Labille-Guiard painting a portrait while teaching two female pupils. According to Katherine Baetjer, it was perceived that the portrait was Labille-Guiard trying to teach other women painting techniques in an attempt to increase the number of women allowed in the Royal Academy, despite the strict four-women quota. This painting caused a stir in the artistic community, but it was also respected. (Baetjer). The work was displayed in the Salon, as

seen in Pietro Antonio Martini's engraving of the Salon in 1785. It was an interesting and provocative choice for Labille-Guiard to create this portrait given that she was working with the royal family at the time, and it is the King who set the quota for women allowed in the Academy. However, the well-received nature of this self-portrait highlights that Labille-Guiard was able to make a name for herself that other male artists respected. Some critics were in disbelief that a woman could paint at such a caliber, going so far as to claim that this portrait was actually a male's work. (Auricchio, Self promotion in Adélaïde Labille-Guiard's..)

The trend of disbelief in women's art ties back to Rousseau's work. In his essay, *Sur la Peinture*, Rousseau claims, "Women in general don't love any art, aren't knowledgeable about any, and do not have any genius. They are able to succeed in little works, they are able to acquire knowledge ... but that celestial fire ... that genius ... that burning eloquence, those sublime raptures ... are always lacking." (Rousseau, *Sur la Peinture*, 5). Through her work and her critics, it is clear that Labille-Guiard did in fact have a genius that is found in large works. Her works were exhibited alongside the male works of the Salon. Even a controversial piece of work, *Self Portrait with Two Pupils*, was included in the Salon. Labille-Guiard's messaging may be polarizing, but her talent and prowess are undeniable. To be a woman trying to succeed in a career path of men, one must not be afraid of the critics and the status quo to make themselves known. Just by sharing her art publicly and embracing her talent, Labille-Guiard broke the status quo Rousseau embedded in the art community.



Labille-Guiard, Adélaïde. *Self-Portrait with Two Pupils*, 1785, Metropolitan Museum of Art, New York

In reviewing women artists and patrons within the context of Rousseau's France, it is necessary to introduce the ideals of Mary Wollstonecraft, who held contrasting beliefs to Rousseau within feminist circles. Rousseau believed in a superior culture of fraternity and therefore found himself popular amongst men. This isn't to say that some women didn't approve of his messaging also, given that women patrons such as Marie Antoinette and Marie Adélaïde embraced the preachings of Rousseau through their commissioned paintings. However, for women artists, it would be remiss to solely explain their actions of social disobedience based on contempt for Rousseau's ideas on womanhood. Wollstonecraft opposed Rousseau's ideals of the

Enlightenment by arguing that women were capable of rational thought, and the institutional oppression of barring women from education was the means men used to impose a gendered hierarchy on society to convince citizens that women were incapable of rational thought. These ideas are highlighted in her text, *A Vindication of the Rights of Women*: “*Taught from their infancy that beauty is woman's sceptre, the mind shapes itself to the body, and roaming round its gilt cage, only seeks to adorn its prison.*” (Wollstonecraft). Through this quote, Wollstonecraft argues that the cage women perpetually remain in is not because of their innate subordination to men, but rather because they are taught their only value comes from their service to men. So, instead of attempting to break out of the cage, they simply adorn it. This text is one of the first modern texts to argue that women aren't innately subordinate to men, but rather oppressed through compliance with the patriarchy. Although Wollstonecraft is writing from an English perspective, her ideas permeated French intellectual circles before the French Revolution. *A Vindication of the Rights of Women* is a call out to the women of the world to break the compliant relationship with gendered societal entrapment. Wollstonecraft understood that philosophers such as Rousseau introduced their ideas about women as the natural order so women accepted them, but such philosophies were actually a means to further the patriarchy and male domination (González). Both Wollstonecraft and Rousseau draw their political and social philosophy from modern liberalism, but they each place the focus on their respective gender. Wollstonecraft's ideals for women shine through both Labille-Guiard and Vigée LeBrun. In the ways these two artists disobey Rousseau's ideals, they embody Wollstonecraft's principles on women's empowerment. Vigée LeBrun embraces Wollstonecraft's economic philosophy.

Wollstonecraft believed that, beyond limited access to education, a lack of opportunities for women to participate in economic markets creates gender barriers that result in a culture of female inferiority.(MacKenzie). As stated previously, at the crest of the French Revolution, Vigée LeBrun fled France without her husband and children to pursue her art career in European countries. She became a successful artist throughout multiple countries and increased her cultural influence and relevance because of the strides she made as a sole woman. Labille-Guiard embraced Wollstonecraft's educational beliefs. Labille-Guiard's infamous painting, *Self Portrait with Two Pupils* not only is a feminist statement piece, it also gives the viewer a glimpse into her career as an educator. Similar to Wollstonecraft, Labille-Guiard makes a statement with this painting that there should be more Academy-trained female artists. As a woman who attended the Academy, she is transferring the lessons and skills she learned to educate up-and-coming female artists. This painting not only demonstrates Labille-Guiard's talent as an artist but also as a woman trying to break down patriarchal values.

Rousseau's Enlightenment ideals did not dictate how every person in society behaved, but his principles were heavily popularized and were even accepted by women living elite life, such as French royalty. It would be impossible to label one philosopher as the dictator of culture during this period, but Rousseau's widespread influence across both gender and class suggests that he was one of the most influential figures during the French Enlightenment. Although Marie Antoinette and Marie Adélaïde were women who lived a life unlike most French women due to

their powerful royal status, they still embraced elements of Rousseau's ideas. Their method for embracing his preachings may not have been exactly what Rousseau intended, but the through-lines of Rousseau's core principles that a woman's most important role in society is in her private sphere was present in the public personas they both tried to curate through portraiture. Both women found art to be a useful tool in cultivating an image of power by embracing elements of their domestic identity. Even though Vigée LeBrun and Labille-Guiard did not adhere to Rousseau's ideals of a woman's career and position in relation to motherhood, they too embraced his teachings about social mobility and the disruption of a predestination-based class system by helping art patrons curate a domestic image of themselves to change public opinions. These artists embraced the aspects of Rousseau's teaching that hierarchy creates unnatural divisions in society, even though they disregarded Rousseau's theories on gendered division. Rousseau's ideas were popularized but not fully accepted. People took parts of ideas they liked from Rousseau and dismissed other parts of his narrative. Women were able to find power through degrading certain tenets of his teachings about social mobility and using art as a medium for exercising self-determination.

Works Cited:

- Auricchio, Laura. "Eighteenth-Century Women Painters in France." The Met's Heilbrunn Timeline of Art History, Aug. 2004, www.metmuseum.org/toah/hd/18wa/hd_18wa.htm. Accessed 25 Aug. 2024.
- Auricchio, Laura, and Paul Getty. Adélaïde Labille-Guiard: Artist in the Age of Revolution. J. Paul Getty Museum, 2009, books.google.com/books?hl=en&lr=&id=XJkvjgxwby4C&oi=fnd&pg=PP1&dq=adelaide+labille+guiard&ots=ggXGFL5JwH&sig=kBOF_Zjkl9GI6PpXq12RkXwEfw#v=onepage&q=adelaide%20labille%20guiard&f=false. Accessed 25 Aug. 2024.
- Auricchio, Laura. "Self-Promotion in Adélaïde Labille-Guiard's 1785 'Self-Portrait with Two Students.'" The Art Bulletin, vol. 89, no. 1, 2007, pp. 45–62. JSTOR, <http://www.jstor.org/stable/25067300>. Accessed 25 Aug. 2024.
- Baetjer, Katharine. "Adélaïde Labille-Guiard (1749–1803)." The Met's Heilbrunn Timeline of Art History, 2016, www.metmuseum.org/toah/hd/lagui/hd_lagui.htm. Accessed 25 Aug. 2024.
- Blacker, J. G. C. "Social Ambitions of the Bourgeoisie in 18th Century France, and Their Relation to Family Limitation." Population Studies, vol. 11, no. 1, 1957, pp. 46–63. JSTOR, <https://doi.org/10.2307/2172509>. Accessed 25 Aug. 2024.
- Cameron, Vivian. "Gender and Power: Images of Women in Late 18th-Century France." History of European Ideas, vol. 10, no. 3, 1989, pp. 309–32, [https://doi.org/10.1016/0191-6599\(89\)90131-9](https://doi.org/10.1016/0191-6599(89)90131-9). Taylor & Francis. Accessed 25 Aug. 2024.
- Fermon, Nicole. Domesticating Passions. Wesleyan University Press, 1997.
- Gates, Eugene, and Karla Hartl. The Women in Music Anthology. The Karpalova Society, 2021, p. 5.
- Getty Museum. "Portrait of Louis XIV." Getty.edu, 22 Feb. 2024, www.getty.edu/art/collection/object/103RA8. Accessed 25 Aug. 2024.
- González, Matilde Martín. "MARY WOLLSTONECRAFT AND THE 'VINDICATION OF THE RIGHTS OW WOMAN': POSTMODERN FEMINISM vs. MASCULINE ENLIGHTENMENT." Atlantis, vol. 19, no. 2, 1997, pp. 177–83. JSTOR, <http://www.jstor.org/stable/41055470>. Accessed 25 Aug. 2024.
- Hyde, Melissa, et al. Women, Art and the Politics of Identity in Eighteenth-Century Europe. London New York Routledge, 2003.
- Karvouni, Evangelia. "Elisabeth Louise Vigée Le Brun: A Historical Survey of a Woman Artist in the Eighteenth Century." Virtual Commons - Bridgewater State University, 2014, vc.bridgew.edu/jiws/vol15/iss2/18. Accessed 25 Aug. 2024.
- Kavanagh, Julia. Woman in France during the Eighteenth Century: With Portraits. 1850. Smith, pp. 61–78, books.google.com/books?id=_Mk5AAAACAAJ&dq=female+portraits+france+18th+century&lr=&source=gbs_navlinks_s. Accessed 25 Aug. 2024.
- Labille-Guiard, Adélaïde. Self-Portrait with Two Pupils, 1785, Metropolitan Museum of Art, New York
- MacKenzie, Catriona. "Reason and Sensibility: The Ideal of Women's Self-Governance in the Writings of Mary Wollstonecraft." Hypatia, vol. 8, no. 4, 1993, pp. 35–55. JSTOR, <http://www.jstor.org/stable/3810368>. Accessed 25 Aug. 2024.

- Montfort, Catherine R. "Self-Portraits, Portraits of Self: Adélaïde Labille-Guiard and Elisabeth Vigée Lebrun, Women Artists of the Eighteenth Century." *Pacific Coast Philology*, vol. 40, no. 1, 2005, pp. 1–18. JSTOR, <http://www.jstor.org/stable/25474166>. Accessed 25 Aug. 2024.
- Murray, J. Clark. "Rousseau: His Position in the History of Philosophy." *The Philosophical Review*, vol. 8, no. 4, 1899, pp. 357–70. JSTOR, <https://doi.org/10.2307/2176197>. Accessed 25 Aug. 2024.
- Nattier, John. *Portrait of Madame Adélaïde Holding a Book of Music, 1758*, Château de Versailles, France
- Rigaud, Hyacinthe. *Portrait of King Louis XIV, 1701*, Musée de Louvre, Paris
- Rosenfeld, Jason. "The Salon and the Royal Academy in the Nineteenth Century ." *The Met's Heilbrunn Timeline of Art History*, Oct. 2004, www.metmuseum.org/toah/hd/sara/hd_sara.htm. Accessed 25 Aug. 2024.
- Rousseau, Jean. *Sur La Peinture*. G. P. Maisonneuve, 1907.
- Rousseau, Jean-Jacques. *Emile*. David l'Angelier, 1762.
- Strobel, Heidi A. "Royal 'Matronage' of Women Artists in the Late-18th Century." *Woman's Art Journal*, vol. 26, no. 2, 2005, pp. 3–9. JSTOR, <https://doi.org/10.2307/3598091>. Accessed 25 Aug. 2024.
- Vigée LeBrun, Élisabeth Louise. *Marie Antoinette in a Chemise dress, 1783*, Kunsthistorisches Museum, Vienna, Gemäldegalerie
- Vigée LeBrun, Élisabeth Louise. *Marie Antoinette in a Court Dress, 1778*, Kunsthistorisches Museum, Vienna, Gemäldegalerie
- Villers, Marie Denise. *Marie Joséphine Charlotte du Val d'Ognes, 1801*, Metropolitan Museum of Art, New York
- Wang , Junyan. *Reframing Marie Antoinette: The Queen's Transformation through Portraiture during Late 18th-Century France*. Warrick Evans Publishing, 15 Sept. 2023, pp. 1–8, wepub.org/index.php/TSSEHR/article/view/11. Accessed 15 Aug. 2024.
- Wollstonecraft, Mary. *A Vindication of the Rights of Woman*. 1792.

Congestive Heart Failure in Humans and Dogs: A Review By Sakura Nakanishi

Abstract

Cardiovascular diseases are one of the leading causes of death worldwide in both humans and dogs. Specifically, congestive heart failure (CHF) is one disease of concern due to its high mortality rate and prevalence in both species. This review covers cardiovascular diseases in humans and dogs, with a focus on conditions that lead to CHF. A general description of cardiovascular disease pathologies will be given, followed by the genetics of various cardiovascular diseases. Environmental factors such as sunlight exposure, birth month, pollution, and diet have a significant influence on a variety of cardiovascular diseases. In terms of diagnosis, thoracic radiography, electrocardiography, echocardiography, and cardiac biomarkers are used in both species, but the diseases diagnosed with the technologies and the exact tools used can vary. Finally, treatment for cardiovascular diseases in humans and dogs is also quite similar, with drugs like spironolactone, furosemide, benazepril, and atenolol used. However, major differences in dosing, diseases treated with the medication, and method of administration exist for some drugs. In general, much less information is available about cardiovascular diseases in dogs compared to humans. Nonetheless, cardiovascular diseases are very similar in humans and dogs. This highlights the importance of further research on cardiovascular disease in dogs, as that information may be applicable to humans as well, thus contributing to our understanding of the disease in both humans and canines. With more research done with collaboration between human and veterinary medicine, CHF and other cardiovascular conditions will likely become more treatable and maybe even curable in both humans and dogs.

Introduction

Cardiovascular diseases affect the circulatory system, which is composed of the heart and systemic blood vessels. These diseases are the leading cause of death worldwide, accounting for approximately 30% of deaths in humans and 11% of diagnoses in dogs at animal hospitals (Thiriet; Parker et al.). As the most common and severe noncommunicable diseases in both humans and dogs, extensive research has been done in both fields. Specifically, congestive heart failure (CHF) is one disease of concern due to its prevalence and high mortality rate. It is a complex multifactorial disease that affects 1-2% of people worldwide, making it a very important disease for further investigation (Schwinger). CHF is a condition where the heart is no longer able to sufficiently operate at an efficient rate to keep up with the body's oxygen demands (Criteria). Many cardiovascular diseases can lead to CHF, such as various cardiomyopathies and ischemic heart disease. In dogs, myxomatous valvular disease (degenerative mitral valve disease) is the most common cause of CHF (de Madron et al.).

Congestive Heart Failure (CHF)

In CHF, the heart cannot provide enough blood flow to meet the body's oxygen demands when it is at its regular state or it can only do so with higher pressures in the heart (Criteria). In

dogs, CHF is typically presented with a systolic murmur with high intensity over the mitral valve area (de Madron et al.). Other symptoms in dogs include pulmonary edema and effusion, tachypnea, respiratory distress, reduced exercise ability and willingness, cough, syncope, and ascites (Beaumier et al.; Erling and Mazzaferro). In humans, it is usually presented with pulmonary and systemic congestion, fatigue, dyspnea, and reduced exercise ability and willingness (Malik et al.). CHF has a prevalence of 2.6% in the United States (US) and is more common in those with diabetes, hypertension, or obesity (Criteria). In total, about 64.3 million people worldwide have CHF, and it can be the final outcome of a variety of cardiac conditions (Criteria; Malik et al.; Groenewegen et al.). In dogs, most die within 6 to 14 months after diagnosis (de Madron et al.). A study done with 54 dogs with advanced heart failure (HF), the median survival time after diagnosis was around 281 days, supporting the range above (Beaumier et al.). In humans, most live for less than 5 years after being diagnosed (Criteria).

Pathophysiology and/Epidemiology of CHF Causes

Cardiomyopathies

Cardiomyopathies are one cause of HF in both humans and dogs (Criteria). They are associated with dysfunction of either the muscular or electrical mechanisms of the heart (Wexler et al.). Five kinds of cardiomyopathies are discussed in this paper.

Dilated Cardiomyopathy (DCM)

Dilated cardiomyopathy (DCM) is where one or both ventricles of the heart are enlarged and dysfunctional. Symptoms in humans include those typically associated with CHF, such as congestive edema, dyspnea, arrhythmias, and orthopnea (Ciarambino et al.). The disease is more common in men than in women, and has a prevalence of around 36 cases per 100,000 annually (Mahmaljy et al.; Mestroni et al.). In dogs, syncope and exercise intolerance are the most common symptoms, although some are asymptomatic and about 1.1% of dogs have this disease (McCauley et al.). DCM is more common in male dogs than female dogs (Broschk and Distl). Exploring this shared feature of sex predisposition and how exactly it affects risk of DCM in both species could help better identify risk populations and prevent or delay the progression of this disease.

Hypertrophic Cardiomyopathy

Hypertrophic cardiomyopathy is a disease characterized by an unusually thickened myocardium, and is the most common primary cardiomyopathy diagnosed (Ciarambino et al.; Criteria). The estimated prevalence for hypertrophic cardiomyopathy is 1 in 500 in the US (Butzner et al.). Common symptoms are dyspnea when undergoing exertion (90% of symptomatic patients), chest pain on exertion (25-30%), syncopal episodes (15-25%), and pre-syncope episodes (20%) (Ciarambino et al.). However, some people are asymptomatic (Spirito et al.). In dogs, heart murmurs, exercise intolerance, and syncope are the most common clinical signs, although many lack any signs (Schober

et al.). It is estimated that as many as 1 in 7 dogs may have this condition although most of the time, the cases are subclinical (Merck Veterinary Manual).

Restrictive Cardiomyopathy

Restrictive cardiomyopathy is identified by ventricular wall stiffness leading to increased end-diastolic pressure, diastolic dysfunction, and enlarged atria (Ciarambino et al.). Systolic function is usually maintained as the thickness of the myocardium remains the same and dilated ventricles are not present (Ciarambino et al.). Only around 5% of cases of cardiomyopathy are restrictive cardiomyopathy, making it one of the least common cardiomyopathies (Brown et al.). The most common symptoms in humans include pulmonary edema, dyspnea, fatigue, orthopnea, heart palpitations, and chest pain (Ciarambino et al.). Additionally, a systolic murmur, third heart sound, peripheral edema, jugular venous distension, and pulmonary rales can be found upon clinical examination (Ciarambino et al.). Restrictive cardiomyopathy is more common in girls and children have the worst prognosis (Ciarambino et al.). In dogs, the symptoms are similar to those of hypertrophic cardiomyopathy, which include symptoms like heart murmurs and exercise intolerance (Merck Veterinary Manual). Investigating whether restrictive cardiomyopathy is also more common in younger female dogs could help veterinarians better diagnose young dogs that present with cardiovascular symptoms.

Arrhythmogenic Right Ventricular Cardiomyopathy

Arrhythmogenic right ventricular cardiomyopathy is where the heart muscle is substituted by fibrofatty tissue. Right ventricular fibrosis is the most distinctive feature of this disease (Ciarambino et al.; Wexler et al.). In humans, the disease is often present with arrhythmia, sudden cardiac death, and a reduction in systolic function (Ciarambino et al.). In dogs (especially Boxers and English bulldogs), the earliest symptoms are asymptomatic arrhythmias from the right ventricle, while syncope, weakness, and sudden death are clinical signs as the condition worsens (Cunningham and Dos Santos).

Arrhythmias

Arrhythmias are a variety of conditions related to disorders of the heart rhythm and heart rate. Two general types of arrhythmia include supraventricular tachycardia (ex: atrial fibrillation, atrial tachycardia, etc.), which are arrhythmias that start from above the atrioventricular (AV) node, and ventricular tachycardia (ex: ventricular fibrillation, ventricular tachycardia etc.), which are arrhythmias that begin below the AV node (Desai et al.). Some patients are asymptomatic, while others experience symptoms like heart palpitations, syncope, sudden tachycardia, weakness, dizziness, hypotension, and shortness of breath (Desai et al.). Arrhythmias have an expected prevalence of 1.5-5% in the general population, with atrial fibrillation being the most common type (Desai et al.). In dogs, it has a prevalence of around 3.2% (Hellemans et al.).

Myocardial Infarction

Myocardial infarctions (MI) occur when blood flow is decreased or eliminated from a part of the heart muscle, resulting in myocardial death (Ojha and Dhamoon). Common symptoms include chest pain, discomfort in the mandibular or epigastric regions, and pain in the upper extremities. Sweating, dyspnea, syncope, nausea, and abdominal pain are common as well (Ojha and Dhamoon). In some cases, patients are asymptomatic or have mild problems like palpitations, but some cases are severe, resulting in cardiac arrest. It has a prevalence of 3.0% in adults in the US, with white males being most commonly affected (Ojha and Dhamoon). MI are very rare in dogs (Driehuys et al.). Further studies on why domesticated dogs as a species rarely get MI could help in considerations for treatment approaches for humans.

Degenerative Mitral Valve Disease

Degenerative mitral valve disease, also known as myxomatous valve disease, is the most common valvular heart disease in humans (Rajamannan). In 2017, 18.1 million people had degenerative mitral valve disease worldwide (Yadgir et al.). It can lead to a condition known as mitral valve prolapse, and it is characterized by slow and chronic changes to the mitral valve of the heart (Neto et al.; Svensson et al.). Mitral valve prolapse is the degeneration of either one or both mitral valve leaflets (Desai et al.). Specifically, the mitral valve thickens slowly, starting from the edges of the valve, making the valves insufficient (Domanjko-Petric). It is the most common cause of HF in dogs and can also affect the other valves (Domanjko-Petric). Symptoms of degenerative mitral valve disease in dogs include cough, syncope, tachypnea, and dyspnea (Domanjko-Petric). It is more common in male small breed dogs (Domanjko-Petric; Mattin et al.). In humans, fatigue and dyspnea are the most common symptoms of the disease, but some people are asymptomatic (Rajamannan); it occurs most commonly in young females, but males tend to exhibit more symptoms. The mitral valve is degenerated in more than 60% of those with degenerative mitral valve disease in humans (Neto et al.). More research on the influence of sex on degenerative mitral valve disease risk should be done as this difference is found in both humans and dogs and could help with improving prevention and treatment.

Ischemic Heart Disease

Ischemic heart disease is where atherosclerotic plaques form in medium to large sized coronary arteries that obstruct the blood flow to the heart for proper heart contraction (Severino et al.). The hallmark symptoms of ischemic heart disease are fatigue and angina (Criteria). It has a prevalence of around 1.72% of the world's population, and about 126 million people are afflicted (Khan et al.). The disease appears to be relatively rare in dogs (Falk and Jönsson).

Genetics

Humans

Types of Inheritance

Autosomal Dominant

Familial dilated cardiomyopathy, hypertrophic cardiomyopathy, and restrictive cardiomyopathy are autosomal dominant diseases with age-related penetrance (Czepluch et al.). Age-related penetrance is where an individual's likelihood to have a diseased phenotype given that they have the genotype for that disease changes as they age (Criteria).

Genes

Dilated Cardiomyopathy

Over one hundred genes have been found to be linked to many cardiomyopathies in humans (Czepluch et al.; Kaviarasan et al.). The gene LMNA (chromosome 1) which codes for Lamin A/C (nuclear envelope proteins), and the MYH7 gene (chromosome 14) which codes for the heavy chain of beta myosin are found to be associated with DCM. Beta myosin is an important motor protein found in cardiomyocytes that facilitates the contraction of the heart (Barrick and Greenberg). The gene for titin, TTN (chromosome 2) as well as the gene for troponin T, TNNT2 (chromosome 1) are also found to be implicated in DCM (Czepluch et al.). Overall, DCM has many different inheritance patterns and mutations: MYH6 (3-4%), BAG3 (2-3%), as well as other genes like SCN5A, DES, FKT, DND, NEXN, and PRDM which frequently have large duplications or deletions (Kaviarasan et al.). RNF207 and PRKAA2 are also found to be implicated (Niskanen et al.). For familial genetic forms, which account for about 30-48% of cases of DCM, the predominant inheritance pattern is autosomal dominant (Mestroni et al.). A R9C and L39stop-PLN mutation in the PLN gene is associated with DCM. Additional mutations in the PLN gene, such as the Leu 39 stop mutation, R14del, R9H, and R9L mutations are found to be associated with DCM (Kranias et al.).

Hypertrophic Cardiomyopathy

This disease is the most common heart disease that has a wide variety of genetic components (Vrablik et al.). Mutations found in the TNNI3 (3-5%) and TNNT2 (3-5%) genes, which code for Troponin I and T respectively, as well as TPM1 (1-3%), which codes for alpha-tropomyosin, and MYL3, which codes for myosin light chain 3, are all implicated in hypertrophic cardiomyopathy (Czepluch et al.; Kaviarasan et al.). Additionally, almost 70% of mutations found in patients with hypertrophic cardiomyopathy are found in the MYH7 (20-30%) and MYBPC3 genes (20-30%), which code for beta myosin heavy chain and cardiac myosin binding protein C, respectively (Czepluch et al.; Kaviarasan et al.). Other genes that are most frequently mutated include NEXN and CAV3 (Kaviarasan et al.). More than 11 genes coding for proteins in heart muscle sarcomeres are linked to the disease (Vrablik et al.).

Restrictive Cardiomyopathy

Mutations found in the genes TNNI3 (Troponin I), MYH7 (beta myosin heavy chain), and DES (desmin) can lead to restrictive cardiomyopathy (Czepluch et al.). The MYBPC3 gene is often mutated too (Kaviarasan et al.). However, some cases are caused by a de novo mutation (Czepluch et al.).

Arrhythmogenic Right Ventricular Cardiomyopathy

Mutations found in the PKP2 genes are often found in those with arrhythmogenic right ventricular cardiomyopathy. DSP, JUP, DSG2, and DSC2 are other genes that are also commonly found to be mutated (Czepluch et al.; Kaviarasan et al.).

Ischemic Heart Disease

The number one cause of chronic HF is ischemic heart disease (Malik et al.). This disease is known to be associated with the LPA locus and the chromosome 9p21 locus (Kessler and Schunkert). Over 200 loci have been reported to be linked to ischemic heart disease, including those involved in progression of plaques, platelet function, and initiation of plaque formation in blood vessels (Kessler and Schunkert).

Dogs

Cardiovascular disease is quite common in dogs - two different studies have shown that about 11% of dogs at veterinary hospitals are diagnosed with a cardiovascular disease. Among those, several are hereditary and prevalent in certain dog breeds. The molecular mechanism and heredity of cardiovascular disease in dogs has been difficult to study as these diseases often are multifactorial, polygenic, and have different penetrance depending on age as opposed to simple Mendelian. Onset of cardiovascular disease is also late, resulting in breeding having already occurred before it is realized that the dog has a cardiac condition (Parker et al.). This makes developing techniques to better determine canine cardiovascular disease genetics very important.

Types of Inheritance

Autosomal Dominant

An organism only needs one copy of the mutant allele for the trait to be present in the diploid organism. Cardiovascular diseases like tricuspid dysplasia in the Labrador, DCM in the Doberman Pinscher, and arrhythmogenic right ventricular cardiomyopathy in the Boxer are autosomal dominant (Parker et al.).

Autosomal Recessive

In autosomal recessive traits, the organism must have two copies of the mutant allele for the trait to be exhibited in the phenotype of the organism. One cardiovascular condition known to have an autosomal recessive inheritance pattern is the juvenile form of familial dilated cardiomyopathy in the Portuguese Water Dog (Parker et al.).

X-linked Recessive

In X-linked recessive traits, more males than females typically exhibit the trait. The gene is found on the X chromosome, so males, who have sex chromosomes XY, only need one mutant allele on the X chromosome to have the diseased phenotype; females, on the other hand, have sex chromosomes XX, so hence, they need both of their X chromosomes to contain the mutant allele in order to have the disease, making affected females much less common than affected males. Adult onset familial dilated cardiomyopathy in Great Danes is thought to be X-linked (Parker et al.).

Polygenic

Polygenic diseases are caused by more than one gene working together to result in the disease phenotype. Conotruncal heart defects in the Keeshond dog, and degenerative mitral valve disease in Dachshunds and Cavalier King Charles Spaniels are thought to be polygenic (Parker et al.). In Cavalier King Charles Spaniels, the prevalence of degenerative mitral valve disease is close to 100% by 11 years of age (O'Brien et al.). Patent ductus arteriosus in Poodles and subvalvular aortic stenosis in Newfoundlands are also polygenic (Parker et al.).

Genes

Dilated Cardiomyopathy

In Dobermanns, DCM is linked to two mutations - a 16-base pair (bp) deletion on pyruvate dehydrogenase kinase 4 (PDK4) gene's donor splice site on intron 10, and a missense mutation on the titin (TTN) gene. The deletion in pyruvate dehydrogenase has 60-68% penetrance and is autosomal dominant, while the missense mutation on TTN has a prevalence of 58% in this breed (Gaar-Humphreys et al.). In humans, low PDK4 expression is associated with the formation of lung tumors, while high PDK4 expression is associated with colon cancer (PDK4 Pyruvate Dehydrogenase Kinase 4 [Homo Sapiens (Human)] - Gene - NCBI). RNF207 and PRKAA2 are also found to be associated with the disease in Dobermanns (Niskanen et al.). In humans, PRKAA2 is associated with many different cancers, such as breast, kidney, liver, gastric, and ovarian cancer (Ouyang et al.). 24-29% of Irish wolfhounds have DCM, with more males affected than females, and the disease being autosomal dominant in nature. Single nucleotide polymorphisms (SNPs) have been linked to DCM in this breed. SNPs on ARHGAP8, PDE3B, FSTL5, one in between introns on chromosome 1 and 17, and one on chromosome 37 in between the MOGATI and ACSL genes have been identified (Gaar-Humphreys et al.). An autosomal recessive mutation on chromosome 8 in dogs can cause juvenile DCM in Portuguese Water Dogs (Gaar-Humphreys et al.). In humans, PRR5-ARHGAP8, the ortholog of ARHGAP8 in dogs, is associated with binge eating behavior in bipolar disorder (McElroy et al.). FSTL5, in humans, is a biomarker for poor prognosis in

medulloblastoma, the most common pediatric malignant brain tumor (Kingwell). They may also be a good target for treating hepatocellular carcinoma (Zhang et al.).

Adult-Onset Cardiomyopathy

Boxers have a specific type of adult-onset cardiomyopathy called “boxer cardiomyopathy” that can be caused by a mutation in the striatin gene (STRN). However, not all Boxers with adult-onset cardiomyopathy studied have this mutation, so there are likely other mutations present (Gaar-Humphreys et al.). In humans, gene variants in striatin are associated with salt sensitivity of blood pressure (Gholami et al.).

Arrhythmias, Left Ventricular Dilation, and Early Sudden Cardiac Death

A substitution in the PLN gene (R9H mutation) in Welsh Springer Spaniels, has been linked to arrhythmia, dilation of the left ventricle, and early sudden cardiac death. The penetrance for this mutation is high (Gaar-Humphreys et al.). The PLN gene is found on chromosome 6 in humans, and instead of arginine like in dogs and other animals, normally has alanine at position 27. Several mutations in the PLN gene are associated with cardiovascular disease in humans as well (Kranias et al.).

Gene	Diseases Implicated	
	Humans	Dogs
PDK4	Lung tumors Colon cancer	Dilated cardiomyopathy (Dobermann)
PRKAA2	Cancers (breast, kidney, liver, ovarian, etc.)	Dilated cardiomyopathy (Dobermann)
ARHGAP8	Binge eating behavior (in bipolar disorder)	Dilated cardiomyopathy (Irish Wolfhound)
FSTL5	Medulloblastoma Hepatocellular carcinoma	Dilated cardiomyopathy (Irish Wolfhound)
STRN	Salt sensitivity of blood pressure	Adult-onset cardiomyopathy (Boxers)
PLN	Cardiovascular diseases	Cardiovascular diseases (Welsh Springer Spaniel)

Figure 1: Genes implicated in cardiovascular disease in dogs and the respective diseases they are associated with in humans.

Human/Dog Genetic Comparison

The R9H mutation in the PLN gene is found in both humans and dogs, but penetrance is lower in humans than in dogs (Gaar-Humphreys et al.). The gene TTN is associated with DCM in both dogs and humans. Dogs seem to have more monogenic cardiovascular diseases than

humans. Additionally, there are many more genes in humans that are associated with more than one disease while in dogs, most of the genes only correspond to one condition. For example, the gene *TNNI3* is associated with DCM, hypertrophic cardiomyopathy, and restrictive cardiomyopathy in humans. This is most likely due to much less research having been done in canine cardiovascular disease genetics compared to in humans, so it is probable that many of the genes that affect multiple diseases in humans also do so in dogs. However, one study found that out of 19 genes associated with DCM in humans, only 1 gene (*DMD*) was also associated with DCM in dogs, suggesting that fewer genes may be implicated in some diseases in dogs compared to humans. It is important to note though, that these studies had small sample sizes (2-66 dogs of each breed) and only studied a few (1-5) breeds, thus more research needs to be done to confirm these results. In this study, one gene (*PDK4*) was shown to be implicated in DCM in dogs but not in humans; this was the only gene studied that had this result (Simpson et al.).

It is noteworthy that the same gene is often associated with different conditions in humans and dogs. Interestingly, some genes are both implicated in cardiovascular disease in the two species, but different ones. For example, *PLN* is associated with DCM in humans, while it is associated with arrhythmias, left ventricular dilation, and early sudden cardiac death in dogs.

Although there is less information out there about dog cardiovascular disease genetics, it remains a source of information and potential for human medicine. In fact, a study done performing genetic analysis of the human genome for the genes *RNF207* and *PRKAA2* (implicated in DCM in dogs), found that these genes are also implicated in DCM in humans as well (Niskanen et al.). This highlights the importance of promoting further research in veterinary cardiovascular disease genetics as it can help enhance the knowledge we have about both human and dog cardiovascular disease.

Some inheritance patterns for the same disease differ by breed in dogs but not by race or ethnicity in humans. Further exploring the differences between ethnicity and genetic predisposition to cardiovascular disease in humans may help patients make better lifestyle choices in order to reduce their risk of developing these debilitating conditions. Specific mechanisms and genes implicated in some cardiovascular diseases need much more research in both humans and dogs.

Environmental Factors

Humans

As with many diseases, various environmental risk factors have been linked to cardiovascular disease, and they also contribute to the progression and severity of the disease. There are different types of environments and the ones explored in this paper are the following: natural, social, and personal.

Natural Environment

Circadian Rhythm

Circadian rhythm disruptions are shown to increase the risk of severe cardiovascular disease due to increased inflammation, blood pressure, and other changes

related to heart gene expression and protein concentrations (Bhatnagar). Additionally, a person is three times more likely to suffer a MI in the early morning than at night, suggesting that severe cardiovascular problems are associated with the circadian rhythm. Symptoms of cardiovascular disease are also more severe in the nighttime (Bhatnagar). This suggests that maintaining one's circadian rhythm can be very important in helping reduce cardiovascular disease risk.

Seasons

Risk factor measurements for cardiovascular disease such as blood pressure and concentrations of high density lipoprotein (HDL), low density lipoprotein (LDL), and blood glucose are higher in the winter than in the summer. The influence of the seasons is also reflected in the fact that 53% more MI cases are reported in the winter months in the US than in the summer months (Bhatnagar). One reason for this may be that in the winter, the elderly are more likely to experience a worsening of pre-existing disease or respiratory infections, which can lead to the acute phase response being triggered. Lower temperatures also destabilize vulnerable lesions, which can lead to occlusive thrombosis and plaque rupture. It also increases coronary artery and vascular resistance as well as blood pressure, which have been associated with MI. On the other hand, high temperatures are also associated with increased cardiovascular disease risk. During extreme heat events, blood flow to the skin increases in order to utilize the effects of evaporative cooling from water, which results in electrolyte imbalances, dehydration, more viscous blood, and a hypercoagulable state. These conditions can lead to various cardiovascular conditions such as arrhythmias, stroke, and acute coronary events (Desai et al.). The significance of both extremes (high and low) in temperature has been shown in other studies. For example, in a 40-year retrospective study that looked at 32 million cardiovascular disease deaths from 27 countries, 2.2 and 9.1 excess deaths per 1000 cardiovascular deaths occurred on hot and cold days, respectively (Alahmad et al.). As the Earth's temperature continues to rise due to the acceleration of global warming, extreme weather and temperature is expected to become more prevalent, making further research in this area extremely important.

Sunlight Exposure

One reason the seasons may be related to changes in cardiovascular risk is due to the difference in the amount of sunlight available at each time of the year. Studies have shown that high sunlight exposure in early life pushes back cardiovascular disease in individuals by as much as 0.6-2.1 years, and spending more time outside is associated with decreased cardiovascular-related death, presumably because vitamin D synthesis occurs in the presence of UV light from the sun. This is likely due to vitamin D's role in regulating inflammation, blood pressure, vascular smooth muscle proliferation, blood volume, and vascular calcification (Danik and Manson).

This difference in vitamin D synthesis is seen with changes in latitude as well as the seasons. There is a gradual increase in blood pressure as one moves away from the equator, both north and south, which correlates with a decrease in UVB radiation. Similarly, UVB radiation is lower in the winter, and blood pressure is higher during these times. The exact link between sunlight and cardiovascular health is still unclear, and UV radiation may not be the only related factor. Other factors of note include the proximity of individuals to natural areas as well as the elevation individuals reside in, with both having a negative correlation to cardiovascular disease risk (Danik and Manson).

Social Environment

Socioeconomic Factors

One environment unique to humans is the social environment. In particular, living in less affluent neighborhoods has been associated with higher cardiovascular disease risk, even when factors like job, personal income, genetic cardiovascular risk, and education level are removed (Bhatnagar). Reasons for this connection are unclear, but some factors can include: the price and availability of types of food, availability of cigarettes, access to healthcare, and access to transportation (Bhatnagar). Psychological reasons, such as attachment, social relationships, and satisfaction with one's neighborhood may be of importance as well (Bhatnagar). The relationship between psychological conditions and cardiovascular disease may be indirect, such as in the case of depression; depression is associated with cardiovascular disease risk factors like smoking and obesity (Mulle and Vaccarino). Psychological stress appears to also have an impact, with meta-analyses finding that social isolation and stress related to work were associated with a 50% and 40% increase in incident cardiovascular disease risk respectively. On the other hand, people with good mental health have been found to experience a decrease in cardiovascular disease risk. People who have a sense of purpose or are more optimistic have been associated with having a 17% and 35% decreased risk of cardiovascular disease events respectively (Levine et al.).

Pollution

Approximately 70 to 80% of early deaths are due to particulate matter air pollution causing cardiovascular events. Why this is the case is still largely unknown, but this connection between particulate matter and cardiovascular disease has been shown in many studies (Bhatnagar). For example, a longitudinal study in Europe conducted over 11.5 years found that a 100 ng/m³ increase in PM₁₀ (particulate matter less than 10 μm in size) was linked to a 6% increase in heart related events. A 50 ng/m³ increase in PM_{2.5} was shown to have a greater increase in coronary events of 18% (Du et al.). Due to PM exposure as well as volatile organic chemicals, roadway proximity has been linked to increased risk of cardiovascular problems like increased mortality due to stroke, acute

HF, or MI. Environmental noise is also known to increase heart rate and blood pressure (Bhatnagar).

Personal Environment

Immediate Surroundings Throughout Development

The personal environment is also an important aspect of the environment to note. This environment starts in the uterus; an unfavorable uterine environment is correlated with an increased risk of ischemic heart disease. This connection between personal environment and cardiovascular disease risk continues throughout childhood and into adulthood. Overweight children more commonly become adults who are obese, and the increase in obese children has led to increased prevalence of risk factors of cardiovascular disease including high cholesterol and blood pressure. These conditions are predictive of early atherosclerosis and other cardiovascular issues (Bhatnagar).

Diet and Physical Activity

Diet influences all the major risk factors for cardiovascular disease. For instance, increased salt intake increases blood pressure, saturated fats increase levels of cholesterol in the blood, and trans fats make LDLs more likely to contribute to atherosclerosis (Bhatnagar). Physical activity is very important not just for reducing cardiovascular disease risk, but also for maintaining overall health. Exercise that is moderate to high in intensity delays cardiovascular disease by one to three years and increases life expectancy by 1.3 to 3.7 years compared to those who are sedentary (Bhatnagar).

Drugs

Smoking is shown to be associated with an increased risk of cardiovascular disease - smokers are ten times more likely to develop peripheral artery disease and two times more likely to develop coronary disease compared to nonsmokers (Bhatnagar). They are also more likely to experience strokes, MI, HF, sudden cardiac death, and arrhythmias, likely due to the amount of inhaled pollutants (Bhatnagar). Alcohol is also positively correlated with cardiovascular conditions like MI and hypertension, which are associated with HF. However, light drinking is shown to reduce HF risk, so genetics, drinking habits, and other factors may play a role (Djoussé and Gaziano).

Many environmental factors, from social factors to physical factors, are associated with cardiovascular disease. However, much of the mechanisms or reasons as to why these factors are linked remains to be uncovered. Further study of the connection between environment and cardiovascular disease may help enhance prevention and treatment.

Dogs

Like most diseases found in any organism, genetics plays a role in the development of cardiovascular diseases in dogs, including CHF, but so do environmental factors. These environmental factors can result in an increased or decreased risk of an organism developing a disease in combination with one's genetic predisposition of developing the disease.

One interesting environmental factor that is shown to have a role in cardiovascular disease development in canines is birth month. It is important to note that this birth month-cardiovascular risk association depended on the dog breed's genetic predisposition to cardiovascular disease. One study used information about 129,778 dogs of 253 different breeds from the Orthopedic Foundation of Animals, which included factors like the animals' breed, date of birth, and the 'final conclusion' to a cardiac test. Taking the breed of the animal and its genetic predisposition to cardiac conditions into account, the study found that in dog breeds not genetically inclined to develop cardiovascular disease, the risk of developing such a disease was higher in dogs born in the summer than in other months, with dogs with a July birth date having the highest risk. Contrastingly, for dogs that are genetically predisposed to develop cardiovascular disease such as mitral valve disease, persistent right aortic arch, or tricuspid dysplasia, there was a small decrease in the likelihood of those dogs developing such a disease if they were born in the summer. Interestingly, in these dogs, September and December were the birth months that had the highest proportion of dogs with a genetic predisposition having developed a cardiovascular disease. The study showed that for purebred dogs, the month in which the animal was born was more predictive of whether the dog developed a cardiovascular disease rather than the genetic predisposition of cardiovascular disease (Boland et al.).

Additionally, vitamin D is vital for proper heart function, and reduced vitamin D, specifically serum 25 (OH)D levels have been shown to be associated with CHF (Kraus et al.). Composition of diet, such as deficiencies in taurine and carnitine are related to nutritional DCM, and transition from a gluten free to gluten inclusive diet in canines is shown to reduce levels of high-sensitivity cardiac troponin I concentrations, which is related to a decrease in cardiac problems (Haimovitz et al.). Related to nutrition, obesity in dogs leads to heart malfunctions like ventricular hypertrophy (increased wall thickness) and high blood pressure, which then subsequently decreased following weight loss in a study (Partington et al.).

Comparison

Vitamin D abundance and obesity are both directly related to risk of cardiovascular diseases. However, dogs cannot synthesize vitamin D with UV activation, so this relationship is implicated in nutrition and diet instead of light availability due to the seasons, altitude, or geographic location as seen in humans. Seasons affect cardiovascular disease in both dogs and humans but in different ways; birth month is related in dogs but not in humans, and temperature differences due to the seasons is related in humans but not in dogs for cardiovascular disease.

Diagnosis

Thoracic Radiography

Humans

Thoracic radiographs are used to diagnose cardiac contour due to an enlarged heart. Findings such as edema at the base of the lungs, vascular congestion, and an enlarged heart silhouette point towards CHF (Malik et al.; Khalil and Alzahrani). The most commonly seen abnormality in those with HF is cardiomegaly, which can be suspected if the cardiac thoracic ratio (the ratio between heart width and thoracic cage width) is greater than 0.5 (Siwik et al.). However, it is only moderately selective at diagnosing CHF; in other words, the presence of acute pulmonary congestion signs will indicate acute CHF, but lack of this symptom cannot exclude acute CHF (Mueller-Lenke et al.). Additionally, increased ventricular size indicates DCM, and an enlarged atria and venous congestion indicates restrictive cardiomyopathy (Khalil and Alzahrani).

Dogs

Thoracic radiography is the most commonly performed diagnostic test for dogs suspected to have a cardiac disease because the size of the heart can provide abundant insight into potential conditions (Hoque et al.). When these images are taken and used for diagnosis, they are compared with those of a healthy dog of the same breed or by measuring cardiothoracic ratios and cardiac dimensions (Hoque et al.). One commonly used measurement, known as the vertebral heart scale (VHS), compares cardiac silhouette dimensions with thoracic vertebral body length; in dogs, the healthy range is 8.7-10.7, while it tends to be elevated in dogs with heart diseases. However, this diagnostic test has low specificity and sensitivity (Hoque et al.).

Comparison

Thoracic radiography is used in a more qualitative manner in humans than in dogs; in humans, the presence of edema in the lungs or relative differences between heart chambers and valves in healthy individuals versus diseased individuals is used for diagnosis, while in dogs, there is a greater focus on measurement (ex: vertebral heart scale). Although ratios are used in diagnosis in both dogs and humans, the exact ratios used (cardiac thoracic ratio versus vertebral heart scale) are different.

Electrocardiography

Humans

Electrocardiograms (ECGs) are used to diagnose arrhythmias, heart chamber enlargements, prior MI, and intraventricular conduction delay. Arrhythmogenic right ventricular cardiomyopathy, for example, is implicated in epsilon waves identified by electrocardiography (usually 12-lead) (Malik et al.).

Dogs

ECGs are used to diagnose arrhythmias and disturbances in cardiac conduction. Additionally, an increase in R wave amplitude, QRS duration, and Q wave depth in lead II can indicate cardiomyopathy in Cocker Spaniels (Hoque et al.). A 6-lead ECG is the standard in veterinary medicine (Trofimiak and Slivinska).

Comparison

The ECG is a valuable tool for diagnosing arrhythmia and heart conduction issues in both humans and dogs. However, the equipment used is slightly different; in humans, the ECG is a 12-lead system, while in dogs, it is a 6-lead system.

Echocardiography (Echo)

Humans

Echo is the first test done on patients suspected to have HF. This test looks at ventricular function, structural problems in chambers and valves of the heart, and allows for visualization of the motion of the walls of the heart (Malik et al.). It is the main method for identifying and differentiating between different types of cardiomyopathy (Khalil and Alzahrani).

Dogs

Echo is one of the most versatile imaging tools for diagnosing cardiovascular disease (Hoque et al.). It can be used to predict CHF in dogs who have DCM and mitral valve disease (Schober et al.). Echocardiography can be used to identify mitral or aortic regurgitation, fluid around the heart, and other abnormalities that can aid in diagnosis (Singh et al.).

Comparison

Echo is a commonly done cardiovascular disease test in both humans and dogs. In both species, three types of echo are used: 2D echo, which is used to look at left ventricle volume and valvular disease; 3D echo, which is used to measure chamber volume and function as well as regurgitation; and Doppler echo, which is used to measure regurgitation, check left ventricular filling, and assess right atrial and pulmonary artery pressure.

Cardiac Biomarker

Humans

Cardiac biomarkers are important heart specific molecules that are present in a diseased state, and serum B-type natriuretic peptide (BNP), N-terminal pro-BNP (NT-ProBP), and troponin I are often used to evaluate for cardiovascular conditions. BNP levels are used to determine the risk of mortality in people with HF. However, it can be dependent on other factors besides cardiovascular disease, such as hypothyroidism, renal

problems, and obesity (Malik et al.). Consistently high levels of troponin-I in the blood suggest injury to the myocardium, which can aid in diagnosis (Malik et al.).

Dogs

The two most commonly measured cardiac biomarkers in dogs are BNP and cardiac troponin I (cTnI). BNP and its prohormone NT-proBNP, are increased in many cardiac conditions, such as DCM, mitral valve disease, and hypertrophic cardiomyopathy, as well as some non-cardiovascular diseases, such as pulmonary hypertension (Gavazza et al.). Quantifying cardiac troponin helps to quantify myocardial injury level. Cardiac troponin is a sensitive but not specific marker. There are many causes for changes in cardiac troponin levels, including kidney disease, pancreatitis, leptospirosis, DCM, and arrhythmogenic right ventricular cardiomyopathy (Gavazza et al.).

Comparison

The same cardiac biomarkers are used to help aid in diagnosing cardiovascular diseases in both humans and dogs. Additionally, the specificity of both BNP and troponin-I is low in both species; a wide variety of conditions (including non-cardiovascular diseases) are associated with elevated levels of these two molecules, resulting in the need of additional tests to narrow down diagnoses.

Treatments for Cardiovascular Disease

A variety of drugs and therapies are used to treat cardiovascular diseases, including surgeries (Doenst et al.). Here, we will focus on four commonly prescribed drug treatments.

Drugs

Spironolactone

Humans

Spironolactone is used to treat HF and hypertension (high blood pressure) in humans. It is also used to treat noncardiovascular conditions like primary hyperaldosteronism and edema secondary to cirrhosis or a nephrotic syndrome (Patibandla et al.). It belongs to a drug class known as the mineralocorticoid receptor antagonists, and binds to aldosterone receptors in the distal tubules and the collecting duct of the nephron in the kidney, preventing aldosterone from fulfilling its role. Thus, sodium reabsorption is blocked, which reduces water retention while increasing potassium retention, helping reduce blood pressure (Patibandla et al.). Spironolactone is given orally in tablets, and the dose is dependent on what it is being used to treat. Decreases in kidney function and increased levels of potassium in the blood (serum) should be monitored regularly (Patibandla et al.).

Dogs

The American College of Veterinary Internal Medicine (ACVIM) recommends a dose of spironolactone 2.0 mg/kg orally every 12 to 24 hours for dogs with stage C degenerative mitral valve disease (Masters et al.).

Comparison

Much less research and information is available for the use of spironolactone in dogs compared to humans. In dogs, spironolactone is mainly used to treat advanced degenerative mitral valve disease, while in humans, it is used to treat both cardiovascular (high blood pressure and HF with preserved ejection fraction) and noncardiovascular disorders (acne and hyperaldosteronism). Additionally, dosing is weight dependent in dogs, while it is a near constant value for humans although it is condition dependent.

Furosemide

Humans

Furosemide is used to treat edema and volume overload due to CHF; it is also used in a non-cardiovascular context by treating these same symptoms for liver and renal failure. This drug is a loop diuretic, meaning that it acts by being a competitive inhibitor of the NKCC2 cotransporter, which is found on the membrane of the thick ascending limb of the loop of Henle in the nephron. By blocking this cotransporter by competing with chloride, sodium and chloride reabsorption in the loop of Henle is prevented, resulting in potassium not being able to be reabsorbed. This in turn, leads to the removal of calcium and magnesium ions and water from the body (Huxel et al.; Khan et al.). For patients with volume overload and acutely decompensated heart failure (ADHF) that have not used diuretics before, a starting dose of 20-40 mg intravenously is recommended. It is then advised to titrate the dose based on the response. Furosemide at high doses (80mg per day or more), can cause a decrease in thyroid hormone as it can inhibit the binding of thyroid hormone to thyroid binding protein (Khan et al.).

Dogs

Furosemide is the only loop diuretic the ACVIM recommends for treating CHF in dogs with degenerative mitral valve disease (Chetboul et al.). The drug is usually given orally, and an oral disintegrating film (ODF) can be used to make it easier for dog owners to administer it to their pets (Koh et al.). Additionally, diuretic resistance often develops after 14 days of use, and urine volume increases with use. Furosemide also increases levels of plasma creatinine and blood urea nitrogen (BUN) (Hori et al.). Also, high doses of furosemide and other loop diuretics (greater than 80 mg) predicted mortality in patients with advanced CHF due to diuretic resistance (Shams et al.). The standard dose of furosemide for dogs is 2 mg/kg taken orally every 12 hours (Ames et al.).

Comparison

Furosemide is used to treat specifically CHF due to degenerative mitral valve disease in dogs, but it is used to treat CHF in general, as well as renal and liver failure in humans. Both humans and dogs can experience diuretic resistance due to high doses of furosemide.

Benazepril

Benazepril is an ACE (angiotensin converting enzyme) inhibitor. ACE is an enzyme that converts angiotensin I to angiotensin II, which is a step in the RAAS (renin-angiotensin-aldosterone system) pathway. This ultimately increases blood volume and blood pressure. ACE also breaks down bradykinin, a vasodilator that is found in the body. By inhibiting ACE, Benazepril prevents the enzyme from doing its job, resulting in a decrease in blood pressure and an increase in sodium excretion via the urine (Dahal and Gupta).

Humans

This drug is used to treat hypertension and it can be taken together with thiazide diuretics. It helps reduce the risk of cardiovascular issues like MI and strokes by decreasing blood pressure. It is taken orally (Dahal and Gupta). Lisinopril is another commonly prescribed ACE inhibitor used to treat hypertension (Chen et al.).

Dogs

Benazepril is used to treat CHF and hypertension in dogs (Benazepril - an Overview | ScienceDirect Topics). It is also used to reduce proteinuria in dogs that have chronic kidney disease (King et al.). Taking both benazepril and furosemide is more effective for treating CHF due to degenerative mitral valve disease (Coffman et al.). Some evidence shows that benazepril can help decrease the risk of cardiomyopathy in dogs (Benazepril - an Overview | ScienceDirect Topics). The drug is administered orally at a dose of 0.25-0.5 mg/kg every 12-24 hours for treatment of CHF (“Angiotensin-Converting Enzyme Inhibitors for Use in Animals - Pharmacology”). It is also administered orally for treating chronic kidney disease (King et al.).

Comparison

Dosing of benazepril appears to be higher per pound of body weight in dogs than humans for cardiovascular disease treatment. It is used to treat hypertension in humans, while it is used to treat CHF and hypertension in dogs. In both cases, the medication is taken orally, and in children and dogs, dosage depends on the weight of the individual.

Atenolol

This drug is an antagonist of the beta-1 receptor and helps decrease heart muscle contractility, heart rate, and blood pressure. Beta-1 adrenergic receptors are found in smooth

muscle of blood vessels as well as the heart, and by blocking access of catecholamines to these receptors, stimulation by the sympathetic nervous system is prevented (Rehman et al.).

Humans

Atenolol is used to treat angina, hypertension, and acute MI. Off-label uses include prevention of migraines and secondary MI and treatment of arrhythmias and alcohol withdrawal (Rehman et al.). Atenolol can be injected by IV at a concentration of 0.5 mg/mL or taken orally at doses of 25, 50, or 100 mg. It is sometimes combined with chlorthalidone for treating hypertension (Rehman et al.). The recommended dose depends on the condition being treated. Common adverse effects of taking atenolol include sleep problems, erectile dysfunction, and visual, neural, respiratory, digestive, and cardiovascular problems (Rehman et al.).

Dogs

Atenolol is used to treat glomerular disease and cardiovascular diseases in dogs (Subgroup et al.). It is given at a dosage of 0.2-1 mg/kg every 12 hours orally. Dosing is slowly increased as needed, and sudden discontinuation can result in issues. Possible adverse effects include hypotension, myocardial depression, and sinus and AV block bradycardia (“Antiarrhythmics for Use in Animals - Pharmacology”). This drug should not be combined with other drugs that are beta-adrenergic receptor antagonists such as Sotalol and Amiodarone (“Antiarrhythmics for Use in Animals - Pharmacology”).

Comparison

In both humans and dogs, Atenolol is used to treat cardiovascular diseases as well as other diseases, such as alcohol withdrawal symptoms in humans and glomerular disease in dogs. Additionally, adverse side effects in humans relate to many body systems besides the cardiovascular system, while those in dogs tend to be related only to the cardiovascular system. Exploring the potential of IV administration of atenolol in dogs may be interesting as both IV and oral administration are done in humans, but only oral administration is done in dogs.

	Humans		Dogs	
Drug	Disease	Administration	Disease	Administration
Spironolactone	C: Heart Failure, Hypertension	oral	C: Stage C DMVD	oral
	NC: hyperaldosteronism, edema	oral	N/A	N/A
Furosemide	C: CHF	IV	C: CHF with DMVD	oral

	NC: liver failure, renal failure	IV	N/A	N/A
Benazepril	C: hypertension, MI, strokes	oral	C: hypertension, CHF with DMVD	oral
	N/A	N/A	NC: chronic kidney disease	oral
Atenolol	C: angina, hypertension, MI, arrhythmias	IV, oral	C: many	oral
	NC: alcohol withdrawal, migraines	IV, oral	NC: glomerular disease	oral

Figure 2: commonly prescribed medications for cardiovascular diseases. C=cardiovascular, NC=non-cardiovascular.

Conclusion/Discussion

Cardiovascular diseases have complex pathologies and are influenced by a variety of factors including genetics and the environment. These diseases are fairly common in today's society, making the multitude of diagnostic methods and treatment options unsurprising. Despite not being very closely related evolutionarily, humans' and dogs' experiences of cardiovascular diseases are very similar. Many of the same diagnostic tests and treatments are used for both species. For instance, thoracic radiography, echocardiography, and cardiac biomarkers are all used to aid in diagnosing CHF in both organisms, and drugs like spironolactone, furosemide, and atenolol are used to treat CHF in both. However, there are major differences between the two species when it comes to the dosing of certain drugs, various environmental factors, and genes implicated in some cardiovascular conditions. Further studies that compare cardiovascular diseases between humans and dogs can help identify existing treatment options and diagnostic tests in one organism that can be tested and used in the other if successful. These studies may also elucidate various aspects of pathophysiology, genetics, and the influence of the environment on cardiovascular disease in dogs or humans that is already known to be true in one species that can be applied to the other, helping advance both veterinary and human medicine. More research should be done regarding cardiovascular disease in dogs in general, perhaps by referencing what we know about human cardiovascular disease, and further exploration of the genetics and environmental factors that influence these complex diseases in canines is needed as information regarding these two aspects are very lacking. Perhaps with further research done with collaboration between veterinary and human medicine, cardiovascular disease, especially CHF, will become a much more treatable and maybe even curable condition in both humans and dogs.

Works Cited

- Alahmad, Barrak, et al. "Associations Between Extreme Temperatures and Cardiovascular Cause-Specific Mortality: Results From 27 Countries." *Circulation*, vol. 147, no. 1, Jan. 2023, pp. 35–46. PubMed Central, <https://doi.org/10.1161/CIRCULATIONAHA.122.061832>.
- Ames, Marisa K., et al. Effects of High Doses of Enalapril and Benazepril on the Pharmacologically Activated Renin-Angiotensin-Aldosterone System in Clinically Normal Dogs. Dec. 2015. avmajournals.avma.org, <https://doi.org/10.2460/ajvr.76.12.1041>.
- "Angiotensin-Converting Enzyme Inhibitors for Use in Animals - Pharmacology." Merck Veterinary Manual, <https://www.merckvetmanual.com/pharmacology/systemic-pharmacotherapeutics-of-the-cardiovascular-system/angiotensin-converting-enzyme-inhibitors-for-use-in-animals>. Accessed 16 Aug. 2024.
- "Antiarrhythmics for Use in Animals - Pharmacology." Merck Veterinary Manual, <https://www.merckvetmanual.com/pharmacology/systemic-pharmacotherapeutics-of-the-cardiovascular-system/antiarrhythmics-for-use-in-animals>. Accessed 27 July 2024.
- Barrick, Samantha K., and Michael J. Greenberg. "Cardiac Myosin Contraction and Mechanotransduction in Health and Disease." *The Journal of Biological Chemistry*, vol. 297, no. 5, Oct. 2021, p. 101297. PubMed Central, <https://doi.org/10.1016/j.jbc.2021.101297>.
- Beaumier, Amelie, et al. "Clinical Findings and Survival Time in Dogs with Advanced Heart Failure." *Journal of Veterinary Internal Medicine*, vol. 32, no. 3, 2018, pp. 944–50. PubMed Central, <https://doi.org/10.1111/jvim.15126>.
- Benazepril - an Overview | ScienceDirect Topics. <https://www.sciencedirect.com/topics/veterinary-science-and-veterinary-medicine/benazepril>. Accessed 27 July 2024.
- Bhatnagar, Aruni. "Environmental Determinants of Cardiovascular Disease." *Circulation Research*, vol. 121, no. 2, July 2017, pp. 162–80. PubMed Central, <https://doi.org/10.1161/CIRCRESAHA.117.306458>.
- Boland, Mary Regina, et al. "Cardiovascular Disease Risk Varies by Birth Month in Canines." *Scientific Reports*, vol. 8, no. 1, May 2018, p. 7130. www.nature.com, <https://doi.org/10.1038/s41598-018-25199-w>.
- Broschk, C., and O. Distl. "[Dilated cardiomyopathy (DCM) in dogs--pathological, clinical, diagnosis and genetic aspects]." *DTW. Deutsche tierärztliche Wochenschrift*, vol. 112, no. 10, Oct. 2005, pp. 380–85.
- Brown, Kristen N., et al. "Restrictive Cardiomyopathy." StatPearls, StatPearls Publishing, 2024. PubMed, <http://www.ncbi.nlm.nih.gov/books/NBK537234/>.
- Butzner, Michael, et al. "Stable Rates of Obstructive Hypertrophic Cardiomyopathy in a Contemporary Era." *Frontiers in Cardiovascular Medicine*, vol. 8, Jan. 2022, p. 765876. PubMed Central, <https://doi.org/10.3389/fcvm.2021.765876>.
- Chen, RuiJun, et al. "Comparative First-Line Effectiveness and Safety of ACE (Angiotensin-Converting Enzyme) Inhibitors and Angiotensin Receptor Blockers: A Multinational Cohort Study." *Hypertension*, vol. 78, no. 3, Sept. 2021, pp. 591–603. ahajournals.org (Atypon), <https://doi.org/10.1161/HYPERTENSIONAHA.120.16667>.
- Chetboul, V., et al. "Short-Term Efficacy and Safety of Torasemide and Furosemide in 366 Dogs

- with Degenerative Mitral Valve Disease: The TEST Study.” *Journal of Veterinary Internal Medicine*, vol. 31, no. 6, 2017, pp. 1629–42. PubMed Central, <https://doi.org/10.1111/jvim.14841>.
- Ciarambino, Tiziana, et al. “Cardiomyopathies: An Overview.” *International Journal of Molecular Sciences*, vol. 22, no. 14, July 2021, p. 7722. PubMed Central, <https://doi.org/10.3390/ijms22147722>.
- Coffman, Melissa, et al. “Clinical Efficacy of a Benazepril and Spironolactone Combination in Dogs with Congestive Heart Failure Due to Myxomatous Mitral Valve Disease: The BENazepril Spironolactone Study (BESST).” *Journal of Veterinary Internal Medicine*, vol. 35, no. 4, 2021, pp. 1673–87. Wiley Online Library, <https://doi.org/10.1111/jvim.16155>.
- Criteria, Institute of Medicine (US) Committee on Social Security Cardiovascular Disability. “Heart Failure, Cardiomyopathy, and Right Heart Failure.” *Cardiovascular Disability: Updating the Social Security Listings*, National Academies Press (US), 2010. www.ncbi.nlm.nih.gov, <https://www.ncbi.nlm.nih.gov/books/NBK209980/>.
- Cunningham, S. M., and L. Dos Santos. “Arrhythmogenic Right Ventricular Cardiomyopathy in Dogs.” *Journal of Veterinary Cardiology*, vol. 40, Apr. 2022, pp. 156–69. ScienceDirect, <https://doi.org/10.1016/j.jvc.2021.07.001>.
- Czepluch, Frauke S., et al. “Genetic Determinants of Heart Failure: Facts and Numbers.” *ESC Heart Failure*, vol. 5, no. 3, Feb. 2018, pp. 211–17. PubMed Central, <https://doi.org/10.1002/ehf2.12267>.
- Dahal, Sujata S., and Mohit Gupta. “Benazepril.” *StatPearls*, StatPearls Publishing, 2024. PubMed, <http://www.ncbi.nlm.nih.gov/books/NBK549885/>.
- Danik, Jacqueline S., and JoAnn E. Manson. “Vitamin D and Cardiovascular Disease.” *Current Treatment Options in Cardiovascular Medicine*, vol. 14, no. 4, Aug. 2012, pp. 414–24. PubMed Central, <https://doi.org/10.1007/s11936-012-0183-8>.
- de Madron, Eric, et al. “Survival and Echocardiographic Data in Dogs with Congestive Heart Failure Caused by Mitral Valve Disease and Treated by Multiple Drugs: A Retrospective Study of 21 Cases.” *The Canadian Veterinary Journal*, vol. 52, no. 11, Nov. 2011, pp. 1219–25.
- Delwarde, Constance, et al. “Genetics and Pathophysiology of Mitral Valve Prolapse.” *Frontiers in Cardiovascular Medicine*, vol. 10, Feb. 2023, p. 1077788. PubMed Central, <https://doi.org/10.3389/fcvm.2023.1077788>.
- Desai, Yash, et al. “Heat and the Heart.” *The Yale Journal of Biology and Medicine*, vol. 96, no. 2, June 2023, pp. 197–203. PubMed Central, <https://doi.org/10.59249/HGAL4894>.
- Djoussé, Luc, and J. Michael Gaziano. “Alcohol Consumption and Heart Failure.” *Current Atherosclerosis Reports*, vol. 10, no. 2, Apr. 2008, pp. 117–20.
- Doenst, Torsten, et al. “Cardiac Surgery 2022 Reviewed.” *The Thoracic and Cardiovascular Surgeon*, vol. 71, May 2023, pp. 356–65. www.thieme-connect.de, <https://doi.org/10.1055/s-0043-57228>.
- Domanjko-Petric, Aleksandra. “Myxomatous Mitral Valve Disease in Dogs - an Update and Perspectives.” *Macedonian Veterinary Review*, vol. 38, Mar. 2015. ResearchGate, <https://doi.org/10.14432/j.macvetrev.2014.11.026>.
- Driehuys, S., et al. “Myocardial Infarction in Dogs and Cats: 37 Cases (1985-1994).” *Journal of the American Veterinary Medical Association*, vol. 213, no. 10, Nov. 1998, pp. 1444–48.

- Du, Yixing, et al. "Air Particulate Matter and Cardiovascular Disease: The Epidemiological, Biomedical and Clinical Evidence." *Journal of Thoracic Disease*, vol. 8, no. 1, Jan. 2016, pp. E8–19. PubMed Central, <https://doi.org/10.3978/j.issn.2072-1439.2015.11.37>.
- Erling, Peter, and Elisa Mazzaferro. "Left-Sided Congestive Heart Failure in Dogs: Treatment and Monitoring of Emergency Patients." *Compendium (Yardley, PA)*, vol. 30, Mar. 2008, pp. 94–104.
- Falk, T., and L. Jönsson. "Ischaemic Heart Disease in the Dog: A Review of 65 Cases." *The Journal of Small Animal Practice*, vol. 41, no. 3, Mar. 2000, pp. 97–103. PubMed, <https://doi.org/10.1111/j.1748-5827.2000.tb03173.x>.
- Gaar-Humphreys, Karen R., et al. "Genetic Basis of Dilated Cardiomyopathy in Dogs and Its Potential as a Bidirectional Model." *Animals : An Open Access Journal from MDPI*, vol. 12, no. 13, June 2022, p. 1679. PubMed Central, <https://doi.org/10.3390/ani12131679>.
- Gavazza, Alessandra, et al. "Canine Traditional Laboratory Tests and Cardiac Biomarkers." *Frontiers in Veterinary Science*, vol. 7, June 2020, p. 320. PubMed Central, <https://doi.org/10.3389/fvets.2020.00320>.
- Gholami, Shadi K., et al. "Striatin Gene Variants Are Associated With Salt Sensitivity of Blood Pressure by Mechanisms That Differ in Women and Men." *Hypertension*, vol. 81, no. 2, Feb. 2024, pp. 330–39. ahajournals.org (Atypon), <https://doi.org/10.1161/HYPERTENSIONAHA.123.21955>.
- Groenewegen, Amy, et al. "Epidemiology of Heart Failure." *European Journal of Heart Failure*, vol. 22, no. 8, Aug. 2020, pp. 1342–56. PubMed Central, <https://doi.org/10.1002/ejhf.1858>.
- Haimovitz, Dana, et al. "Effect of Diet Change in Healthy Dogs with Subclinical Cardiac Biomarker or Echocardiographic Abnormalities." *Journal of Veterinary Internal Medicine*, vol. 36, no. 3, 2022, pp. 1057–65. PubMed Central, <https://doi.org/10.1111/jvim.16416>.
- Hellemans, Arnaut, et al. "Diagnosis and Management of Arrhythmias in Dogs: A Cross-Sectional Online Survey among Flemish Veterinary Practitioners." *Veterinary Record Open*, vol. 9, no. 1, 2022, p. e35. Wiley Online Library, <https://doi.org/10.1002/vro2.35>.
- Hoque, Mozammel, et al. *CARDIAC DISEASES IN DOGS*. June 2019.
- Hori, Yasutomo, et al. "Effects of Oral Administration of Furosemide and Torsemide in Healthy Dogs." *American Journal of Veterinary Research*, vol. 68, no. 10, Oct. 2007, pp. 1058–63. PubMed, <https://doi.org/10.2460/ajvr.68.10.1058>.
- Huxel, Chris, et al. "Loop Diuretics." StatPearls, StatPearls Publishing, 2024. PubMed, <http://www.ncbi.nlm.nih.gov/books/NBK546656/>.
- Hypertrophic Cardiomyopathy in Dogs and Cats - Circulatory System - Merck Veterinary Manual. <https://www.merckvetmanual.com/circulatory-system/cardiomyopathy-in-dogs-and-cats/hypertrophic-cardiomyopathy-in-dogs-and-cats>. Accessed 26 July 2024.
- Kaviarasan, Vaishak, et al. "Genetic Predisposition Study of Heart Failure and Its Association with Cardiomyopathy." *The Egyptian Heart Journal*, vol. 74, Jan. 2022, p. 5. PubMed Central, <https://doi.org/10.1186/s43044-022-00240-6>.
- Kessler, Thorsten, and Heribert Schunkert. "Coronary Artery Disease Genetics Enlightened by Genome-Wide Association Studies." *JACC: Basic to Translational Science*, vol. 6, no. 7, July 2021, pp. 610–23. PubMed Central, <https://doi.org/10.1016/j.jacbts.2021.04.001>.
- Khalil, Hassan, and Talal Alzahrani. "Cardiomyopathy Imaging." StatPearls, StatPearls

- Publishing, 2024. PubMed, <http://www.ncbi.nlm.nih.gov/books/NBK541056/>.
- Khan, Moien AB, et al. “Global Epidemiology of Ischemic Heart Disease: Results from the Global Burden of Disease Study.” *Cureus*, vol. 12, no. 7, p. e9349. PubMed Central, <https://doi.org/10.7759/cureus.9349>.
- King, J. N., et al. “Effects of Benazepril on Survival of Dogs with Chronic Kidney Disease: A Multicenter, Randomized, Blinded, Placebo-Controlled Clinical Trial.” *Journal of Veterinary Internal Medicine*, vol. 31, no. 4, 2017, pp. 1113–22. PubMed Central, <https://doi.org/10.1111/jvim.14726>.
- Kingwell, Katie. “FSTL5—a New Prognostic Biomarker for Medulloblastoma.” *Nature Reviews Neurology*, vol. 7, no. 11, Nov. 2011, pp. 598–598. www.nature.com, <https://doi.org/10.1038/nrneuro.2011.156>.
- Koh, Suk-Kyu, et al. “Pharmacokinetics and Diuretic Effect of Furosemide after Single Intravenous, Oral Tablet, and Newly Developed Oral Disintegrating Film Administration in Healthy Beagle Dogs.” *BMC Veterinary Research*, vol. 17, no. 1, Sept. 2021, p. 295. PubMed, <https://doi.org/10.1186/s12917-021-02998-4>.
- Kranias, Evangelia G., et al. “PLN Foundation.” *Circulation Research*, vol. 123, no. 12, Dec. 2018, pp. 1276–78. ahajournals.org (Atypon), <https://doi.org/10.1161/CIRCRESAHA.118.314014>.
- Kraus, M. S., et al. “Relation of Vitamin D Status to Congestive Heart Failure and Cardiovascular Events in Dogs.” *Journal of Veterinary Internal Medicine*, vol. 28, no. 1, 2014, pp. 109–15. PubMed Central, <https://doi.org/10.1111/jvim.12239>.
- LaHaye, Stephanie, et al. “Genetics of Valvular Heart Disease.” *Current Cardiology Reports*, vol. 16, no. 6, 2014, p. 487. PubMed Central, <https://doi.org/10.1007/s11886-014-0487-2>.
- Levine, Glenn N., et al. “Psychological Health, Well-Being, and the Mind-Heart-Body Connection: A Scientific Statement From the American Heart Association.” *Circulation*, vol. 143, no. 10, Mar. 2021, pp. e763–83. ahajournals.org (Atypon), <https://doi.org/10.1161/CIR.0000000000000947>.
- Mahmaljy, Hadi, et al. “Dilated Cardiomyopathy.” *StatPearls*, StatPearls Publishing, 2024. PubMed, <http://www.ncbi.nlm.nih.gov/books/NBK441911/>.
- Malik, Ahmad, et al. “Congestive Heart Failure.” *StatPearls*, StatPearls Publishing, 2024. PubMed, <http://www.ncbi.nlm.nih.gov/books/NBK430873/>.
- Masters, Allison K., et al. “Comprehensive Characterization of the Effect of Mineralocorticoid Receptor Antagonism with Spironolactone on the Renin-Angiotensin-Aldosterone System in Healthy Dogs.” *PLOS ONE*, vol. 19, no. 2, Feb. 2024, p. e0298030. PubMed Central, <https://doi.org/10.1371/journal.pone.0298030>.
- Mattin, M. J., et al. “Prevalence of and Risk Factors for Degenerative Mitral Valve Disease in Dogs Attending Primary-care Veterinary Practices in England.” *Journal of Veterinary Internal Medicine*, vol. 29, no. 3, 2015, pp. 847–54. PubMed Central, <https://doi.org/10.1111/jvim.12591>.
- McCauley, Sydney R., et al. “Review of Canine Dilated Cardiomyopathy in the Wake of Diet-Associated Concerns.” *Journal of Animal Science*, vol. 98, no. 6, June 2020, p. skaa155. PubMed Central, <https://doi.org/10.1093/jas/skaa155>.
- McElroy, Susan L., et al. “Bipolar Disorder with Binge Eating Behavior: A Genome-Wide Association Study Implicates PRR5-ARHGAP8.” *Translational Psychiatry*, vol. 8, no. 1,

- Feb. 2018, p. 40. PubMed, <https://doi.org/10.1038/s41398-017-0085-3>.
- Mestroni, Luisa, et al. "GENETIC CAUSES OF DILATED CARDIOMYOPATHY." *Progress in Pediatric Cardiology*, vol. 37, no. 1–2, Dec. 2014, pp. 13–18. PubMed Central, <https://doi.org/10.1016/j.ppedcard.2014.10.003>.
- Mueller-Lenke, N., et al. "Use of Chest Radiography in the Emergency Diagnosis of Acute Congestive Heart Failure." *Heart*, vol. 92, no. 5, May 2006, pp. 695–96. PubMed Central, <https://doi.org/10.1136/hrt.2005.074583>.
- Mulle, Jennifer Gladys, and Viola Vaccarino. "Cardiovascular Disease, Psychosocial Factors, and Genetics: The Case of Depression." *Progress in Cardiovascular Diseases*, vol. 55, no. 6, 2013, pp. 557–62. PubMed Central, <https://doi.org/10.1016/j.pcad.2013.03.005>.
- Neto, Felipe Lazar, et al. "Myxomatous Degeneration of the Mitral Valve." *Autopsy & Case Reports*, vol. 8, no. 4, Nov. 2018, p. e2018058. PubMed Central, <https://doi.org/10.4322/acr.2018.058>.
- Niskanen, Julia E., et al. "Identification of Novel Genetic Risk Factors of Dilated Cardiomyopathy: From Canine to Human." *Genome Medicine*, vol. 15, no. 1, Sept. 2023, p. 73. BioMed Central, <https://doi.org/10.1186/s13073-023-01221-3>.
- O'Brien, M. J., et al. "Genetics of Canine Myxomatous Mitral Valve Disease." *Animal Genetics*, vol. 52, no. 4, 2021, pp. 409–21. Wiley Online Library, <https://doi.org/10.1111/age.13082>.
- Ojha, Niranjana, and Amit S. Dhamoon. "Myocardial Infarction." StatPearls, StatPearls Publishing, 2024. PubMed, <http://www.ncbi.nlm.nih.gov/books/NBK537076/>.
- Ouyang, Yan, et al. "AMPK α 2 Promotes Tumor Immune Escape by Inducing CD8⁺ T-Cell Exhaustion and CD4⁺ Treg Cell Formation in Liver Hepatocellular Carcinoma." *BMC Cancer*, vol. 24, no. 1, Mar. 2024, p. 276. BioMed Central, <https://doi.org/10.1186/s12885-024-12025-y>.
- Parker, Heidi G., et al. "Finding Cardiovascular Disease Genes in the Dog." *Journal of Veterinary Cardiology : The Official Journal of the European Society of Veterinary Cardiology*, vol. 8, no. 2, Nov. 2006, pp. 115–27. PubMed Central, <https://doi.org/10.1016/j.jvc.2006.04.002>.
- Partington, C., et al. "The Effect of Obesity and Subsequent Weight Reduction on Cardiac Structure and Function in Dogs." *BMC Veterinary Research*, vol. 18, Sept. 2022, p. 351. PubMed Central, <https://doi.org/10.1186/s12917-022-03449-4>.
- Patibandla, Saikrishna, et al. "Spironolactone." StatPearls, StatPearls Publishing, 2024. PubMed, <http://www.ncbi.nlm.nih.gov/books/NBK554421/>.
- PDK4 Pyruvate Dehydrogenase Kinase 4 [Homo Sapiens (Human)] - Gene - NCBI. <https://www.ncbi.nlm.nih.gov/gene/5166>. Accessed 17 Aug. 2024.
- Rajamannan, Nalini M. "Myxomatous Mitral Valve Disease Bench to Bedside: LDL-Density-Pressure Regulates Lrp5." *Expert Review of Cardiovascular Therapy*, vol. 12, no. 3, Mar. 2014, pp. 383–92. PubMed Central, <https://doi.org/10.1586/14779072.2014.893191>.
- Rehman, Baryiah, et al. "Atenolol." StatPearls, StatPearls Publishing, 2024. PubMed, <http://www.ncbi.nlm.nih.gov/books/NBK539844/>.
- Schober, K. e., et al. "Detection of Congestive Heart Failure in Dogs by Doppler Echocardiography." *Journal of Veterinary Internal Medicine*, vol. 24, no. 6, 2010, pp. 1358–68. Wiley Online Library, <https://doi.org/10.1111/j.1939-1676.2010.0592.x>.

- Schwinger, Robert H. G. "Pathophysiology of Heart Failure." *Cardiovascular Diagnosis and Therapy*, vol. 11, no. 1, Feb. 2021, pp. 263–76. PubMed Central, <https://doi.org/10.21037/cdt-20-302>.
- Severino, Paolo, et al. "Ischemic Heart Disease Pathophysiology Paradigms Overview: From Plaque Activation to Microvascular Dysfunction." *International Journal of Molecular Sciences*, vol. 21, no. 21, Oct. 2020, p. 8118. PubMed Central, <https://doi.org/10.3390/ijms21218118>.
- Shams, Elham, et al. "Diuretic Resistance Associated With Heart Failure." *Cureus*, vol. 14, no. 1, p. e21369. PubMed Central, <https://doi.org/10.7759/cureus.21369>.
- Simpson, Siobhan, et al. "Genetics of Human and Canine Dilated Cardiomyopathy." *International Journal of Genomics*, vol. 2015, 2015, p. 204823. PubMed Central, <https://doi.org/10.1155/2015/204823>.
- Singh, P., et al. "Echocardiography as an Approach for Canine Cardiac Disease Diagnosis." *Veterinary World*, vol. 7, Nov. 2014, pp. 960–65. ResearchGate, <https://doi.org/10.14202/vetworld.2014.960-965>.
- Siwik, Dominika, et al. "Diagnosing Lung Abnormalities Related to Heart Failure in Chest Radiogram, Lung Ultrasound and Thoracic Computed Tomography." *Advances in Respiratory Medicine*, vol. 91, no. 2, Feb. 2023, pp. 103–22. PubMed Central, <https://doi.org/10.3390/arm91020010>.
- Spirito, P., et al. "Prognosis of Asymptomatic Patients with Hypertrophic Cardiomyopathy and Nonsustained Ventricular Tachycardia." *Circulation*, vol. 90, no. 6, Dec. 1994, pp. 2743–47. ahajournals.org (Atypon), <https://doi.org/10.1161/01.CIR.90.6.2743>.
- Subgroup, IRIS Canine GN Study Group Standard Therapy, et al. "Consensus Recommendations for Standard Therapy of Glomerular Disease in Dogs." *Journal of Veterinary Internal Medicine*, vol. 27, no. s1, 2013, pp. S27–43. Wiley Online Library, <https://doi.org/10.1111/jvim.12230>.
- Svensson, Mikael, et al. "Myxomatous Mitral Valve Disease in Large Breed Dogs: Survival Characteristics and Prognostic Variables." *Veterinary Sciences*, vol. 11, no. 3, 3, Mar. 2024, p. 136. www.mdpi.com, <https://doi.org/10.3390/vetsci11030136>.
- Thiriet, Marc. "Cardiovascular Disease: An Introduction." *Vasculopathies*, vol. 8, Feb. 2019, pp. 1–90. PubMed Central, https://doi.org/10.1007/978-3-319-89315-0_1.
- Trofimiak, R. M., and L. H. Slivinska. "Electrocardiographic Evaluation of Cardiac Activity in Dogs with CHF." *Scientific Messenger of LNU of Veterinary Medicine and Biotechnologies. Series: Veterinary Sciences*, vol. 22, no. 98, 98, Aug. 2020, pp. 100–07. nvlvet.com.ua, <https://doi.org/10.32718/nvlvet9818>.
- Vrablik, Michal, et al. "Genetics of Cardiovascular Disease: How Far Are We from Personalized CVD Risk Prediction and Management?" *International Journal of Molecular Sciences*, vol. 22, no. 8, Apr. 2021, p. 4182. PubMed Central, <https://doi.org/10.3390/ijms22084182>.
- Wexler, Randy, et al. "Cardiomyopathy: An Overview." *American Family Physician*, vol. 79, no. 9, May 2009, pp. 778–84.
- Yadgir, Simon, et al. "Global, Regional, and National Burden of Calcific Aortic Valve and Degenerative Mitral Valve Diseases, 1990–2017." *Circulation*, vol. 141, no. 21, May 2020, pp. 1670–80. ahajournals.org (Atypon), <https://doi.org/10.1161/CIRCULATIONAHA.119.043391>.
- Zhang, Deng-Yong, et al. "Follistatin Like 5 (FSTL5) Inhibits Epithelial to Mesenchymal Transition in Hepatocellular Carcinoma." *Chinese Medical Journal*, vol. 133, no. 15, Aug. 2020, pp. 1798–804. PubMed, <https://doi.org/10.1097/CM9.0000000000000847>.

Third Order Fibonacci Sequences By Stefan Maxim

Abstract:

I begin by introducing a special class of third-order Fibonacci sequences. I then investigate some connections with the standard Fibonacci sequences.

AMS SUBJECT CLASS [2020]: 11B37 (Primary), 11B39 (Secondary).

Keywords: Standard (second order) Fibonacci sequence, third order Fibonacci sequence, asymptotic behavior.

Introduction

By a Fibonacci sequence, we mean any sequence in which a certain term is the sum of the two preceding ones. So, if (F_n) is such a sequence, then its (second order) recurrence formula is written as follows:

$$(FS-2) F_{n+2} = F_{n+1} + F_n$$

(or, equivalently: $F_{n+2} - F_{n+1} - F_n = 0$), for all $n \in N := \{0, 1, \dots\}$.

To obtain the general solution of this recurrence, one starts with the condition that

$(F_n = \lambda^n; n \in N)$ should fulfill (FS). This yields that $\lambda^{n+2} - \lambda^{n+1} - \lambda^n = 0$, $\forall n \in N$; or, equivalently, $\lambda^2 - \lambda - 1 = 0$.

Let λ_1 and λ_2 be the roots of this equations; then,

$$\lambda_1 = (1/2)(1 + \sqrt{5}) \approx 1.632, \lambda_2 = (1/2)(1 - \sqrt{5}) = 1 - \lambda_1 \approx -0.632.$$

Note that, as a direct consequence of this,

$$0 < |\lambda_2| < 1 < |\lambda_1|; \text{ hence, } 0 < |\lambda_2/\lambda_1| < 1.$$

Furthermore, note that the relations above can be obtained without any approximation technique through the original Viète relations,

$$(V\text{-rela}) \lambda_1 + \lambda_2 = 1, \lambda_1 \lambda_2 = -1,$$

and are deductible through its direct factorization:

$$\lambda^2 - \lambda - 1 = (\lambda - \lambda_1)(\lambda - \lambda_2);$$

Using these facts, one can obtain the general solution for the recurrence (FS-2):

$$(FS-2\text{-sol}) F_n = C_1 \lambda_1^n + C_2 \lambda_2^n, n \in N,$$

where C_1 and C_2 are constants. To determine their values, one can fix the values of F_0 and F_1 and solve. To ensure consistency with its recursive property, the natural choices for F_0 and F_1 would be

$$(FS-01) F_0 = 1, F_1 = 1.$$

By substituting back into (FS-2), one obtains the system

$$(FC-2\text{-sys}) C_1 + C_2 = 1, C_1 \lambda_1 + C_2 \lambda_2 = 1;$$

The determinant of this system is

$$\Delta = \det ((1, 1); (\lambda_1, \lambda_2))^T = \lambda_2 - \lambda_1 \neq 0.$$

(Where \top signifies the transpose operation). Using the Cramer rule, we obtain a unique solution (D_1, D_2) of (FC-2-sys), expressed as

$$D_1 = (\lambda_2 - 1)/(\lambda_2 - \lambda_1), D_2 = (1 - \lambda_1)/(\lambda_2 - \lambda_1).$$

Consequently, the solution of (FS-2) +(FS-01) has the form

$$F_n = D_1 \lambda_1^n + D_2 \lambda_2^n, n \in N;$$

We then say that (F_n) is a *Fibonacci sequence of order two*; in short: a *2-order Fibonacci sequence*. Note that, by its very definition, all elements of the sequence (F_n) are natural numbers. For example, the first 10 elements of this sequence are

$$F_0 = 1, F_1 = 1, F_2 = 2, F_3 = 3, F_4 = 5, F_5 = 8, \\ F_6 = 13, F_7 = 21, F_8 = 34, F_9 = 55.$$

Now that we have established how we define the Fibonacci sequence, we can begin to discuss the asymptotic behavior of the sequence. However, before we begin, we must note some preliminary facts.

Let (U_n) and (V_n) both be sequences over $R_+^0 :=]0, \infty[$. We can say that

(ad-1) (U_n) is *asymptotically equivalent* to (V_n) , if $\lim_n(U_n/V_n) = \gamma$, for some $\gamma \in R_+^0$

(ad-2) (U_n) is *asymptotic inferior* to (V_n) , if $\lim_n(U_n/V_n) = 0$ (ad-3) (U_n) is *asymptotic superior* to (V_n) , if $\lim_n(U_n/V_n) = \infty$.

Proposition 1.1. *Let (U_n) be a geometric sequence with a common ratio of λ_1 over R_+^0 : $U_n = U_0 \lambda_1^n, n \in N$, where $U_0 > 0$.*

Then,

(11-1) *the Fibonacci sequence (F_n) is asymptotic equivalent with the progression (U_n) , since $\lim_n(F_n/U_n) = D_1/U_0 (> 0)$*

(11-2) *As a consequence, $\lim_n(F_{n+1}/F_n) = \lambda_1$; so that, at least asymptotically, (F_n) behaves like a geometric series with a common ratio of λ_1 .*

Proof. (I) According to our established definition, $\lim_n(F_n/U_n) = \lim_n(D_1 \lambda_1^n + D_2 \lambda_2^n)/(U_0 \lambda_1^n) = \lim_n[(D_1/U_0) + (D_2/U_0)(\lambda_2/\lambda_1)^n] = D_1/U_0$.

From here, if we denote $\xi = \lambda_2/\lambda_1$, we have (see above)

$$0 < |\xi| < 1; \text{ from which we observe, } \lim_n \xi^n = 0.$$

Thus, the first part is proved. Note that, as a direct consequence, $\lim_n(U_n/F_n) = 1/\lim_n(F_n/U_n) = 1/(D_1/U_0) = U_0/D_1$.

(II) From the previous stage, one has (under our notations)

$$\lim_n(F_{n+1}/F_n) = \lim_n[(F_{n+1}/U_{n+1})(U_{n+1}/U_n)(U_n/F_n)] = (D_1/U_0)(\lambda_1)(U_0/D_1) = \lambda_1;$$

thus, proving the second part. \square

Statement of the problem

To depict the function of 2-order Fibonacci sequence (F_n) , we can use it as the abstract model of numbering the pairs of rabbits [5] concerning a sequence of time units (months, for example). In

what follows, we will begin by making a certain modification to the aforementioned process. Specifically, we will denote G_n = the number of rabbit pairs at the n -th month, $n \in N$.

The conditions below are acceptable, for each $n \in N$:

(cond-1) in the time interval $]n, n+(3/2)[$, each rabbit pair (from the class of all G_n pairs) mates, resulting in a single rabbit

(cond-2) in the time interval $]n+(3/2), n+3[$, each pair (from the class of all G_n pairs) mates

(again), resulting in another single rabbit (cond-3) in the time interval $]n, n+3[$, each pair (from the class of all G_n pairs) mates with a new rabbit pair.

As a consequence of this, the abstract formula for our problem now becomes

$$(GS-3) \quad G_{n+3} = G_{n+2} + G_n$$

(or, equivalently:

$$G_{n+3} - G_{n+2} - G_n = 0), \quad n \in N.$$

Where the initial conditions imposed are (in accordance to our previous convention)

$$(GS-012) \quad G_0 = 1, G_1 = 1, G_2 = 1.$$

The sequence (G_n) fulfilling (GS-3) and (GS-012) will be referred to as a *Fibonacci sequence of order three*; in short: a *3-order Fibonacci sequence*.

Solution of (GS-3) + (GS-012)

To get the general solution for the recurrence (GS-3), one starts with the condition that $(G_n = \mu^n; n \in N)$ should fulfill (GS-3). This results in

$$(CE-3) \quad \mu^{n+3} - \mu^{n+2} - \mu^n = 0, \quad \forall n \in N; \text{ or, equivalently, } \mu^3 - \mu^2 - 1 = 0.$$

Through an analysis of the associated function $g : R \rightarrow R$ defined as $g(\mu) = \mu^3 - \mu^2 - 1$, $\mu \in R$, it follows that

(r-sol) its characteristic equation (CE-3) has only a single real solution μ_1 , which is approximated as $\mu_1 = 3/2$

(c-sol) the characteristic equation (CE-3) has two (conjugate) complex solutions $\mu_2 = \alpha + i\beta$, $\mu_3 = \alpha - i\beta$.

As for μ_1 's actual value, we observe that

$$g(1.46) = -0.019, \quad g(1.47) = 0.015;$$

so, a more precise value for the real solution would be $\mu_1 = 1.47$. That being said, for the sake of simplicity in our calculation, we will continue to refer to μ_1 's previous value $\mu_1 = 1.5 = 3/2$.

As for the real constants (α, β) appearing in our equations for (μ_2, μ_3) , they can be obtained using the Viète relations. To this end, let us begin with the factorization

$$\mu^3 - \mu^2 - 1 = (\mu - \mu_1)(\mu - \mu_2)(\mu - \mu_3).$$

A direct identification of these polynomials returns the relations $\mu_1 + \mu_2 + \mu_3 = 1$, $\mu_1\mu_2 + \mu_1\mu_3 + \mu_2\mu_3 = 0$, $\mu_1\mu_2\mu_3 = 1$.

From the first and third equations, we can conclude that

$\mu_2 + \mu_3 = 1 - \mu_1 = 1 - 3/2 = -1/2$, $\mu_2\mu_3 = 1/\mu_1 = 1/(3/2) = 2/3$; that is (using the representation consistent with how we defined μ_2 and μ_3)

$$2\alpha = -1/2, \alpha^2 + \beta^2 = 2/3.$$

The solution of this system is

$$(\text{solu}) \alpha = -1/4, \beta = \sqrt{\left(\frac{2}{3} - \frac{1}{16}\right)} = (1/4)\sqrt{\left(\frac{29}{3}\right)}$$

This, finally, gives us the general solution of our recurrence (GS-3)

$$(\text{GS-3-sol}) G_n = K_1\mu_1^n + K_2\mu_2^n + K_3\mu_3^n, n \in N.$$

where, for now, (K_1, K_2, K_3) are complex numbers. From here, the particular solution of the recurrence that fulfills (GS-012) can be obtained by passing our imposed conditions on this equation as follows:

$$(\text{GS-3-sys}) K_1\mu_1^0 + K_2\mu_2^0 + K_3\mu_3^0 = 1, \\ K_1\mu_1^1 + K_2\mu_2^1 + K_3\mu_3^1 = 1, K_1\mu_1^2 + K_2\mu_2^2 + K_3\mu_3^2 = 1.$$

The determinant of this system is

$$\Delta = \det((1, 1, 1), (\mu_1, \mu_2, \mu_3), (\mu_1^2, \mu_2^2, \mu_3^2))^T = \\ (\mu_2 - \mu_1)(\mu_3 - \mu_1)(\mu_3 - \mu_2) \neq 0.$$

By the Cramer rule, we therefore get a unique (complex) solution (K_1^*, K_2^*, K_3^*) of

(GS-3-sys), expressed as

$$K_1^* = [(\alpha - 1)^2 + \beta^2]/[(\alpha - \mu_1)^2 + \beta^2], \\ K_2^* = [(1 - \mu_1)(\alpha - 1 - i\beta)]/[(\alpha - \mu_1 + i\beta)(-2i\beta)] = S + iT, \\ K_3^* = [(1 - \mu_1)(\alpha - 1 + i\beta)]/[(\alpha - \mu_1 - i\beta)(2i\beta)] = S - iT,$$

where the real pair (S, T) depends on these equations. As a consequence, the solution of the problem (GS-3)+(GS-012) has the form

$$(\text{GS-3-sol}) G_n = K_1^*\mu_1^n + K_2^*\mu_2^n + K_3^*\mu_3^n, n \in N;$$

where (K_1^*, K_2^*, K_3^*) is the triple of complex numbers given above. However, to provide a more appropriate solution, we have to transform our complex solutions (μ_2, μ_3) of (EC-3) into polar form. To this end, remember that for every complex number, we have the polar representation

$$\alpha + i\beta = \rho(\cos\theta + i\sin\theta), \text{ where} \\ \rho = \sqrt{\alpha^2 + \beta^2}, \cos\theta = \alpha/\rho, \sin\theta = \beta/\rho,$$

From this, we will derive the well-known Moivre formula (see, for instance, Brown et al [1, Ch 14, Sect 14-5])

$$(\alpha + i\beta)^n = \rho^n (\cos n\theta + i\sin n\theta), n \in N.$$

Passing to the complex solutions of (CE-3), we have

$$-1/4 + i(1/4)\sqrt{29/3} = \rho(\cos\theta + i\sin\theta), \\ -1/4 - i(1/4)\sqrt{29/3} = \rho(\cos\theta - i\sin\theta), \text{ where} \\ \rho = \sqrt{2/3}, \cos\theta = (-1/4)\sqrt{3/2}, \sin\theta = (1/4)\sqrt{29/2}; \\ \text{hence, } \theta = \arccos(-1/4)\sqrt{3/2};$$

and this gives the expression of their powers

$$\begin{aligned}\mu_2^n &= (\alpha + i\beta)^n = \rho^n(\cos n\theta + i \sin n\theta), \\ \mu_3^n &= (\alpha - i\beta)^n = \rho^n(\cos n\theta - i \sin n\theta), \quad n \in N.\end{aligned}$$

By substituting this all back into the expression (GS-3-sol) of the particular solution, we obtain, after some small calculations

$$(GS-3-sol-real) \quad G_n = H_1\mu_1^n + H_2\rho^n \cos n\theta + H_3\rho^n \sin n\theta, \quad n \in N;$$

where (H_1, H_2, H_3) is a triple of real numbers. Note that, unfortunately, this representation is not exact since the values of (μ_1, μ_2, μ_3) we used are only loose approximations of the exact solutions of the characteristic equation (CE-3). That being said, by using some techniques of solving polynomial equations of degree 3, we can

write the exact representation of these solutions; namely

$$\begin{aligned}\mu_1 &= (1/3)(1 + \sqrt[3]{A} + \sqrt[3]{B}), \\ \mu_2 &= (1/3)[1 - P\sqrt[3]{A} - Q\sqrt[3]{B}], \quad \mu_3 = (1/3)[1 - Q\sqrt[3]{A} - P\sqrt[3]{B}],\end{aligned}$$

where, for simplicity, we denoted

$$\begin{aligned}A &= (1/2)(29 - 3\sqrt{93}), \quad B = (1/2)(29 + 3\sqrt{93}), \\ P &= (1/2)(1 - i\sqrt{3}), \quad Q = (1/2)(1 + i\sqrt{3}).\end{aligned}$$

As a result, an “exact” representation of the solutions, related to the one in (GS3-sol-real) are available. However, from a practical perspective, these expressions hold little value since they cannot be used in their current crude form. Thus, an approximation of these (as is shown in GS-3-sol-real) is preferable.

Finally, note that despite the complicated expression for our solution (littered with algebraic and trigonometric functions) all elements of the sequence (G_n) are still natural numbers. For example, the first 10 elements of this sequence are

$$G_0 = 1, G_1 = 1, G_2 = 1, G_3 = 2, G_4 = 3, G_5 = 4, G_6 = 6, G_7 = 9, G_8 = 13, G_9 = 19.$$

Lastly, an algebraic theory for these numbers is still waiting to be discovered. Hopefully, future developments in this field will yield some results that will change that.

Asymptotic properties

In the following, we will discuss an asymptotic behavior of the solution (GS-3sol-real) for the third order Fibonacci sequence (G_n)

Let us first return to the Fibonacci sequence of order two, (Fn).

Its characteristic equation is written as:

$$\begin{aligned}f(\lambda) &:= \lambda^2 - \lambda - 1 = 0; \text{ with the roots:} \\ \lambda_1 &= (1/2)(1 + \sqrt{5}) \approx 1.632, \quad \lambda_2 = (1/2)(1 - \sqrt{5}) \approx -0,632.\end{aligned}$$

On the other hand, for the Fibonacci sequence of order three (G_n) , its characteristic equation is written as

$$g(\mu) := \mu^3 - \mu - 1 = 0; \text{ with the roots: } \mu_1 \approx 1.5, \mu_2 = \alpha + i\beta, \mu_3 = \alpha - i\beta;$$

where (α, β) are given by the precise relations above. Note that, according to some classical results expressed in Stewart et al [3, Ch 4, Sect 4-1]

$$g'(\mu) = 3\mu^2 - 2\mu, \mu \in R; \text{ thus, } g \text{ is increasing on the interval } [2/3, \infty[.$$

This, along with

$$g(1) = -1, g(\lambda_1) = \lambda_1^3 - \lambda_1^2 - 1 = \lambda_1^2(\lambda_1 - 1) - 1 = (\lambda_1 + 1)(\lambda_1 - 1) - 1 = \lambda_1^2 - 2 = \lambda_1 - 1 > 0$$

tells us that

$$1 < \mu_1 < \lambda_1; \text{ thus } 0 < \rho = \sqrt{2/3} < 1 < \mu_1 < \lambda_1.$$

Proposition 4.1. *Let (V_n) be a geometric series with a common ratio of μ_1 over*

R_+^0 :

$$V_n = V_0 \mu_1^n, n \in N, \text{ where } V_0 > 0.$$

Then,

(41-1) *the 3-order Fibonacci sequence (G_n) is asymptotic equivalent with the progression (V_n) , since, as established in (ad-1), $\lim_n(G_n/V_n) = H_1/V_0 (> 0)$*

(41-2) *As a consequence, we obtain that $\lim_n(G_{n+1}/G_n) = \mu_1$; so that, asymptotically, (G_n) acts as a geometric series with a common ratio of μ_1*

(41-3) *the 3-order Fibonacci sequence (G_n) is asymptotic inferior with respect to the 2-order Fibonacci sequence (F_n) , in the sense that: $\lim_n(G_n/F_n) = 0$.*

Proof. (I) Denote, for simplicity,

$$\sigma = \rho/\mu_1; \text{ hence, } 0 < \sigma < 1, \lim_n \sigma^n = 0.$$

According to this definition, we get that

$$\lim_n(G_n/V_n) = \lim_n(H_1 \mu_1^n + H_2 \rho^n \cos n\theta + H_3 \rho^n \sin n\theta) / (V_0 \mu_1^n) = \lim_n[(H_1/V_0) + (H_2/V_0)\sigma^n \cos n\theta + (H_3/V_0)\sigma^n \sin n\theta] = H_1/V_0,$$

where we imply that

$$(\text{imp-1}) |\sigma^n \cos n\theta| \leq \sigma^n, \forall n, \text{ and } \lim_n \sigma^n = 0 \text{ imply } \lim_n \sigma^n \cos n\theta = 0,$$

$$(\text{imp-2}) |\sigma^n \sin n\theta| \leq \sigma^n, \forall n, \text{ and } \lim_n \sigma^n = 0 \text{ imply } \lim_n \sigma^n \sin n\theta = 0,$$

Thereby proving the first part. As a result, $\lim_n(V_n/G_n) = 1/\lim_n(G_n/V_n) = 1/(H_1/V_0) = V_0/H_1$.

(II) By a previous observation, all elements in the sequence (G_n) are natural numbers;

Specifically, positive numbers. In this case, we can conclude that $\lim_n(G_{n+1}/G_n) =$

$$\lim_n[(G_{n+1}/V_{n+1})(V_{n+1}/V_n)(V_n/G_n)] =$$

$$(H_1/V_0)(\mu_1)(V_0/H_1) = \mu_1; \text{ thereby proving the second part as well.}$$

(II) Remember that we introduced the geometric series

$$U_n = U_0 \lambda_1^n, n \in N.$$

Furthermore, we also showed that $\lim_n(F_n/U_n) = D_1/U_0 (> 0)$.

Finally, through a previous comparison relation, we showed that

$$0 < \mu_1 < \lambda_1; \text{ hence, } 0 < \eta := \mu_1/\lambda_1 < 1.$$

Combining all of these, we can conclude that

$$\lim_n(G_n/F_n) = \lim_n(G_n/V_n)(V_n/U_n)(U_n/F_n) = (H_0/V_0)(V_0/U_0)(\eta^n)(U_0/D_1) = 0;$$

□

As for the import of our results, it may prove useful when we wish to study our sequence (G_n) for large values of n . It should also be useful when evaluating the quotient sequence (G_n/F_n) for the same values of n . As a consequence of this last property, the sequence (G_n) may better describe the growth progression in question. Further aspects may be found in Mitchell et al [2], Waldschmidt [4], and the references therein.

Works Cited

1. R. G. Brown, M. P. Dolciani, R. H. Sorgenfrey, and R. B. Kane, Algebra and Trigonometry (Structure and Method), Book 2, McDougall Littell, Evanston-Boston-Dallas, 2000.
2. C. Mitchell, C. Potters, and J. Ecker Fibonacci sequence: definition, how it works, and how to use it, <https://www.investopedia.com/terms/f/fibonaccilines.asp>.
3. J. Stewart, D. Clegg, S. Watson, Calculus (Ninth Edition), Cengage Learning Inc., 2021.
4. M. Waldschmidt, Linear recurrence relations, Yogiakarta CIMPA School UGM, February 27, 2020, <https://www.imj-prg.fr/~michel.waldschmidt>.
5. Wikipedia, Fibonacci sequences, <https://en.wikipedia.org/wiki/Fibonacci-sequence>.



Source: National Gallery of Art.¹

Segregated Economies: The Lingering Effects of African American Exclusion on Southern Prosperity By Cooper Martin

Introduction

While slavery was abolished on December 6, 1865, President Lyndon B. Johnson did not sign the Civil Rights Act, ensuring integration of public spaces, until ninety-nine years later. Yet, the negative effects of those ninety-nine years still linger today. The abolition of slavery had immediate and profound impacts on the economies of Southern states. Emancipation impacted land utilization, profitability, economic growth, labor shortages, and overall agricultural productivity. However, the economic consequences stretched beyond Reconstruction. Even after the Civil War ended and slavery was officially abolished, the inability to fully integrate African Americans into society and the economy had arguably more severe impacts on Southern economies. Slavery was a major factor in the Southern states' economic growth, but the inability to integrate freedmen into society and the economy resulted in Southern states lagging behind the Northern states in many categories, including education, income, and overall economic growth. These trends can be traced back to the end of slavery and the systemic racism and segregation that followed.

Abstract

To fully understand the effect of the inability to gainfully integrate African Americans into the economy after the Civil War, the time span from early slavery and the peak of slavery until the Jim Crow period and Reconstruction era must be taken into consideration. The slavery period defines the true dominance slavery had on Southern economies and how its abolition negatively impacted Southern economies. Furthermore, the Jim Crow period and Reconstruction era demonstrate the segregation and barriers African Americans faced and the overall lack of integration. Lastly, the 2020s show how Southern economies continue to lag behind Northern

¹ National Gallery of Art. "Civil Rights Movement." Accessed April 16, 2024. <https://www.nga.gov/features/slideshows/civil-rights.html>

economies, stemming from poor incorporation of the large Black population. All of these periods are significant in understanding the overall impact slavery had on Southern economies.

Analysis

Prior to the abolition of slavery, the South relied heavily on slave labor for agricultural production. “In 1860, the South was still predominantly agricultural, highly dependent upon the sale of staples to a world market. By 1815, cotton was the most valuable export in the United States; by 1840, it was worth more than all other exports combined.”² Southern states were so invested in their agricultural sector because of their ability to mass produce cotton. They were so successful that “the Southern states produced two-thirds of the world's supply of cotton.”³ This success translated to immense profits for Southern states, primarily due to slavery. The ability to have free labor maximized profitability as paying wages was not required, allowing for significant Southern agricultural growth and financial success. The graph below shows the growth of the cotton industry due to slavery and the immediate drop off of cotton produced, from 1,019 thousand metric tons produced in 1861 to sixty-four thousand metric tons in 1864, demonstrating the impact slavery had on the large cotton market.

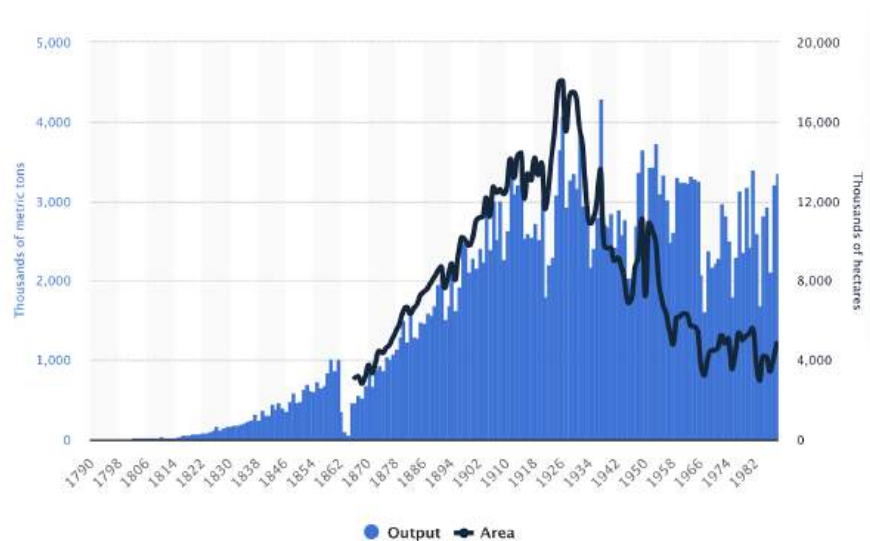


Fig 1: Output of cotton by thousands of metric tons from 1790 to 1982⁴

The purchase and use of slaves was similar to that of a commodity or stock. The purpose behind buying a slave was the expectation that a slave would yield higher profits with less time consumed. The price of slaves increased as the market for agriculture grew, therefore connecting slave prices with agriculture as a whole. The graph below displays the average price of a slave from 1800 until 1860.

² Benjamin T. Arrington. “Industry and Economy during the Civil War.” National Parks Service. Accessed April 16, 2024. <https://www.nps.gov/articles/industry-and-economy-during-the-civil-war.htm>.

³ Arrington “Industry and Economy Civil War.”

⁴ “Cotton Output and the Production Area in the United States from 1790 to 1988,” Statista, Accessed April 27, 2024, <https://www.statista.com/statistics/1070570/us-cotton-output-area-historical/>.

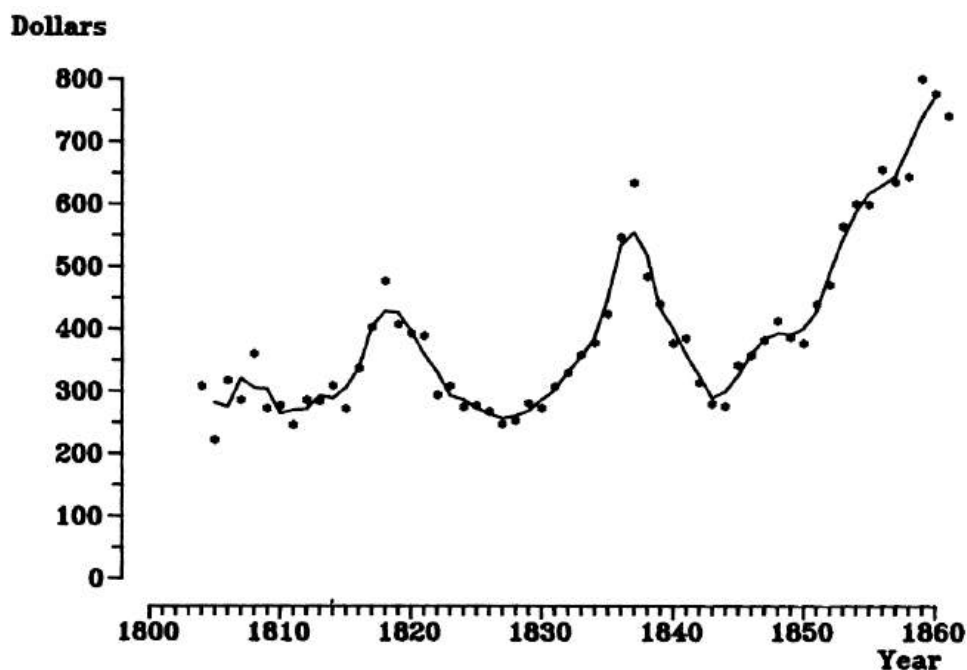


Fig 2: Average price of a slave from 1800 to 1860 in dollars⁵

Because slave labor was the basis for immense agricultural profits in the South, the farming markets followed the price of slaves. The figures below demonstrate the growth of the agricultural sector.

	<i>Percentage of Labor Force in Farming</i>	
	1800	1860
South	82	84
North and West	68	40

Fig 3: Percent of labor force in farming in 1800 and 1860⁶

⁵ Roger Ransom and Richard Sutch, "Capitalists without Capital: The Burden of Slavery and the Impact of Emancipation," *Agricultural History* 62, no. 3 (1988): 133–60. Accessed April 17, 2024, <http://www.jstor.org/stable/3743211>.

⁶ Stanley Lebergott, "Labor Force and Employment," 1800–1960. Edited by Dorothy S. Brady, *Studies in Income and Wealth*, NBER, 1966. Accessed April 17, 2024, <https://www.nber.org/system/files/chapters/c1567/c1567.pdf>.

TABLE 1
THE LABOR FORCE, BY INDUSTRY AND STATUS 1800-1960
(thousands)

	Total	Free	Slave	Agriculture	Fishing	Mining	Construction	Manufacturing			Trade	Transport		Teachers	Domestic Service
								Total Persons Engaged	Cotton Textile Wage Earners	Primary Iron and Steel Wage Earners		Ocean Vessels	Railway		
1800	1,900	1,370	530	1,400	5	10			1	1		40		5	40
1810	2,330	1,590	740	1,950	6	11		75	10	5		60		12	70
1820	3,135	2,185	950	2,470	14	13			12	5		50		20	110
1830	4,200	3,020	1,180	2,965	15	22			55	20		70		30	160
1840	5,660	4,180	1,480	3,570	24	32	290		72	24	350	95	7	45	240
1850	8,250	6,280	1,970	4,520	30	102	410	1,200	92	35	530	135	20	80	350
1860	11,110	8,770	2,340	5,880	31	176	520	1,530	122	43	890	145	80	115	600
1870	12,930			6,790	28	180	780	2,470	135	78	1,310	135	160	170	1,000
1880	17,390			8,920	41	280	900	3,290	175	130	1,930	125	416	230	1,130
1890	23,320			9,960	60	440	1,510	4,390	222	149	2,960	120	750	350	1,580
1900	29,070			11,680	69	637	1,665	5,895	303	222	3,970	105	1,040	436	1,800
1910	37,480			11,770	68	1,068	1,949	8,332	370	306	5,320	150	1,855	595	2,090
1920	41,610			10,790	53	1,180	1,233	11,190	450	460	5,845	205	2,236	752	1,660
1930	48,830			10,560	73	1,009	1,988	9,884	372	375	8,122	160	1,659	1,044	2,270
1940	56,290			9,575	60	925	1,876	11,309	400	485	9,328	150	1,160	1,086	2,300
1950	65,470			7,870	77	901	3,029	15,648	350	550	12,152	130	1,373	1,270	1,995
1960	74,060			5,970	45	709	3,640	17,145	300	530	14,051	135	883	1,850	2,489

^a Persons engaged (employees, wage earners, salaried, self-employed, and unpaid family workers), unless otherwise specified. Aged ten and over.

Fig 4: The labor force by industry and status from 1800 to 1960⁷

Figure two demonstrates the percentage of the labor force in farming. The high and growing percentage of the farming labor force in the South shows how important agriculture was for their economies compared to Northern states. This is because Northern states had a lower and decreasing percent of the labor force involved with farming. Figure three demonstrates the growth of the agriculture sector in more detail, showing that from the 11,110,000 employee population in 1860, 5,880,000 or 53% of people were involved in agriculture. The difference between this graph's 53% and figure 2's 84% could be explained by the definition of persons engaged and labor force, varying the count of people total. Overall, the use of low cost slave labor heavily impacted Southern economies, due to the large percent of slaves involved in agriculture and the impact of agriculture growth on Southern economies.

Because of how profitable slavery was for Southern states, the abolition of slavery and freedom of slaves negatively impacted Southern economies. The overall production and financial returns of many important resources, such as tobacco, cotton, rice, and wool, were greatly impacted by no longer being able to use free or low cost labor. In 1860, the production of tobacco in Kentucky was 108,126,840 pounds, whereas in 1880, the production was 105,305,860 pounds.⁸ While this change may not seem large, one must take into consideration the growth of the population over this period. In 1860, the population was 1,155,684, whereas in 1880, the

⁷ Lebergott, "Labor Force and Employment"

⁸ US Census Bureau. "Decennial Census Official Publications," Last updated December 16, 2021, <https://www.census.gov/programs-surveys/decennial-census/decade/decennial-publications.1860.html#list-tab-799609106>; US Census Bureau, "Decennial Census Official Publications," Last updated December 16, 2021, <https://www.census.gov/programs-surveys/decennial-census/decade/decennial-publications.1880.html#list-tab-799609106>

population was 1,648,690, representing a 493,042 person gain or 42.6% of the 1860 population.⁹ This increase, compared to the decrease in tobacco output, shows the impact that emancipation had on the agricultural sector and overall the economies for Southern states. However, tobacco is not the only resource that experienced a decrease in output due to emancipation.

The production of rice in South Carolina also dropped greatly. In 1860, South Carolina was the largest rice producer, producing 119,100,628 pounds of rice.¹⁰ However, in 1880, South Carolina only produced 32,304,825 pounds of rice.¹¹ The population increase during this period was 291,869 people.¹² This demonstrates the 73% decrease of rice output, while having a 41.5% increase in population. The decrease of overall agricultural output reflects the effects of the abolition of slavery as the ability to produce and yield profits from agriculture greatly decreased without free labor. This can be seen in figure 1 with the cotton output falling from about 1019 metric tons of cotton output in 1960 to just sixty-eight metric tons output in 1964. Not only did agricultural output decrease with emancipation, but the total value of farms decreased as well. As seen in the graph below, without slavery generating immense profits, the wealth of farms also dropped.

Table 2. Wealth and Farm Output on Southern Farms, Current Dollars, 1860 and 1870, Eight Countries

	<i>Number Farms Sampled</i>	<i>Average Value per Farm</i>			<i>Wealth-Output Ratio</i>	
		<i>Personal Estate</i>	<i>Total Wealth</i>	<i>Value of Output</i>	<i>Mean</i>	<i>Median</i>
1860:						
All Farms	570	14,576	24,881	2,478	8.9	7.3
Plantations	125	45,394	81,609	7,905	9.0	9.7
Other Slave Farms	262	9,114	13,345	1,215	9.7	9.0
Non-Slave Farms	183	1,346	2,648	577	4.6	3.9
1870:						
All Farms	755	772	2,766	1,470	2.5	1.4
White Farms	555	976	3,682	1,702	3.4	2.1
Black Farms	200	203	226	827	0.2	0.2

Fig 5: Wealth and farm output of Southern farms in 1860 and 1870¹³

The average total wealth of farms dropped nearly 22,000 dollars per farm, severely impacting Southern agriculture and economies.

⁹ US Census, "Decennial Census Official Publication," 1860; US Census, "Decennial Census Official Publications," 1880

¹⁰ US Census, "Decennial Census Official Publications," 1860

¹¹ US Census, "Decennial Census Official Publications," 1880

¹² US Census, "Decennial Census Official Publications," 1860; US Census, "Decennial Census Official Publications," 1880

¹³ Roger and Sutch, "Capitalists without Capital"

The barriers put in place against the full participation of freedmen in the economy after the abolition hindered African Americans from successfully integrating into society and the economy. Many barriers were put in place from around 1865 all the way to 1964. These included Black Codes, Jim Crow laws, lynchings, and the 3/4ths clause. These became the root of the disparity between Southern and Northern states, and resulted in Southern states struggling economically to this day. Southern states and citizens were still heavily racist and did not approve of freedom for African Americans. Instead of overcoming these ideas, they implemented many laws to continue to segregate Blacks from Whites. For example, during the Reconstruction Era, “the Equal Justice Initiative has documented more than 2,000 Black victims killed during the Reconstruction era, from 1865 to 1876. This is a staggering figure compared to the more than 4,400 victims documented for the 74-year era of racial terror lynching that spans 1877 to 1950.”¹⁴ In the eleven-year period after emancipation, nearly half as many African Americans were killed, compared to the seventy-four year period of lynching. This demonstrates the hostility towards emancipation and White people’s inability to overcome the racist barriers and integrate African Americans into society, resulting in negative impacts on the economies of Southern states.

Some civil rights activists, such as Booker T. Washington, a prominent author and educator fighting for equality, even went as far as accepting segregation as the only way to gain freedom and eventual equality for African Americans. His ideology to fight racism was “in all things that are purely social we can be as separate as the fingers, yet one as the hand in all things essential to mutual progress.”¹⁵ He believed that if Blacks and Whites worked together to improve the economy, the advancement of Southern states would be inevitable and enable Blacks to prove their value and strength through technical education and increased positive impact on the labor markets and economies as a whole. Overall, Washington was willing to initially accept segregation as long as it ultimately offered equality. However, his desire for equality did not come for many years, due to the continuous lack of integration and segregation.

Segregation began being implemented through Black Codes. Black Codes were laws passed in the South beginning in 1865 to hinder African Americans from getting jobs, owning property, and voting.¹⁶ Black Codes foreshadowed the much more serious discriminatory laws to come, the Jim Crow laws. Segregation was legally imposed in the South through Jim Crow laws. These laws, which were sets of rules and restrictions to segregate Blacks and Whites, came with significant penalties allowing for Southern governments to “impose legal punishments on people for consorting with members of another race.”¹⁷ These laws were implemented nearly

¹⁴ Equal Justice Initiative, “Documenting Reconstruction Violence,” June 15, 2020,

<https://eji.org/report/reconstruction-in-america/documenting-reconstruction-violence/#chapter-3-intro>.

¹⁵ Louis R. Harlan, “Booker T. Washington, 1856-1915,” Documenting the American South, Accessed April 18, 2024, <https://docsouth.unc.edu/fpn/washington/bio.html>.

¹⁶ National Geographic Society, “The Black Codes and Jim Crow Laws,” National Geographic Education, accessed May 2, 2024, <https://education.nationalgeographic.org/resource/black-codes-and-jim-crow-laws/#>.

¹⁷ National Parks Service, “Jim Crow Laws,” Accessed April 20, 2024, https://www.nps.gov/malu/learn/education/jim_crow_laws.htm.

everywhere from about 1880 until 1960.¹⁸ These places included barbers, burial grounds, and parks in Georgia, bathrooms and buses in Alabama, child custody and lunch counters in South Carolina, cohabitation in Florida, hospital entrances in Mississippi, parks in Georgia, telephone booths in Oklahoma, text books in North Carolina, theaters in Virginia, and numerous others.¹⁹ All of these places required Blacks and Whites to be separate and mandated that Whites could not serve Blacks.

Many barriers were put in place to segregate Black people and limit rights, such as the right to vote. The right to vote has always been challenging to possess and still is to this day. Women's right to vote was only passed in 1920 and Black men in 1869. In 1975, non-English speaking citizens were allowed to vote. Just two years ago, the United States Senate failed to pass an act that would restore many protections removed in the 2013 *Shelby County v. Holder* and the 2021 *Brnovich v. DNC* decisions.²⁰ These were not the only battles for voting rights fought. Before the end of slavery in 1865, African Americans, if they could vote, would not be counted as a whole vote. In 1787, the U.S. Government passed a clause that only counted slaves as $\frac{3}{5}$ of a vote. This was due to the fact that the South wanted to include slaves in voting, but the North did not want to include them. Due to the large population of slaves in the South, their votes greatly influenced national voting, allowing Southern states to dominate politics and pass any bills they wanted, possibly resulting in slavery never ending. The North and South agreed on a $\frac{3}{5}$ clause that counted each slave as a partial vote. This allowed Southern states to gain more representatives in voting without controlling elections as a whole, as seen in the figure below.

State	Non-slave Population	Slave Population	# of Reps. 1793	# of Reps. from Counting 3/5 of the Slaves 1793
Vt.	85,539	0	2	0
N.H.	141,727	158	4	0
Mass.	475,327	0	14	0
R.I.	67,877	948	2	0
Conn.	235,182	2,764	7	0
N.Y.	318,796	21,324	10	0
N.J.	172,716	11,423	5	0
Pa.	430,636	3,737	13	0
Del.	50,209	8,887	1	0
Md.	216,692	103,036	8	2
Va.	454,983	292,627	19	5
Ky.	61,247	12,430	2	0
N.C.	293,179	100,572	10	2
S.C.	141,979	107,094	6	2
Ga.	53,284	29,264	2	0
Total	3,199,373	694,264	105	11

Fig 6: Voting representatives by state in 1793 including the $\frac{3}{5}$ clause²¹

¹⁸ “Examples of Jim Crow Laws - Oct. 1960 - Civil Rights,” Jim Crow Museum, accessed April 30, 2024, <https://jimcrowmuseum.ferris.edu/links/misclink/examples.htm>.

¹⁹ National Parks Service, “Jim Crow Laws.”

²⁰ Rotimi Adeoye and Johanna Silver, “The Fight for Voting Rights: How the Past Informs the Current Discriminatory Landscape,” American Civil Liberties Union, February 23, 2022, <https://www.aclu.org/news/voting-rights/the-fight-for-voting-rights-how-the-past-informs-the-current-discriminatory-landscape>.

²¹ “The Impact of the Three-Fifths Clause on Representation in U.S. House of Representatives, 1793,” Center for the Study of the American Constitution, February 12, 2021, <https://csac.history.wisc.edu/2021/02/12/the-impact-of-the-three-fifths-clause-on-representation-in-u-s-house-of-representatives-1793/>.

Once the 15th Amendment clause ended, African American men (women were not yet allowed to vote in the United States) still had trouble voting, as they were required to adhere to numerous restrictions in order to vote, many of which a majority of the population lacked. Each Southern state differed in their voting rules; however, they were all unified in making it as difficult as possible for Blacks to vote. For example, in Alabama, Louisiana, Mississippi, North Carolina, South Carolina, Virginia, and Tennessee, an African American was required to read and write in order to register to vote.²² However, before emancipation, Whites did not grant African Americans access to education in order to maintain the racial barrier and keep their superiority.²³ This made it nearly impossible for African Americans to gain the necessary education to vote. Not only was education necessary to gain voting rights, but also in order to develop the race as a whole. Historian and civil rights activist, W. E. B. Du Bois, believed higher education was the only way to justify Blacks' integration into the economy and society. He states, "Men we shall have only as we make manhood the object of the work of the schools—intelligence, broad sympathy, knowledge of the world that was and is, and of the relation of men to it—this is the curriculum of that Higher Education which must underlie true life."²⁴ This idea of gaining higher education in order to cohere with society was difficult to achieve due to racist education policies. These policies resulted in lower overall education in the South due to a portion of the population not attending school during reconstruction. Those policies impacted funding, importance of attending school, level of education and many more factors contributing to differing education rates.

Even after emancipation, Blacks often faced challenges going to school, resulting in much lower education rates, as seen in the figures below.

²² African American Pamphlet Collection, "To the Colored Men of Voting Age in the Southern States," Philadelphia: Press of E.A. Wright, 190-?, Pdf, Accessed April 18, 2024, <https://www.loc.gov/item/92838850/>.

²³ National Park Service, "African Americans and Education During Reconstruction: The Tolson's Chapel Schools," U.S. Department of the Interior, Accessed April 18, 2024, <https://www.nps.gov/articles/african-americans-and-education-during-reconstruction-the-tolson-s-chapel-schools.htm#:~:text=Before%20Emancipation%2C%20whites%20generally%20denied,the%20former%20slave%2Dholding%20states.>

²⁴ W.E.B. Du Bois, "The Talented Tenth," from *The Negro Problem: A Series of Articles by Representative Negroes of To-day* (New York, 1903).

Figure 1.-- Percent of 5- to 19-year-olds enrolled in school, by race: 1850 to 1991

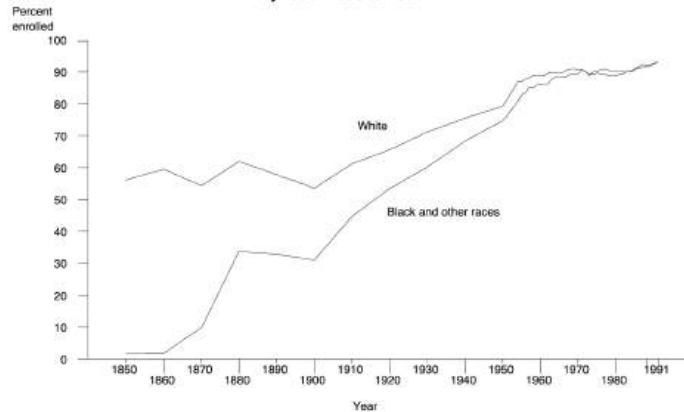


Fig 7: Percent of 5 to 19 year olds enrolled in school by race in 1850 to 1991²⁵

Table 2.—School enrollment of 5– to 19–year-olds per 100 persons, by sex and race: 1850 to 1991

Year	Both sexes			Male			Female		
	Total	White	Black and other races ¹	Total	White	Black and other races ¹	Total	White	Black and other races ¹
1	2	3	4	5	6	7	8	9	10
1850	47.2	56.2	1.8	49.6	59.0	2.0	44.8	53.3	1.8
1860	50.6	59.6	1.9	52.6	62.0	1.9	48.5	57.2	1.8
1870	48.4	54.4	9.9	49.8	56.0	9.6	46.9	52.7	10.0
1880	57.8	62.0	33.8	59.2	63.5	34.1	56.5	60.5	33.5
1890	54.3	57.9	32.9	54.7	58.5	31.8	53.8	57.2	33.9
1900 ²	50.5	53.6	31.1	50.1	53.4	29.4	50.9	53.9	32.8
1910 ²	59.2	61.3	44.8	59.1	61.4	43.1	59.4	61.3	46.6
1920 ²	64.3	65.7	53.5	64.1	65.6	52.5	64.5	65.8	54.5
1930 ^{2,3}	69.9	71.2	60.3	70.2	71.4	59.7	69.7	70.9	60.8
1940	74.8	75.6	68.4	74.9	75.9	67.5	74.7	75.4	69.2
1950	78.7	79.3	74.8	79.1	79.7	74.7	78.4	78.9	74.9

Fig 8: School enrollment of 5 to 19 year olds per 100 persons by sex and race from 1850 to 1950²⁶

African Americans lacked proper education due to slavery, resulting in a majority of the population not meeting voting requirements. During slavery, almost no African Americans were enrolled in school or receiving education at all. Furthermore, after emancipation, many African Americans were still denied education up until 1872 when only a handful of kids were enrolled in school.²⁷ The restrictions placed on Black education resulted in lower overall education levels for southern states due to the lack of integration of Blacks into education. However, education was not the only restriction placed on Blacks. In some states, including Alabama, Louisiana, and South Carolina, African Americans needed to own land in order to vote, which at the time was

²⁵ Thomas D. Snyder, ed, “120 Years of American Education: A Statistical Portrait,” National Center for Educational Statistics, January 1993, <https://nces.ed.gov/pubs93/93442.pdf>.

²⁶ Snyder, “120 Years of American Education.”

²⁷ “African Americans and Education During Reconstruction: The Tolson’s Chapel Schools,” National Parks Service, accessed May 2, 2024, <https://www.nps.gov/articles/african-americans-and-education-during-reconstruction-the-tolson-s-chapel-schools.htm>

nearly impossible.²⁸ “Mississippi and many other states passed laws that obstructed Black ownership of real estate.”²⁹ These Southern states set many restrictions for African Americans to vote, only to make the restrictions nearly impossible to meet.

In 1964, African Americans were legally accepted into society and the economy. This was due to the Civil Rights Act of 1964. This act overturned Jim Crow laws and allowed African Americans equality in restaurants, workplaces, education, and all other public facilities.³⁰ While equality may have been granted to African Americans in 1964, the inability to integrate African Americans into Southern societies and economies is still present today, impacting many statistics, such as lower education, income, and overall socioeconomic synchronization. This can be seen through maps of levels of education and state mathematics, science, reading, and writing tests, demonstrating how Southern states currently lag behind other states in overall education due to segregation throughout history.

The education below high school from 2017 to 2021 has a dramatic line between Northern and Southern states, with the Southern states having larger percentages of 25+ year olds lacking education beyond high school. States, such as Mississippi, Louisiana, Arkansas, Texas, and Alabama, have percentages as high as 15.8. Other Southern states, including North and South Carolina, Tennessee, Georgia, and Florida, have percentages as high as 11.9. The abnormally high percentages of less educated individuals stems from the unwillingness to integrate African Americans into society and especially providing education in and around the Reconstruction Era. However, one must take into consideration other factors, including attitude toward education, school attendance, class rigor, overall involvement with education, background of agricultural education, taxpayers’ money in the south put towards education, and the Black to White student ratio. One study takes many of these factors into consideration and concluded, “although racial differences in literacy and school attendance decreased in successive cohorts, the racial gap in average educational attainment remained persistently large. And, although successive cohorts of Black children were educated in better and better schools, racial differences in the quality of schooling followed a U-shaped pattern over time.”³¹ This study demonstrates the correlation between lack of integration post slavery and the lower education rates while still considering many possible factors.

²⁸ Pamphlet Collection, “Colored Men Voting Southern.”

²⁹ Spencer D. Wood, “Landowners, African American,” Mississippi Encyclopedia, July 11, 2017, <https://mississippiencyclopedia.org/entries/african-american-landowners/>.

³⁰ “The Civil Rights Act of 1964: A Long Struggle for Freedom,” Library of Congress, accessed May 2, 2024, <https://www.loc.gov/exhibits/civil-rights-act/epilogue.html>

³¹ Robert A. Margo, “Race and Schooling in the South: A Review of the Evidence,” in *National Bureau of Economic Research* (Chicago, IL: University of Chicago Press, 1990), 6–32, <https://www.nber.org/system/files/chapters/c8792/c8792.pdf>.

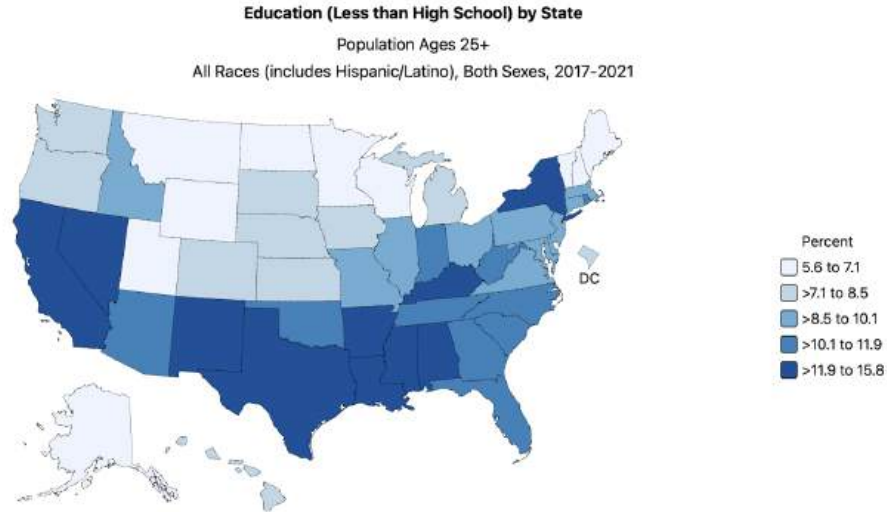


Fig 9: Education less than high school by state including ages 25+ and all races and sexes

Diving more specifically into the education sector, statistics such as grade 8 average scale scores for writing, reading, science, and mathematics, shows that Southern states severely fall behind those of the Northern states. In all 4 figures below, states including North and South Carolina, Tennessee, Georgia, Alabama, Mississippi, Louisiana, Arkansas, and Oklahoma, significantly fall behind in all of these scores, once again stemming from the continued reluctance to integrate African American students into society and education.

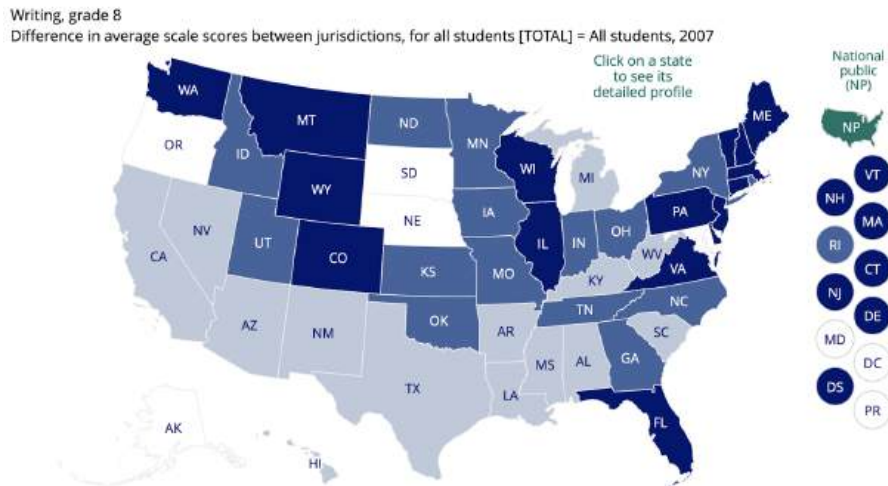


Fig 10: Average grade 8 writing scale scores for all students in 2007 between jurisdictions³²

³² The Nation's Report Card, "State Performance Compared to the Nation: Grade 8 Writing, 2007," Accessed April 20, 2024.
<https://www.nationsreportcard.gov/profiles/stateprofile?chort=2&sub=WRI&sj=AL&sfj=NP&st=MN&year=2007R3>

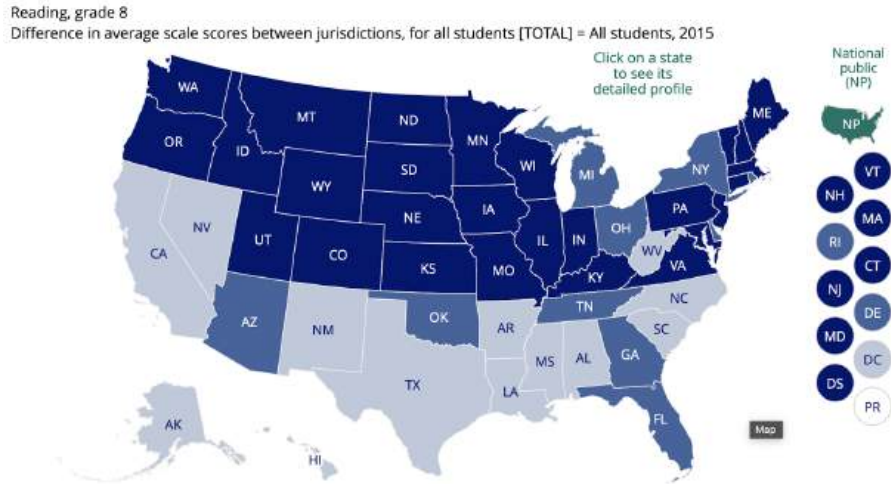


Fig 11: Average grade 8 reading scale scores for all students in 2015 between jurisdictions³³

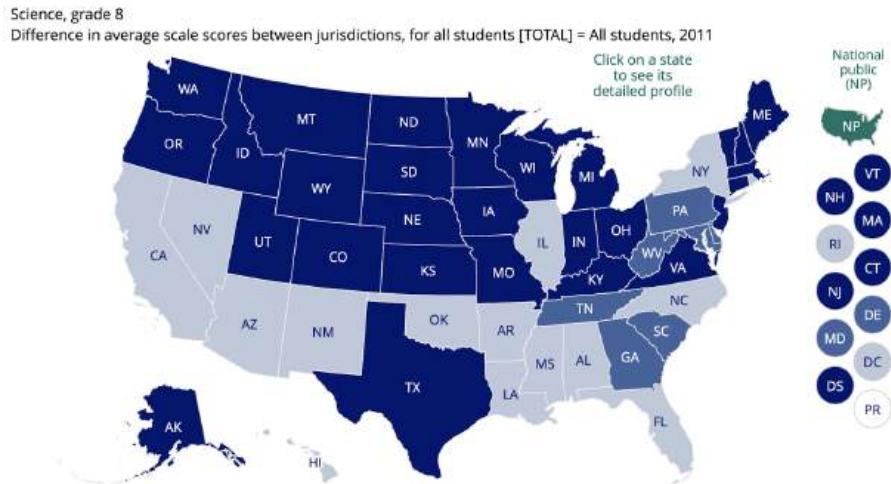


Fig 12: Average grade 8 science scale scores for all students in 2011 between jurisdictions³⁴

³³ The Nation's Report Card, "State Performance Compared to the Nation: Grade 8 Reading, 2015," Accessed April 20, 2024.

<https://www.nationsreportcard.gov/profiles/stateprofile?chort=2&sub=RED&sj=AL&sfj=NP&st=MN&year=2015R3>

³⁴ The Nation's Report Card, "State Performance Compared to the Nation: Grade 8 Science, 2011," Accessed April 20, 2024.

<https://www.nationsreportcard.gov/profiles/stateprofile?chort=2&sub=SCI&sj=AL&sfj=NP&st=MN&year=2011R3>

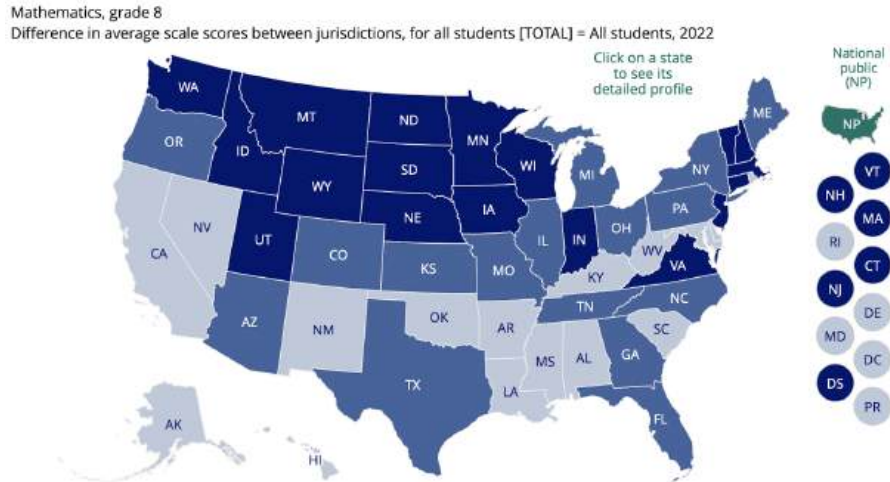
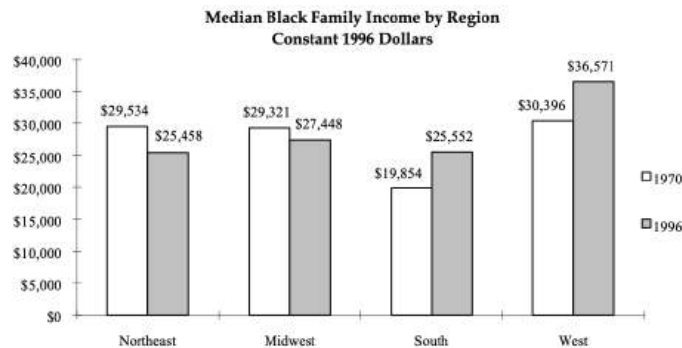


Fig 13: Average grade 8 mathematics scale scores for all students in 2022 between jurisdictions³⁵

The unwillingness to integrate African Americans into society not only impacted education, but also caused income disparity and poverty. The Southern states had a much lower median family income due to African Americans not being accepted or treated equally in society during Reconstruction. This prevented them from being able to get a strong education and jobs, resulting in a lower average income still seen today. In addition, this correlates with greater poverty rates due to the lack of money to support families. However, the Southern states standing out in statistics, such as poverty and income is not new. Even in 1970 and 1996, Southern states were falling behind other regions in median Black family income. This shows the persistent gap caused by the lack of integration of African Americans into the economy and the toll that took on Southern states, lowering the economic status of that region for the years to follow. This can be seen in the figures below with Southern states lagging behind other regions from 1977 through 2021.



³⁵ The Nation's Report Card, "State Performance Compared to the Nation: Grade 8 Mathematics, 2022," Accessed April 20, 2024.
<https://www.nationsreportcard.gov/profiles/stateprofile?chort=2&sub=MAT&sj=AL&sfj=NP&st=MN&year=2022R>

Fig 16: Median Black family income by region in 1970 and 1996 measured by 1996 dollars³⁶

Southern states fell behind every other region in median income for African Americans. This was due to the lack of integration into the economy after slavery, resulting in an overall lower median income for Blacks. This brought the whole state's income rates down and poverty rates up, as seen in the more recent figures below.

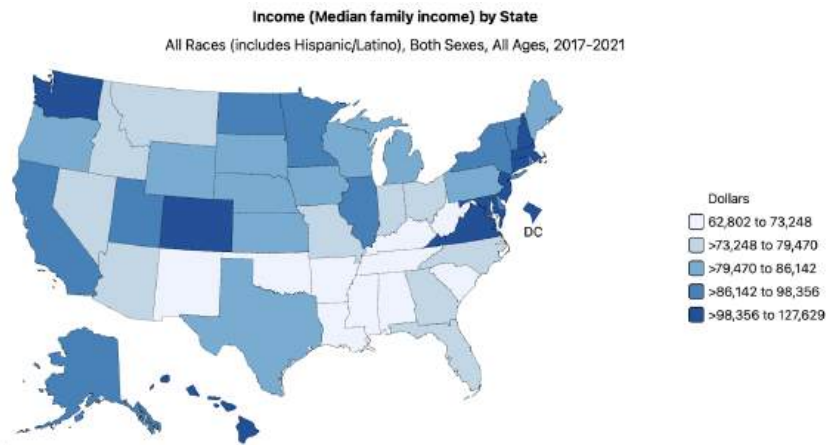


Fig 14: Median family income by state including all races, sexes, and ages from 2017 to 2021³⁷

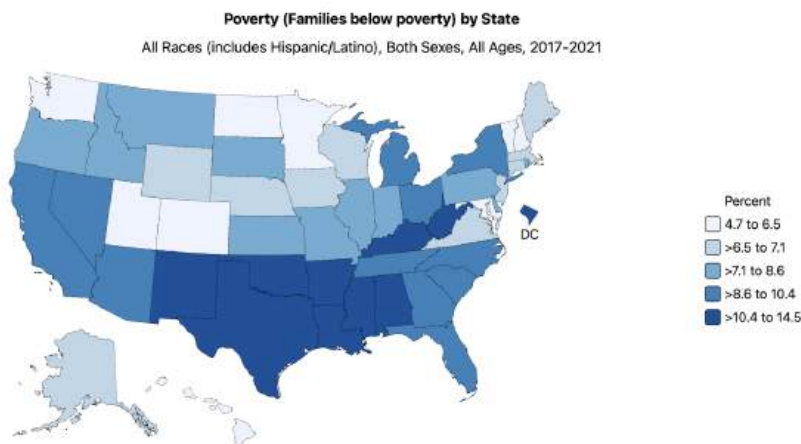


Fig 15: Percent of families living below the poverty line by state including all races, sexes, and ages from 2017 to 2021³⁸

³⁶ "Income and Wealth in the South," MDC, May 1988, <https://www.mdcinc.org/wp-content/uploads/2017/11/income.pdf>.

³⁷ National Institute on Minority Health and Health Disparities, "Income Map," Accessed April 20, 2024, https://hdpulse.nimhd.nih.gov/data-portal/social/map?socialtopic=030&socialtopic_options=social_6&demo=00010&demo_options=income_3&race=00&race_options=race_7&sex=0&sex_options=sexboth_1&age=001&age_options=ageall_1&statefips=00&statefips_options=area_states.

³⁸ National Institute on Minority Health and Health Disparities, "Poverty Map," Accessed April 20, 2024, https://hdpulse.nimhd.nih.gov/data-portal/social/map?socialtopic=080&socialtopic_options=social_6&demo=00007&demo_options=poverty_3&race=00&race_options=race_7&sex=0&sex_options=sexboth_1&age=001&age_options=ageall_1&statefips=00&statefips_options=area_states

States, including Oklahoma, Arkansas, Louisiana, Mississippi, Alabama, South Carolina, Kentucky, and Tennessee, all have a max income of \$73,248, putting them behind nearly every other state on median family income. Due to the lower income in these states, all but Tennessee and South Carolina lead the nation in poverty rates, with their max being 14.5%. The inability to integrate African Americans into society, stemming back to the abolition of slavery and the laws created to segregate African Americans from society and the economy that followed, resulted in lower education rates and median income, as well as higher poverty rates, representing an overall poor socioeconomic synchronization with the rest of the nation.

Conclusion

As detailed above, the abolition of slavery in 1856 had long-lasting impacts on the economies of Southern states, extending well beyond the immediate aftermath of slavery. While the loss of slave labor and fall of the strong Southern agricultural sector negatively impacted economies in the short term, segregation and the lack of integration of African Americans cost them employment opportunities, voting rights, education, and land ownership. This significantly hurt the Southern economies in the long run, even spanning to today and likely into the foreseeable future. The racist barriers and laws originating during Reconstruction and popularized thereafter caused prolonged racism and segregation well past emancipation. It was not until 99 years later when the Civil Rights Act was signed did segregation legally and fully end. This left Southern economies lagging behind Northern economies in education and income equality, but surpassing them in poverty rates. This obvious difference in important economic indicators can be traced directly back to slavery and the South's inability to overcome racist societal norms and accept African Americans as equal citizens of our country. Overcoming the long history of exclusion and segregation is still a real ongoing challenge, which is likely to continue to impact the economic and societal prosperity of the South until African Americans are truly equal citizens in our country.

Works Cited

- Adeoye, Rotimi, and Johanna Silver. "The Fight for Voting Rights: How the Past Informs the Current Discriminatory Landscape." American Civil Liberties Union, February 23, 2022. <https://www.aclu.org/news/voting-rights/the-fight-for-voting-rights-how-the-past-informs-the-current-discriminatory-landscape>.
- African American Pamphlet Collection. "To the Colored Men of Voting Age in the Southern States." Philadelphia: Press of E.A. Wright, 190-?, Pdf. Accessed April 18, 2024 <https://www.loc.gov/item/92838850/>.
- "African Americans and Education During Reconstruction: The Tolson's Chapel Schools." National Parks Service. Accessed May 2, 2024. <https://www.nps.gov/articles/african-americans-and-education-during-reconstruction-the-tolson-s-chapel-schools.html>.
- Arrington, Benjamin T. "Industry and Economy during the Civil War." National Parks Service. Accessed April 16, 2024. <https://www.nps.gov/articles/industry-and-economy-during-the-civil-war.htm>.
- "Cotton Output and the Production Area in the United States from 1790 to 1988." Statista. Accessed April 27, 2024. <https://www.statista.com/statistics/1070570/us-cotton-output-area-historical/>.
- Equal Justice Initiative. "Documenting Reconstruction Violence." June 15, 2020. <https://eji.org/report/reconstruction-in-america/documenting-reconstruction-violence/#chapter-3-intro>.
- "Examples of Jim Crow Laws - Oct. 1960 - Civil Rights." Jim Crow Museum. Accessed April 30, 2024. <https://jimcrowmuseum.ferris.edu/links/misclink/examples.htm>.
- Harlan, Louis R. "Booker T. Washington, 1856-1915." Documenting the American South, Accessed April 18, 2024, <https://docsouth.unc.edu/fpn/washington/bio.html>.
- "Income and Wealth in the South." MDC, May 1988. <https://www.mdcinc.org/wp-content/uploads/2017/11/income.pdf>.
- Lebergott, Stanley. "Labor Force and Employment." 1800-1960. Edited by Dorothy S. Brady. Studies in Income and Wealth. NBER, 1966. Accessed April 17, 2024. <https://www.nber.org/system/files/chapters/c1567/c1567.pdf>.
- Margo, Robert A. "Race and Schooling in the South: A Review of the Evidence." In National Bureau of Economic Research, 6-32. Chicago, IL: University of Chicago Press, 1990. <https://www.nber.org/system/files/chapters/c8792/c8792.pdf>.
- National Gallery of Art. "Civil Rights Movement." Accessed April 16, 2024. <https://www.nga.gov/features/slideshows/civil-rights.html>
- National Geographic Society. "The Black Codes and Jim Crow Laws." National Geographic Education. Accessed May 2, 2024. <https://education.nationalgeographic.org/resource/black-codes-and-jim-crow-laws/#>.
- National Institute on Minority Health and Health Disparities. "Income Map." Accessed April 20, 2024. https://hdpulse.nimhd.nih.gov/data-portal/social/map?socialtopic=030&socialtopic_options=social_6&demo=00010&demo_options=income_3&race=00&race_options=race_7&sex=0&sex_options=sexboth_1&age=001&age_options=ageall_1&statefips=00&statefips_options=area_states.
- National Institute on Minority Health and Health Disparities. "Poverty Map." Accessed April 20, 2024. https://hdpulse.nimhd.nih.gov/data-portal/social/map?socialtopic=080&socialtopic_options=social_6&demo=00007&demo_options=poverty_3&race=00&race_options=race_7&sex=0&sex_options=sexboth_1&age=001&age_options=ageall_1&statefips=00&statefips_options=area_states

- National Park Service “African Americans and Education During Reconstruction: The Tolson’s Chapel Schools.” U.S. Department of the Interior. Accessed April 18, 2024.
<https://www.nps.gov/articles/african-americans-and-education-during-reconstruction-the-tolson-s-chapel-schools.htm#:~:text=Before%20Emancipation%2C%20whites%20generally%20denied,the%20former%20slave%2Dholding%20states.>
- National Parks Service. “Jim Crow Laws.” National Parks Service. Accessed April 20, 2024.
https://www.nps.gov/malu/learn/education/jim_crow_laws.htm.
- Ransom, Roger, and Richard Sutch. “Capitalists without Capital: The Burden of Slavery and the Impact of Emancipation.” *Agricultural History* 62, no. 3 (1988): 133–60.
<http://www.jstor.org/stable/3743211>.
- Snyder, Thomas D., ed. “120 Years of American Education: A Statistical Portrait.” National Center for Educational Statistics, January 1993. <https://nces.ed.gov/pubs93/93442.pdf>.
- “The Civil Rights Act of 1964: A Long Struggle for Freedom.” Library of Congress. Accessed May 2, 2024. <https://www.loc.gov/exhibits/civil-rights-act/epilogue.html>.
- “The Impact of the Three-Fifths Clause on Representation in U.S. House of Representatives, 1793.” Center for the Study of the American Constitution, February 12, 2021.
<https://csac.history.wisc.edu/2021/02/12/the-impact-of-the-three-fifths-clause-on-representation-in-u-s-house-of-representatives-1793/>.
- The Nation’s Report Card. “State Performance Compared to the Nation: Grade 8 Mathematics, 2022.” Accessed April 20, 2024.
<https://www.nationsreportcard.gov/profiles/stateprofile?chort=2&sub=MAT&sj=AL&sfj=NP&st=MN&year=2022R3>
- The Nation’s Report Card. “State Performance Compared to the Nation: Grade 8 Reading, 2015.” Accessed April 20, 2024.
<https://www.nationsreportcard.gov/profiles/stateprofile?chort=2&sub=RED&sj=AL&sfj=NP&st=MN&year=2015R3>
- The Nation’s Report Card. “State Performance Compared to the Nation: Grade 8 Science, 2011.” Accessed April 20, 2024.
<https://www.nationsreportcard.gov/profiles/stateprofile?chort=2&sub=SCI&sj=AL&sfj=NP&st=MN&year=2011R3>
- The Nation’s Report Card. “State Performance Compared to the Nation: Grade 8 Writing, 2007.” Accessed April 20, 2024.
<https://www.nationsreportcard.gov/profiles/stateprofile?chort=2&sub=WRI&sj=AL&sfj=NP&st=MN&year=2007R3>
- US Census Bureau. “Decennial Census Official Publications.” Last updated December 16, 2021.
<https://www.census.gov/programs-surveys/decennial-census/decade/decennial-publications.1860.html#list-tab-799609106>
- US Census Bureau. “Decennial Census Official Publications.” Last updated December 16, 2021.
<https://www.census.gov/programs-surveys/decennial-census/decade/decennial-publications.1880.html#list-tab-799609106>
- W.E.B. Du Bois. “The Talented Tenth.” from *The Negro Problem: A Series of Articles by Representative Negroes of To-day* (New York, 1903).
- Wood, Spencer D. “Landowners, African American.” *Mississippi Encyclopedia*. July 11, 2017.
<https://mississippiencyclopedia.org/entries/african-american-landowners/>.

The Impact of GenAI on International Relations, and Proposed Countermeasures to be Employed by the Foreign Service to Adapt in a New Geopolitical Paradigm

By Sydney Garber

Abstract

The practice of international relations, as carried out by The Foreign Service, requires a delicate hand. At its core, it is about navigating a wide range of political, legal, economic, humanitarian, and other variables to help smooth interactions among sovereign states and place one's country in the best-possible position to flourish going forward. A new element is now being added to the mix, which will radically alter how these objectives are accomplished and deliver new obstacles to overcome almost immediately. This paper will dig into the specific challenges presented by Generative Artificial Intelligence (GenAI) to diplomats and supporting servicepeople, and it will offer evidence-based recommendations on steps that can be taken by the Foreign Service to thrive in a future that will be dominated by this black box technology.

Introduction

Technology innovation has influenced foreign affairs throughout history. In the last millennium, the introduction of the printing press, photography, and radio not only made it easier to share information, but also misinformation at scale. In the years leading up to World War II, for instance, German propaganda minister Goebbels emphasized the importance of a recent technical advance saying he wanted “a radio that reaches across our borders to give the world a picture of our character, our life, and our work” (Goebbels).

Today, *Information Technology* (IT) is the form of innovation most impacting our international relations. China's Digital Silk Road is a clear example of how tech is currently shaping global perspectives. By delivering infrastructure and standards across borders, the country is now in a strong position to spread its ideals. This push to expand their digital influence can possibly impact foreign states in many ways, including compromising monetary and security-related data (Carr).

Beware the Black Box: GenAI

The next wave of IT – led by GenAI – is here, and it will present the single greatest challenge to The Foreign Service in the future. Sometimes called “black box” technology because it can mimic human thinking without complete transparency, it will test existing diplomatic practices, expose new risks, and require altogether different approaches to cooperation and securing our secrets.

Deepfakes are one obvious example of how this innovation will affect international relations. Because GenAI can produce cheap, convincing content, it is being used as a tool of mass deception. For example, a recent video appeared of Ukrainian President Zelenskiy asking his soldiers to lay down their weapons and return to their families (Milmo). And in March 2023, as some top U.S. banks were failing, a clip of President Biden appeared where he expresses

concern about the country's overall banking stability (McKenzie). This perceived news can impact markets, trade agreements, and conflict resolution. Perhaps that is why the U.S. Department of Homeland Security said that deepfakes will increasingly “pose a threat for individuals and industries, including potential largescale impacts to nations, governments, businesses, and society” (U.S. Dept. of Homeland Security).

Another risk posed by this emerging technology lies in the realm of *cybersecurity*. Hackers have been using ChatGPT to write more effective malware, data-scraping practices are creating new privacy concerns, and phishing attacks are growing more sophisticated with the aid of this enhanced intelligence (Castro). Security risks even apply to government employees using GenAI-powered tools as part of their daily work. For instance, the state of California published an analysis in late 2023 that said, “GenAI systems can be susceptible to unique attacks and manipulations, such as poisoning of AI training datasets, evasion attacks, and interference attacks” (State of California). Because of the potential consequences, the notion of cyber diplomacy is gaining traction and was highlighted at one of last year's top cybersecurity conferences by an ambassador for the U.S. Bureau of Cyberspace and Digital Policy (Fick). The rise of cyber technologies – now fueled by GenAI – is in fact so closely connected to international relations that the American Foreign Service Association (AFSA) described it as, “one of the central foreign policy issues of the 21st century” (Painter).

A final critical challenge posed to The Foreign Service by this innovation comes from how it may be used as a *tool of influence*. While its benefits are numerous, GenAI also has many shortcomings to be aware of such as its tendency to reflect existing biases from the data it mines – and even amplify those stereotypes (Nicoletti). A speaker at a U.S. Advisory Commission on Public Diplomacy event reinforced this saying, “ChatGPT is a large language model, it's not sentient, and it's not conscious...this means that it will confidently generate content that sounds authoritative even if it's inaccurate, or inappropriate, or biased” (ACPD). Adding to this caution, the Global Diplomatic Forum noted, “while AI can provide valuable insights, complex decision-making in diplomacy often involves ethical considerations, moral judgments, and a deep understanding of human values...these aspects, integral to diplomatic decision making, are beyond the capabilities of AI” (Global Diplomatic Forum). Failing to be conscious of these warnings can lead to breakdowns in negotiations or escalated political tensions.

Proposed Countermeasures: The “Four-S” Strategy

GenAI is here to stay. The good news is that it can benefit The Foreign Service in many ways, but its use – without guardrails – can also lead to serious consequences. Following are a few ways to help put protections into place, each of which present their own unique challenges:

1) Skepticism: Trust but Verify

AI can be complex and difficult to validate. But at the same time, this cutting-edge innovation can separate signal from noise and deliver important insights (Lawrence). The key to taking advantage of this powerful tool in diplomacy is a healthy dose of skepticism. Users must

question assumptions, look for inaccuracies, and use other data (when available) to form a complete picture. This can only be accomplished, particularly in the short term, with training. All levels of staff should be educated on the pros and cons of using GenAI, as well as how it can lead to security risks and assessments that don't account for culture or context.

2) Skillset: Hire from the Outside

While training current staff should be a first phase in reacting to this new technology, it's also critical to acknowledge that those outside of government (the private sector) likely have more experience in emerging technologies overall. The AFSA appears to acknowledge this when they discuss how Microsoft's early warning signs can help lead to improved public sector cybersecurity practices, and further encourage coalitions around the world to better take advantage of private sector data (Bachus). A new Lateral Entry Pilot Program, which was launched in 2024 to (in part) give private sector professionals with expertise in cyberspace and emerging technologies an easier path to joining The Foreign Service, is an excellent next step in preparing for the future (State Department). Hopefully this program will evolve from a small test to a larger-scale effort in the coming years. It's even possible that AI may be used to score these candidates on their ability to apply GenAI to their new prospective roles, which has been done by the U.S. State Department to test other qualities in the past (Blaser).

3) Sway: Cautiously Drive Influence

Armed with appropriate skepticism and the right skillset, The Foreign Service must be prepared to put GenAI to work across multiple disciplines. It is already a proven tool for big data mining, and it is increasingly being used in foreign affairs to pinpoint emerging risks and to enter negotiations with well-informed plans (ACPD). The European External Action Service is using it for enhancing communication across languages (Global Diplomatic Forum). And tech company Meta proved GenAI can be successful in game theory when it secretly entered a bot into an online diplomacy contest – where participants anticipate their opponents' moves to outmaneuver them – and it more than doubled the average score of its human competitors (Delcker). Capitalizing on its strengths, this powerful technology can help sway global officials and commerce partners. But users must also be conscious that opposing forces will be utilizing this technology to influence as well, creating new challenges on a different playing field in the future.

4) Standards: Establish Foundations of Trust

The Foreign Service will also have a continuing part to play in helping govern GenAI worldwide. This should include bringing local leaders together at the grassroots level to agree on shared values. Once alignment is achieved regionally, global efforts to adopt standards should be even more effective. The UN Norms for Responsible State Behavior in Cyberspace – guidelines that describe what countries should and should not be doing on the Internet – is a place to start (ASPI). The EU AI Act, which outlines rules for developing and using related tools, could be extended outside the union (Larsen). And guidelines from the U.S. Cybersecurity and

Infrastructure Security Agency already call out the importance of collaborating and communicating with international partners on key AI activities (U.S. CISA). Portions of these diverse frameworks must be unified, advanced, and defended internationally to protect our citizens' privacy, and our country from risks such as state-sponsored IP theft and violations of digital sovereignty (Humphries).

Conclusion

The geopolitical landscape is complex and ever changing, and effectively traversing a wide range of variables that are not entirely under any one nation's control can directly impact its future. A radical innovation stands to impact foreign relations, perhaps even more than any other single action or innovation, creating an entirely new paradigm to navigate in the pursuit of prosperity.

GenAI's ability to emulate human thinking, mislead, amplify, and inflame – all with relative ease, and at scale – presents extreme risk. Adversaries competing for superpower status, rogue states, and criminal non-state actors will all use it to their advantage. Even well-intentioned users can mistakenly introduce risk with its use.

The Foreign Service's ability to tame the black box by promoting skepticism with employees, boosting its overall skillset, cautiously using it to sway, and introducing standards to serve as guardrails around GenAI will have the greatest impact on safeguarding a nation's interests over time.

Works Cited

- ACPD Official Meeting Minutes. "The Use of Artificial Intelligence in Public Diplomacy." 14 June 2023, <https://www.state.gov/acpd-official-meeting-minutes-june-14-2023/>. Accessed 29 January 2024.
- Australian Strategic Policy Institute (ASPI). "The UN Norms of Responsible State Behavior in Cyberspace." March 2022, <https://documents.unoda.org/wp-content/uploads/2022/03/The-UN-norms-of-responsible-state-behaviour-in-cyberspace.pdf>. Accessed 28 January 2024.
- Bachus, Jennifer. "New Bureau, New Cyber Priorities in Foreign Affairs Interview." *The Foreign Service Journal*, November 2022, <https://afsa.org/new-bureau-new-cyber-priorities-foreign-affairs>. Accessed 27 January 2024.
- Blaser, Virginia. "How to Use Artificial Intelligence in Diplomacy." Washington International Diplomatic Academy, 1 October 2023, <https://diplomaticacademy.us/2023/10/01/artificial-intelligence-diplomacy/>. Accessed 12 January 2024.
- Carr, Earl. "Is China Threatening America's Dominance in The Digital Space?" *Forbes*, 20 June 2020, <https://www.forbes.com/sites/earlcarr/2020/06/20/is-china-threatening-americas-dominance-in-the-digital-space>. Accessed 31 January 2024.
- Castro, Chiara. "7 Cybersecurity Predictions to Look out For in 2024." *TechRadar*, 28 December 2023, <https://www.techradar.com/computing/cyber-security/7-cybersecurity-predictions-to-look-out-for-in-2024>. Accessed 10 February 2024.
- Delcker, Janosh. "Where AI Can Replace Diplomats—and Where It Shouldn't." *Interntionale Politik Quarterly*, 2 June 2023, <https://ip-quarterly.com/en/where-ai-can-replace-diplomats-and-where-it-shouldnt>. Accessed 17 January 2024.
- Fick, Nathaniel. "The Role of Partnerships in Advancing Cyber Diplomacy." RSA Conference, 14 April 2023, <https://www.youtube.com/watch?v=jtPQphosnA8>. Accessed 17 February 2024.
- Global Diplomatic Forum. "Digital Diplomacy Unveiled: AI's Role in Replacing Diplomats and the Crucial Boundaries We Must Set." 28 November 2023, <https://www.linkedin.com/pulse/digital-diplomacy-unveiled-ais-role-replacing-pgiye>. Accessed 17 January 2024.
- Goebels, Joseph. "The Radio as the Eighth Great Power." Speech on German Radio. 18 August 1933, <https://research.calvin.edu/german-propaganda-archive/goeb56.htm>. Accessed 6 January 2024.
- Humphries, Matthew. "China Accused of Massive IP Theft, AI Hacking by Five Eyes Intelligence Alliance." *PC Magazine*, 18 October 2023, <https://www.pcmag.com/news/china-accused-of-massive-ip-theft-ai-hacking-by-five-eyes-intelligence>. Accessed 26 January 2024. Accessed 28 January 2024.
- Larsen, Benjamin Cedric. "The Geopolitics of AI and the Rise of Digital Sovereignty."

- Brookings, 8 December 2022,
<https://www.brookings.edu/articles/the-geopolitics-of-ai-and-the-rise-of-digital-sovereignty/>. Accessed 17 January 2024.
- Lawrence, Lizzy. “To Get Better Heart Data from Wearables, Researchers Train AI to Separate Signal from the Noise.” StatNews, 19 July 2023,
<https://www.statnews.com/2023/07/19/heart-data-wearables-apple-watch-health/>. Accessed 10 January 2024.
- McKenzie, Bryan. “Is That Real? Deepfakes Could Pose Danger to Free Elections.” UVA Today, 24 August 2023,
<https://news.virginia.edu/content/real-deepfakes-could-pose-danger-free-elections>. Accessed 12 January 2024.
- Milmo, Dan. “Doctored Sunak Picture is Just Latest in String of Political Deepfakes.” The Guardian, 3 August 2023,
<https://www.theguardian.com/technology/2023/aug/03/doctored-sunak-picture-is-just-latest-in-string-of-political-deepfakes>. Accessed 12 January 2024.
- Nicoletti, Leonardo and Bass, Dina. “Humans are Biased. Generative AI is Even Worse.” Bloomberg Technology, 15 December 2023,
<https://www.bloomberg.com/graphics/2023-generative-ai-bias/>. Accessed 7 January 2024.
- Painter, Chris. “The Rise of the Internet and Cyber Technologies Constitutes One of the Central Foreign Policy Issues of the 21st Century.” The Foreign Service Journal, June 2018, <https://afsa.org/diplomacy-cyberspace>. Accessed 7 January 2024.
- “State Department Announces New Lateral Entry Pilot Program.” U.S. Department of State, 24 January 2024,
<https://www.state.gov/state-department-announces-new-lateral-entry-pilot-program/>. Press Release. Accessed 10 February 2024.
- State of California. “Benefits and Risks of Generative Artificial Intelligence.” November 2023,
https://www.govops.ca.gov/wp-content/uploads/sites/11/2023/11/GenAI-EO-1-Report_FI_NAL.pdf. Accessed 10 February 2024.
- U.S. Cybersecurity & Infrastructure Security Agency (CISA). Artificial Intelligence. 2024.
<https://www.cisa.gov/ai>. Accessed 17 January 2024.
- U.S. Department of Homeland Security. “The Increasing Threat of Deepfake Identities.” 14 September 2021,
https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf. Accessed 31 January 2024.

G.P. Maffei: Shining a New Light on China By Matthew Sennelius

I. Abstract

In the 1580s, while decades of disagreements strained commercial ties between Portugal and China, Giovanni Pietro Maffei (1533–1603) wrote the *Libri XVI Historiarum Indicarum* (“Sixteen Books of Indic Histories”) — in the sixth book of which he included a comprehensive ethnographic description of China. His unusually positive views of China, its culture, and its people are striking. During a time of high Sino–Portuguese tension and, more generally, the widespread notion of European superiority in the wake of the recent discovery of the Americas, Maffei’s high praise for the Chinese in all facets of society present a privileged glimpse into his motive in writing this ethnography. By analyzing the various ways in which Maffei subverts the ‘uncivilized pagan’ stereotype, I evaluate his views on this meeting of cultures to show how his writing was influenced by his desire to depict the Chinese in a positive light.

II. Introduction

In modern scholarship, one of the most under-recognized early European writers on Asian culture was Giovanni Pietro Maffei (1533–1603), an Italian Jesuit historian and writer hailing from Bergamo in Northern Italy, although he spent much of his career in Portugal whose most influential work on Asian history was the *Historiarum Indicarum Libri XVI* (literally meaning “Sixteen Books of Indic Histories”; henceforth the *Historiae Indicae*).

Maffei began work on the *Historiae Indicae* at the request of his patron, Cardinal King Henry I of Portugal.³⁹ Subsequently, he moved to Portugal in 1580 to conduct research on Portuguese activity in the East. Even though Henry died in the same year, Maffei would continue his work at the behest of Henry’s successor and nephew — King Philip I of Portugal (r. 1581–1598; alias King Philip II of Spain, r. 1556–1598). Maffei concluded his research in the Portuguese archives in 1584.⁴⁰ In 1588, Maffei’s *Historiae Indicae* was first published in Florence, but in light of its popularity, it continued to be reprinted all around Europe — in some cases over 160 years following its initial publication.⁴¹

During the years Maffei spent researching lands, cultures, and histories of the East Indies, he corresponded with Jesuit missionaries around Asia. In fact, Maffei himself exchanged letters in 1580 and 1581 with the famous Jesuit Matteo Ricci — letters which Ricci wrote from Cochin and Goa, respectively.⁴² But as he had not yet found success in China, Ricci instead referred Maffei to Alessandro Valignano (1539-1606), another Jesuit priest who was in Japan at the time and wrote a work on Jesuit accomplishments in Asia. Maffei’s correspondence with Valignano

³⁹ In the dedication to King Philip, Maffei writes: “Your Majesty urged me to continue, with equal speed and zeal, the history of Indic matters having been begun by the authority of your uncle (*nomine avunculi Tui*), King Henry.”

⁴⁰ Stefano Andretta, “MAFFEI, Giampietro - Treccani,” Treccani, accessed January 26, 2024, [https://www.treccani.it/enciclopedia/giampietro-maffei_\(Dizionario-Biografico\)/](https://www.treccani.it/enciclopedia/giampietro-maffei_(Dizionario-Biografico)/).

⁴¹ From surviving editions alone, we know that Maffei’s *Historiae Indicae* was reprinted in its original Latin in such places as Cologne (1589, 1593), Venice (1589), Vienna (1752), Antwerp (1605), Florence (1588), Bergamo (1590), and Caen (1614). Translations of all 16 books also appear in French and Italian.

⁴² Matteo Ricci, *Lettere* (Quodlibet, 2001), 21–25 (Letter 4), 33–41 (Letter 6).

ended up informing much of Maffei's ethnography of Japan and some of his ethnography of China.⁴³

Throughout his work, Maffei uses an unusually favorable characterization of Chinese culture to establish a likewise positive view of China in the reader's mind. To this end, he employs three strategies (which I present in the same order in which Maffei employs them in the work itself): classicizing China through explicit comparisons between Chinese and Graeco-Roman cultures; relating China and Europe by emphasizing their shared morality; and connecting pre-existing Chinese religious beliefs with those dominant in the West, thus presenting the Chinese populace as ready for Jesuit proselytization. Furthermore, I will consider how Maffei discusses the events that led to the Sino-Portuguese tensions and the hostility that the Chinese felt towards the Europeans at that time. He significantly downplays China's culpability while at the same time exaggerating the crimes and the reputational damage done there by European bad actors.

Considering how Maffei's time marked Europe's shift into colonialism, his frequent references to Graeco-Roman standards of both material and moral advancement are unusual. By presenting the Chinese as similar to classical societies, as morally mature, and with a strong potential for Christian conversion, Maffei aims for his sixteenth-century European reader to form a generally positive view of the alien and hostile far-East nation, and ultimately become more eager to support proselytization and foster friendlier Sino-European cooperation in the future.

All excerpts of Maffei's writing are my own translations based on the original Latin text unless otherwise noted. Each translation is followed by the page number on which the Latin appears in the 1752 reprint of the *Historiae Indicae*.⁴⁴

III. The History of Sino-European Conflict

Maffei's writings were written and published during the height of Chinese distrust towards Europeans, with the Portuguese receiving particular animosity. According to a recent description of Chinese-Portuguese relations by Michael Schuman,⁴⁵ the causes for this distrust include the following: Portugal's capture of Malacca, a Chinese tributary state; Simão de Andrade's actions in "building a fort on a Chinese island, forbidding other foreigners from trading ahead of him, and then abusing a Ming official who tried to assert control over the situation"; and the Portuguese refusal to cease trading after being expelled by Ming port officials. By 1522, Chinese and Portuguese forces had fought numerous naval battles and the Portuguese had been banned from official trade with the Ming government. These numerous points of disagreement followed by outright military conflict constituted the ultimate turning point in Sino-Portuguese relations.

⁴³ Donald F. Lach, *Asia in the Making of Europe, Volume I* (Chicago: University of Chicago Press, 2010), 803.

⁴⁴ See Giovanni Pietro Maffei, *Joannis Petri Maffei Bergomatis E Societate Jesu Historiarum Indicarum Libri XVI*, *Google Books* (Bernard, 1752), https://books.google.com/books?id=Ts5MAAAAcAAJ&printsec=frontcover&source=gbs_atb#v=onepage&q&f=false.

⁴⁵ See Michael Schuman, *Superpower Interrupted: The Chinese History of the World* (New York: Public Affairs, 2021), 215-216.

Despite the ban, for the following thirty years the Portuguese still found ways to access the abundant wealth of Ming trade through illegal smuggling operations with foreign merchants. Around the same time as Portugal began engaging in illegal activity, the new Ming emperor established a moratorium on maritime trade, effectively abolishing the tributary system, which caused an increase of both illegal trading and piracy. These factors ultimately forced the Chinese to reopen trade to the Portuguese in 1567.

During the maritime trade ban, the Chinese *haidao* (a shortened version of the title meaning “the Vice Official for the Inspection of the Maritime Sector”) allowed the Portuguese to establish a colony on the island of Macau, off the coast of Guangzhou. A subsequent agreement, signed in 1557, permitted them to establish permanent housing in Macau, which they did with energy and efficiency.⁴⁶

The Portuguese colony at Macau was viewed with suspicion by the Chinese. As Schuman writes: “To the Chinese, Macau became a truly foreign place, with strange architecture, stranger people, and unfamiliar religious processions. Some local Chinese saw the settlement in Macau as a bad omen; others complained that it was no longer part of China.” Notably, the Chinese referred to Europeans as “foreign devils.” This pervasive sentiment was so widespread that, in 1580, one European account said that it was “the natural tendency of the Chinese to fear and to bear ill will towards foreigners.”⁴⁷ It is important to note, however, that although after the establishment of Macao, Portugal established official relations with the Mandarins at Guangzhou, the benefits of these relationships extended only to trade and not to further intercultural contact.

In the years between the establishment of Macau and the publication of Maffei’s *Historiae Indicae*, many Jesuit priests came and went from the Portuguese settlement off the coast of Guangzhou. None of them, however, found success in entering the country due to a lack of trust alongside significant cultural and linguistic divides. One particularly unsuccessful attempt to enter the Chinese mainland was made by four “Spanish friars” who, claiming to be Portuguese, compensated for their ignorance of the Mandarin language by bringing an interpreter. Although they were initially welcomed and treated well, their lack of provisions and one of them falling ill forced them to “catch a junk to Manila.” Furthermore, according to the same account “The Mandarins had no intention of granting the Franciscans’ wish.”⁴⁸ Stories like this one painted a picture of a closed-off and stubborn China, a result of past trauma and lasting distrust.

Considering the great amount of communication between Europe and Asia, the ruling Ming dynasty’s hostility towards Europeans — with a particular ire reserved for the Portuguese — would have been a well-known obstacle to accessing the Chinese mainland whether for trade or missionary activity. As such, Maffei’s *Historiae Indicae* — with so much praise and admiration for the Chinese interspersed throughout — would have aimed to reframe the idea of

⁴⁶ R Po-Chia Hsia, *A Jesuit in the Forbidden City: Matteo Ricci, 1552-1610* (Oxford: Oxford University Press, 2013), 54.

⁴⁷ Schuman, *Superpower Interrupted*, 216.

⁴⁸ Hsia, *A Jesuit in the Forbidden City*, 61.

the Chinese people in the collective European consciousness, thus gaining more popular support for Christianization missions eastward.

IV. Maffei's Equation of China and Graeco-Roman Society

Maffei begins his ethnography with a general description of China's geography, people, and culture. Strikingly, he tends to compare Chinese ideas to Greco-Roman analogues in a way that goes beyond his general classicizing terminology in the work.⁴⁹ These comparisons not only made this foreign culture more accessible to an average European reader through familiar terms but also served to extend the contemporary idealization of classical societies onto the Chinese people — i.e., to bias educated Western Europeans in favor of the Chinese.

One notable instance of Maffei portraying the Chinese as moral and upstanding is through highlighting the diligence and hard-working nature of Chinese farmers:

In addition to the exceptional fertility of the land, there is the highest industriousness of the farmers. And there is an infinite multitude [of farmers]: although the population is growing daily and although movement into other foreign regions has been prohibited, it is not permitted for anybody to be idle in such a great crowd of mortals. Idleness is punished not just by private shame or the disgrace and quarrels of each neighbor but also by the public customs and laws. And so rural people allow no part of (their) field to be neglected out of carelessness. (*p. 119*)

Maffei's description of Chinese peasants' diligence and industriousness strongly parallels the Graeco-Roman "fetishization of virtuous poverty and simple living."⁵⁰ As one of the characters in the *Deipnosophistae* of Athenaeus (fl. c. 2nd-3rd C.E.) reflects in discussing the value of the Homeric writings:

For he [i.e., Homer] calculated that desires and pleasures are very powerful, and that those that involve food and drink are the most basic and deeply engrained, and that people who have lived in a consistently frugal manner are orderly and self-disciplined in other aspects of their lives as well. He therefore gives them all a simple way of life, drawing no distinction between kings and commoners, young and old.⁵¹

⁴⁹ Christopher Francese has noted that "Maffei uses classical terms like *decurion* and *praetor* freely throughout the work" in describing non-Western cultures. As such, I will be focusing on moments where his references to the classical world go beyond such standard classicizing uses of language — when he draws significant parallels in moral or material sophistication, and so forth; see Christopher Francese, "A Forgotten Masterpiece of Latin Prose: Maffei's *Historiae Indicae*," *The Classical Outlook* 94, no. 4 (2019): 192–97, <https://www.jstor.org/stable/10.2307/26865152>.

⁵⁰ Dean Hammer, *A Companion to Greek Democracy and the Roman Republic* (Chichester: Wiley-Blackwell, 2015), 85.

⁵¹ Athenaeus and Douglas Olson (ed. & tr.), *The Learned Banqueters* (Loeb Classical Library, 2006), 47.

Similarly, Roman authors often appealed to the “picture of an idealized past society rooted in the values of frugality and self-sufficiency.”⁵² One such example was the legendary Roman figure of Lucius Quinctius Cincinnatus. During a moment of crisis, Cincinnatus is said to have been made dictator of Rome, during which time he fended off another Italic tribe, the Aequi, celebrated a triumph, and promptly gave up his position of power to return to his farm.⁵³

Cincinnatus’ example and its legendary status speak to the high value that the Romans placed upon moderation and frugality — a notion that would have been well-known to Renaissance Europeans, who were attempting to restore Greco-Roman ideals.⁵⁴ By appealing to the admiration that Romans and Greeks had for frugality and simple living, Maffei’s description serves to idealize the Chinese peasantry and establish a strong sense of admiration among his reader base.

Hand-in-hand with the Roman glorification of the poor peasantry went their disdain for the luxury of the upper classes, whose lifestyles of excess they perceived as idle and morally degenerate.⁵⁵ The Roman historian Cassius Dio (c. 165–c. 235) writes, for instance, in his account of Seneca (c. 4 B.C.E.–65 C.E.), the famous Stoic philosopher and tutor of the emperor Nero:

Though finding fault with the rich, [Seneca] himself acquired a fortune of 300,000,000 sesterces; [...] In stating [his great shows of wealth] I have also made clear what naturally went with it — the licentiousness in which he indulged at the very time that he contracted a most brilliant marriage, and the delight that he took in boys past their prime, a practice which he also taught Nero to follow.⁵⁶

In contrast to the common Roman critique of luxury, Maffei’s comparison between the Chinese emperor and Emperor Vespasian (r. 69–79 AD) includes no such criticism:

Nor are there lacking those who assert for certain that in our currency, the king takes in 1,200 times 100,000 [i.e. 120 million] gold coins annually: an amount that the dying Emperor Vespasian, who was very diligent with money (*pecuniae diligentissimus*), hardly left in the Roman treasury. It is altogether a great sum: and the kind of sum in the consideration of which credence quite reasonably struggles (*laboret fides*).⁵⁷ (p. 129)

⁵² Peter Garnsey, *Food and Society in Classical Antiquity* (Cambridge, U.K. ; New York: Cambridge University Press, 1999), 72.

⁵³ Hammer, *Greek Democracy and the Roman Republic*, 74.

⁵⁴ I refer to “Greco-Roman ideals” here to describe the ideals of both Ancient Rome and Ancient Greece, not to suggest that all these ideals were shared between them.

⁵⁵ Garnsey, *Food and Society*, 72.

⁵⁶ Earnest Cary, Herbert Baldwin Foster, and Cassius Dio Cocceianus, *Dio’s Roman History. With an English Translation by Earnest Cary, Ph.D., on the Basis of the Version of H.B. Foster. Gr. & Eng.*, 1914, 57.

⁵⁷ “Credence suffers” means to say that the size of the sum is so difficult to believe or comprehend that it would be difficult to give credence to such reports. This phrase also suggests that the reported figure, while difficult to believe, may be accurate.

On one hand, Maffei makes the obvious material comparison between the wealth of Vespasian and the great riches that the Chinese emperor accumulates annually. This serves as another appeal to more economically inclined readers, portraying China as a land full of wealth from which enterprising Europeans would be able to profit. On the other hand, Maffei explicitly associates the emperor's great tax revenues with the financial prudence of Emperor Vespasian, who was extremely careful with money, rather than associating it with luxury or extravagance, as the Romans sometimes did. In fact, the Romans themselves disagreed about how to judge Vespasian's character. By some, Vespasian was perceived as stingy, but by others fiscally prudent, even frugal. This choice to describe the Chinese emperor in terms of his "carefulness" or "diligence" shows Maffei's desire to paint the Chinese in a positive light by associating them with a positive strand in ancient evaluations of Vespasian. The imperial biographer Suetonius (c. 69 C.E.-after 122 C.E.) notes criticisms of Vespasian but argues for a positive evaluation of his accumulation of wealth in the state coffers:

Others [...] believe that; [Vespasian] was driven by necessity to raise money by spoliation and robbery because of the desperate state of the treasury and the privy purse; [...] This latter view seems the more probable, since he made the best use of his gains, ill-gotten though they were.⁵⁸

Cassius Dio also presents a similar, positive outlook of Vespasian in his work *Roman History*:

In short, [Vespasian] was looked upon as emperor only by reason of his oversight of the public business, whereas in all other respects he was democratic and lived on a footing of equality with his subjects. For example, he indulged in jests like a man of the people and enjoyed jokes at his own expense; and whenever any anonymous bulletins, such as are regularly addressed to the emperors, were posted, if they contained scurrilous references to himself, he would simply post a reply in kind, without showing the least resentment.⁵⁹

The former account by Suetonius justifies Vespasian's use of "ill-gotten" gains through an understanding of the near-collapse of the empire as a result of the excessive spending of Nero, Vespasian's predecessor. In the latter account by Cassius Dio, Vespasian is held in high regard for his humility in interacting with the Roman people at their level — a trait rarely found in Roman emperors. Considering the value that Roman culture placed on leading the simple life, Vespasian's adherence to that life even in the position of ultimate power suggests a goodness and honesty in his character that Maffei, by associating the two, equally implies of the Chinese emperor.

⁵⁸ Gaius Suetonius and John Rolfe, *Suetonius. With an English Translation by J.C. Rolfe Ph. D.* (Harvard University Press, 1914), <https://archive.org/details/suetonius-loeb/page/n1/mode/2up>, 311.

⁵⁹ Cary, *Dio's Roman History*, 281.

Maffei later describes precisely one of the consequences of the emperor's diligence with money in noting: "The king [Emperor of China] impressively and dutifully maintains schools with excellent teachers in almost every town." (p. 125) Ultimately, by describing the nationally provided educational system, Maffei conveys a positive perception of the emperor: he exceeds his duty by providing and caring for his subjects, a notion only strengthened by the implicit contrast to the greed and moral degeneracy of the Roman upper class drawn by his allusion to Vespaian and thus ancient Roman leadership more generally.

Besides praising the Chinese people, he also frames their material sophistication as on par with or even superior to that of the Romans. He does so most explicitly when discussing the Chinese road system, writing: "[the roads in China] are of such cost and vastness that you would not miss (*requiras*) Roman magnificence in this kind [i.e., category]." (p. 121)

This description is especially striking because the Romans were widely renowned for their architecture, and especially their roads. In various accounts from classical and modern European sources alike, Roman roads are held in very high regard for their quality and durability. A monument of the pinnacle of Roman road building, the Via Appia, built in the fourth century BC, remains partially preserved even today. Eight hundred years after the road's construction, the sixth-century Byzantine historian Procopius of Caesarea praised the Via Appia in the following terms:

Now the Appian Way [...] extends from Rome to Capua. And the breadth of this road is such that two wagons going in opposite directions can pass one another, and it is one of the noteworthy sights of the world. [...] And they [i.e., the stones] were fastened together so securely and the joints were so firmly closed, that they give the appearance, when one looks at them, not of having been fitted together, but of having grown together. And after so long a time [...] they have neither separated at all at the joints, nor has any one of the stones been worn out or reduced in thickness, - nay, they have not even lost any of their polish. Such, then, is the Appian Way.⁶⁰

The Via Appia was culturally and strategically significant to the Romans, appearing multiple times in both Cicero's and Livy's works as a vital roadway central to both trade and military endeavors.⁶¹ Thus, in light of both the significance of Roman roads like the Via Appia and their renowned durability, Maffei's praise of Chinese roads becomes much more significant than it may initially seem. By claiming Chinese roads to be of the same quality as the legendary Roman ones, Maffei overturns the idea that the Chinese are less sophisticated than the Europeans. Ultimately, this comparison aims to surprise the reader, who would not expect a foreign people to be able to achieve or surpass the Roman standard, especially in a skill for which Romans were well-known.

⁶⁰ Cornelis van Tilburg, *Traffic and Congestion in the Roman Empire* (Routledge, 2007), 12–13.

⁶¹ Cic. Man. 18.55 <https://www.perseus.tufts.edu/hopper/text?doc=Cic.%20Man.%2018.55&lang=original>, Liv. 9.29 <https://www.perseus.tufts.edu/hopper/text?doc=Liv.%209.29.6&lang=original>

Another section of Maffei's writing that lends itself to a positive reading of the Chinese is the following description of their New Year celebrations:

They set the beginning [of the year] from the new moon of the month of March. They celebrate this day with public merriment; and each person celebrates their own birthday in private celebration. [...] Then, especially at night, they intersperse very extravagant dinner parties with the most exquisite entertainments. They exhibit lavish comedies and tragedies, either with the plots having been made up by the poet for the present purpose, or with the actions having been derived from ancient history. (p. 122)

Primarily of note in the above passage is how Maffei explicitly uses comedies and tragedies as examples of Chinese celebrations. This is important, first, because these two specific genres of stage play held great moral and cultural significance in Greco-Roman society. Traditional Roman celebrations of both comedies and tragedies such as the *ludi scaenici*⁶² and Greek ones like the Dionysia (Festival of Dionysus)⁶³ provide evidence of the social significance of theater in both societies. Thus, this connection deepens the parallels drawn between Greco-Roman society and Chinese society in the mind of the reader.

Also of significance is how Maffei describes some plays as "having been derived from ancient history": many writers and playwrights in the classical era drew from their own history. The Trojan War is the most common example: famous plays like Aeschylus' *Oresteia* and many of the works of Euripides and Sophocles deal with the events, or the aftermath, of this legendary war. In addition, this moment suggests that the Chinese had a level of historical reflectiveness that Europeans once thought exclusive to themselves. Since comedies and tragedies were seen as highly intellectual pursuits which offered insights into and contemplations on society as a whole, for non-European people to be able to do the same would be impressive.

Another significant aspect of the above passage is how Maffei emphasizes the value that the Chinese place on their Lunar New Year celebrations. By describing their festivities as very extravagant, most exquisite, and lavish (using the superlative adjectives *apparatissimis* and *exquisitissimi* and *sumptuosas*, respectively), he repeatedly emphasizes the economic cost of Chinese celebrations, and thus the Chinese dedication to the quality of the entertainment. To the Renaissance European, who looked to Roman and Greek art and society as the paragon of cultural sophistication, Maffei's presentation of Chinese art and customs as comparable to Graeco-Roman standards helped make this seemingly alien nation seem more familiar and more enlightened.

V. Morality in China

⁶² "Ludi Scaenici | Theatrical Performances, Roman Festivals & Religious Rituals | Britannica," [www.britannica.com](https://www.britannica.com/art/ludi-scaenici), accessed January 27, 2024, <https://www.britannica.com/art/ludi-scaenici>.

⁶³ "Great Dionysia," Brown.edu, 2009, https://www.brown.edu/Departments/Joukowsky_Institute/courses/13things/7411.html.

In Roman culture, a woman's value was largely based upon her chastity and modesty (*castitas* and *pudicitia*, respectively). As Rebecca Langlands writes in *Sexual Morality in Ancient Rome*:

[*Pudicitia*] is multidimensional, appearing in the Roman sources as deity, as core civic virtue, as psychological state, as physical state; it is associated with shame and awareness of social boundaries, with honour and bravery, with reputation, with patriotism, with self-control, with paternalistic authority over the sex lives of other people, with personal vulnerability, and with much more.⁶⁴

One example of *pudicitia*'s prescribed importance in Roman historiography appears in Livy's account of the legendary rape of Lucretia: after being raped by Sextus Tarquinius — the son of the final Roman king, L. Tarquinius Superbus — Lucretia kills herself in front of her husband and father, who go on to avenge her. The consequences of Sextus Tarquinius' actions are disastrous for both himself and his father, as Lucretia's father and husband drive them out of Rome, marking the end of Rome's time as a kingdom and the beginning of the Roman Republic.

This historical and cultural context makes it all the more significant when Maffei describes the great value that the Chinese place upon female purity and modesty:

The honor of modesty is great among the noble matrons. They are rarely seen, nor do they go out into public settings, unless they are in a litter, with curtains drawn on all sides, and for the purpose of concealed observation, with small, reticulated openings (made) out of ivory put on the sides. Servants carry [the litter] on their shoulders, with a large household accompanying on all sides. (*p. 122*)

Notably, Maffei uses the same term as the Romans in describing the modesty of Chinese women (*pudicitia*) and the high value that it is accorded in Chinese culture. His use of traditionally Roman terminology harkens back to a Roman sense of morality, thus opposing the commonly held European conception of pagans as amoral, or at least less moral, savages. By portraying Chinese morality to be, in this respect, essentially the same as the Europeans' (or perhaps even more committed to female *pudicitia*), Maffei's account lends the Chinese an air of cultural advancement and moral uprightness that would have been the ideal for a Western culture.

The moral uprightness of Chinese culture surrounding women — to such an extent that they are rarely seen in public settings (*raro visuntur, nec in publicum prodeunt*) — contrasts strongly with the customs of another region recently visited and described by Europeans: the Americas. More specifically, the dress (or lack thereof) of the indigenous peoples was the first thing that some Europeans noticed in exploring "The Indies." As Christopher Columbus reports in his 1493 letter to the Spanish administrator Luis de Santángel:

⁶⁴ Rebecca Langlands, *Sexual Morality in Ancient Rome*. (Cambridge University Press, 2009), 32.

The inhabitants of this island, and of all the others I have seen or of which I have received information, all go naked, the women as well as the men, just as they came into the world, with the exception of some women who cover themselves with a leaf or something made of cotton which they make for that purpose. They do not have any iron or steel or weapons, nor are they capable of using them; not from any deformity of body, but because they are incredible cowards ... for it has sometimes occurred that when I have sent two or three men to some village to speak with the natives they have been greeted by countless Indians and after they saw them arrive they fled so that even the fathers forsook their children [...] they are hopeless cowards.⁶⁵

Columbus' description of both these people's nakedness and their cowardice — the latter illustrated by “fathers [forsaking] their children” — suggests a strong connection between immodesty and moral shortcomings. To a Renaissance European reader, the comparative immorality of the Americans would have contrasted with the high moral standards that Maffei attributes to the Chinese people, primarily in the modesty of women. The difference in morality between these groups in spite of their common non-Christianity (even more, their lack of belief in the Abrahamic God) would establish a conception of the Chinese as, while nonetheless pagans, exceptionally moral pagans.

Beyond Maffei's considerations of Chinese morality in its social context, he also touches upon Chinese morality in their judicial and administrative practices. In particular, he considers it through what would today be called Blackstone's Ratio,⁶⁶ an idea that resonates with a precedent set in the Bible. When Abraham hears of God's plan to strike down the city of Sodom and the people in it, he pleads to God for the sake of the righteous in the city (who include his nephew, Lot, and Lot's family): “Far be it from [God] to do such a thing—to kill the righteous with the wicked, treating the righteous and the wicked alike [...] Will not the Judge of all the earth do right?” At the end of their exchange, Abraham is able to convince God to spare the city “[for] the sake of ten [righteous people].”⁶⁷

In light of this biblical moral precedent, Maffei's description of the Chinese commitment to fairness and diligence in judicial proceedings becomes even more striking:

From early morning until evening, with a few hours omitted for a midday break, a praetor sits in a high tribunal. [...] All things are administered openly. Solitude, secret conversations, and private whispers are forbidden, so that there is less of an opportunity for disgraceful actions and fraud, which hate the light [i.e., transparency]. Repeatedly,

⁶⁵ Other translations choose to call the natives, rather than “incredible cowards” and “hopeless cowards,” “timid to a surprising degree” and “incurably timid.” Regardless of the severity of the language, the connection drawn between their nakedness and immorality is the same; see Margarita Zamora, “Abreast of Columbus: Gender and Discovery,” *Cultural Critique* 17, no. 17 (1990), <https://doi.org/10.2307/1354142>, 137.

⁶⁶ Coined by William Blackstone in his work *Commentaries on the Laws of England*, Blackstone's Ratio is the idea that “[it] is better that ten guilty persons escape than that one innocent suffer.”

⁶⁷ Gn. 18:23–32 (NIV)

they call back those accused of a capital offense for questioning. Not unless the case has been investigated slowly and examined with complete clarity do they condemn [the accused]. Therefore, exceedingly few among such a great multitude of people end [their] lives by the hand of the executioner. (p. 126)

The final sentence in particular calls to mind Abraham's aforementioned deal with God and the legal principle that it is better for many guilty people to be acquitted than for an innocent to be condemned. Maffei highlights how few are executed to show that the Chinese court leans towards leniency for the sake of accuracy rather than overzealous punishment. In this way, he again brings into focus the many ways in which Chinese beliefs and morals resemble those found in European culture and Christian teachings, suggesting the ease with which potential missionaries could convert China, should they be offered the opportunity.

Maffei also shows the Chinese taking many steps to avoid nepotism and unjust appointments whenever possible through the use of the famous Ming dynasty civil service examinations. To demonstrate the fairness of these exams, Maffei describes the precautions that the Chinese go through to prevent all forms of cheating and unfair dealings:

In the most celebrated schools of each kingdom, for [the purpose of the civil service examinations], tables and benches are set up in large and spacious rooms. In these, adolescents summoned from other provinces (so that no favoritism is shown), with their tunics draped over so that they do not bring any book with them, are locked in the morning with writing tablets. There [...] with the doors closed and guards present, each writes as best he can and with the most intent focus, knowing that in such a competition not only dignity [i.e., status] but also each person's fortune is at stake. [...] Those who have carried the palm (*qui palmam tulere*) are called masters with exquisite honor of words: and that day is spent with great pomp and celebration of feasts and games. (p. 125)

Again using classical analogies, Maffei describes the high honor given to the 'winners' of the civil service examinations by referring to them as "those who have carried the palm," referencing the palm leaf as a symbol of Nike, the Greek goddess of victory.⁶⁸

Further, the meticulous system of protecting the integrity of the meritocratic civil service exams (i.e., bringing in adolescents in from other provinces, security against the smuggling in of notes) contrasts with the nepotism that was decried throughout Europe during the period, especially in the context of papal appointments. For example, of his predecessor Pope Calixtus III (r. 1455–1458), Pope Pius II (r. 1468–1464) himself wrote: "Nor did Calixtus escape censure for having preferred the ties of the flesh to the advantage of the church."⁶⁹ Here, Pope Pius II

⁶⁸ The Editors of Encyclopædia Britannica, "Nike | Greek Goddess," in *Encyclopædia Britannica*, August 26, 2014, <https://www.britannica.com/topic/Nike-Greek-goddess>.

⁶⁹ In this and the below instance of Pius II's personal writing, I rely on Richard Hilary, "The Nepotism of Pope Pius II, 1458-1464," *The Catholic Historical Review* 64, no. 1 (January 1978): 34, <https://www.jstor.org/stable/25020238>.

clearly condemns nepotism and the “ties of the flesh,” showing that it was recognized to be immoral and unjust to raise up unworthy people to positions of power by virtue of their connections. In spite of this, Pius himself indulged in nepotism; modern studies have concluded that about a quarter of Pius’ papal appointments could qualify as nepotistic in a broad sense.⁷⁰ Thus, while the head of the Catholic church, who was the authority on morality for the majority of Europe during the Renaissance, was known to widely engage in nepotism, the Chinese aim for meritocracy whenever possible. Furthermore, Maffei explicitly describes the Chinese ruler’s policy towards appointments of office and positions of power:

The King bestows magistracies and prefectures not so much as a benefit or reward, but as a task and duty: nor do the splendid titles of honor and the best revenues often follow a degenerate heir. Each person prepares his own wealth and dignity through effort and diligence: nor do the positions of brave and vigilant men get occupied by indolence and laziness, through the deceit of nobility. (*p. 129*)

In showing that the Chinese, as meritocratic and uncorrupt, actually surpass the Europeans morally, Maffei disproves the stereotype of pagan peoples being uncivilized and morally shallow. Maffei presents detailed evidence that the Chinese, in spite of their lack of Christian doctrine or established Christian communities, are in other cultural, social, and economic aspects either comparable to or better than the Europeans.

VI. Christianity in China

Sixteenth century European culture was dominated religion, especially the conflict between Catholicism and the emerging Protestant sects and the competition to convert far-off peoples who, thanks to the progress of maritime technology and Columbus’ discovery of the Americas less than a century prior, had recently come into close contact with the European world.

Maffei first implies that the Chinese are amenable to Christian conversion by reminding the reader that they are “believed to have formerly received the gospel from the Apostle Thomas” (*accepisse olim ab Apostolo Thoma creduntur*). His full description of China’s limited, apparently vestigial, Christian culture is as follows:

Of course they are believed to have received the gospel in the past from the Apostle Thomas, as was said above: and the image of a tall woman bears witness to this itself, holding a little boy in her lap, from which, even today, they suspend burning candles, as is our custom. But no longer does any mention of the Apostle exist in that place, and they do not know exactly whose likeness it is. (*p. 130*)

⁷⁰ As in the above footnote, this quotation is taken from: Hilary, “Nepotism of Pope Pius II”, 34

Maffei acknowledges scholarly doubts surrounding these reports of the Apostle Thomas' evangelization of China with the use of the phrase "they are believed" (*creduntur*), as even today traditional accounts of St. Thomas' activities in India and China remain disputed. In the same quotation, however, he takes the opportunity to highlight a similarity between Chinese and European religious practices by describing a statue of a tall woman holding a little boy in her arms (*puerulum in sinu foventis*), before which they suspend burning candles, as is our [i.e., the European] custom (*cui ardentis hodieque lychnuchos more nostro suspendunt*). He thus evokes the idea of the Virgin Mary and suggests that this is a (now poorly understood) trace of previous Christian practice in China.

Notably, while Maffei is very specific in his description of the statue, he does not explicitly make the connection between the tall woman with a child and the Virgin Mary with the Christ child Jesus. One possibility is that the statue actually depicted the common Buddhist image of the female Bodhisattva (alternatively called by her Chinese name: *Guanyin* 观音), who is sometimes shown holding a child in her arms.⁷¹ His ambiguity suggests that he might have known that the woman's identity was disputed (*ignorant*, "they do not know," could mean that Chinese informants claim to be ignorant or say that they know but identify the statue incorrectly). By remaining cryptic, he can imply greater Christian influence than there actually may have been while retaining scholarly integrity.

Regardless of his uncharacteristically imprecise description, by highlighting the similarity of their two cultures in their veneration of the Virgin Mary (because the reader would have likely assumed this identity of the woman), Maffei presents the Chinese as people who have already been introduced to Christian doctrine, but whose understanding of it is vestigial. These formerly Christian communities could be assumed to be ripe for proselytization from Europeans.⁷²

After this passage, however, Maffei addresses the reasons for the estrangement of these lapsed Christian outposts. One of his guiding motifs in his writing — and traditional Christian doctrine more generally — is the duty that Christians have to spread the word to non-Christian peoples. As Maffei would have known, the Biblical parable of the farmer found in the book of Mark presents a metaphor of the gospel as a seed and the individual as soil. Although the passage suggests that said seeds can fall upon paths, rocky places, or thorns, causing them to fail and die out, it also claims that "Others, like seed sown on good soil, hear the word, accept it, and produce a crop—some thirty, some sixty, some a hundred times what was sown."⁷³ Maffei takes inspiration from this well-known extended metaphor in his discussion of Christian outposts in China:

⁷¹ "Compassion, Mercy, and Love: Guanyin and the Virgin Mary," Metmuseum.org, 2021, <https://www.metmuseum.org/perspectives/articles/2021/5/virgin-mary-guanyin>.

⁷² More recent discoveries have uncovered evidence of Christian communities in China stemming from a missionary in 650 C.E. see: Athanasius Kircher, and Charles D. Van, *China Illustrata* (Muskogee, Okla: Indian University Press, Cop, 1987).

⁷³ Mk. 4:14-20 (NIV)

Indeed, either because of the very remoteness of the regions [i.e., from centers of Christianity], or because of the long-standing lack of cultivators [*cultores*, which can mean ‘farmers,’ ‘supporters,’ or ‘worshippers’], every seed of Christian piety died out in this land. (p. 130)

Maffei’s choice of words here is significant, as it sidesteps any idea that the Chinese people were inherently or culturally “infertile” for Christianity. Such a line of thought — that foreign, uncivilized pagans lacked the requisite sophistication to understand Christian doctrine, an idea that had become increasingly popular during the Renaissance — would have likely been an assumption held by at least some contemporary readers of the work. Maffei in this way dispels any notion that Christianity’s short-lived existence in China was a result of any deficiency of the Chinese. Rather, he offers two reasons (namely the distance from centers of Christian thought and the lack of priests, or even large communities of believers, i.e., “cultivators” of Christianity) which place the blame on factors outside of the Chinese people’s control. In these ways, by both showing vestiges of Christianity supposedly preserved in China and by suggesting that the disappearance of Christianity is not the fault of the Chinese, Maffei argues convincingly for China’s readiness for missionary work.

Maffei’s description of China curiously fails to provide any explicit reference to Confucianism as its main religion, one which had been established over two millennia prior to the publication of the *Historiae Indicae*. The closest Maffei comes to describing Confucianism is in a reference to laws and statutes (*jura legesque*) that receive serious study. Whether this confusion as to Confucianism’s status as a religion is a result of cross-cultural interpretation or purposeful obfuscation is unclear. In full, his vague description of what seems to be Confucianism is:

They [the Chinese] have laws and statutes written more than two thousand years ago, which have not (as they say) been changed since then. To these, indeed, many eagerly devote themselves because they generally open the way to honors and magistracies; and they inquire among themselves about political matters and the administration of kingdoms. (p. 125)

Of note in this particular description is Maffei’s uncharacteristic ambiguity. While it can be inferred with a high degree of certainty — given what is now known of the Confucian canon and its role in Chinese education and bureaucratic appointments (“honors and magistracies”) — that Maffei is referring to Confucian texts, his usual precision is lost here, much as it is in his description of the Virgin Mary/Bodhisattva statue mentioned above. This sudden change in style suggests that either Maffei sought to hide the true nature of China’s religious landscape or that accounts of China to which Maffei had access did not view Confucianism as a religion, due to it being largely based on “secular” rules and laws and seeming to be an organic part of Chinese society. Regardless of how this interpretation came to be, however, the effect is the same. The

reader, whether a European intellectual, Jesuit official, or King Philip I himself, is prevented from seeing Confucianism — presented as “laws and statutes” through which ambitious office-seekers study to gain knowledge of political science — as in competition with Christianity in any way, or as a threat to missionary work in the region.

VII. Maffei’s Portrayal of Chinese Hostility

Maffei closes his ethnography by addressing the Chinese people’s hostility towards Europeans. In the following passage, he describes the trustworthiness and success of a Portuguese traveler named Fernão Pires de Andrade (Lat. *Fernandus Andradius*) and the good impression that he left on the Chinese:

This deed [i.e., Fernão’s display of trustworthiness by announcing his departure well in advance, so that he could pay off all his debts] seemed to be the action of a great and innocent soul, and from the character of one man, they [the Chinese] doubtlessly judged the virtue of the whole nation. (*p. 133*)

This account might seem to present the Chinese as admirably trusting of foreign peoples: from the actions of one exceptionally honest individual alone, they assume the best of all foreigners. Indeed it might even portray them as naive, accepting the foreigners at face value with little to no suspicion. By suggesting that the Chinese are naive, Maffei would play to the assumption of some Renaissance European readers that pagan people are simple-minded, although he himself doesn’t endorse this view (and gives readers reasons to doubt this view elsewhere).

Regardless of which interpretation one subscribes to, both lead to the same conclusion: that the Chinese — as trusting and benevolent, even perhaps naive, as they are — must have good reason for their current hatred of foreign Christians. Maffei lays out such reasons directly afterwards, writing:

They [i.e., Portuguese sailors who came to China following Fernão’s departure] built a castle and set up guns and guards, assumed authority, defrauded public business dealings and claimed the business strictly for themselves, inflicted injuries and insults upon both foreigners and locals, raped young women, and purchased male and female slaves from kidnappers; it was a shame and disgrace not only to their own people and excellent king but also to Christian discipline and the most sacred teachings of (their) ancestors. Therefore within a few days, with an evil demon inciting, they became worthy of being considered not as allies or companions but as monstrous pirates and enemies. (*p. 133*)

Here Maffei indicates that Chinese hostility towards and wariness of all Europeans is warranted, considering the fairly recent depraved actions of bad actors there. Indeed, Maffei places the full force of blame on these unidentified European traders — framing the Chinese as

reasonably wary of them. He even condemns the actions of the Europeans as “a disgrace to their most excellent king, their own people [and] Christian discipline and the most sacred teachings of (their) ancestors” — directing readers to be appalled by these Portuguese sailors’ actions and swaying them to side with the Chinese on this issue, subverting what would have been the dominant negative European view of China at the time.

Moreover, Maffei is careful in his description of these Portuguese sailors, namely by neglecting to identify them, instead choosing to refer to them as “other ship captains” (“*aliorum [...] navarchorum*”). Other accounts of European activity in China indicate that one of the main perpetrators among these “other ship captains” was the Portuguese Simão de Andrade⁷⁴ — the brother of Fernão de Andrade, but a figure who seems to be mentioned nowhere in the *Historiae Indicae*. Maffei likely omitted mention of Simão to avoid outright condemning the Portuguese adventurer.⁷⁵ Additionally, Simão’s misdeeds and the ensuing fallout would have been sensitive subjects for the Portuguese king, to whom the book was dedicated. Ultimately, Simão’s omission from Maffei’s account serves to detach the actions of these “other ship captains” from any one nation, instead urging all Christians to be better behaved.

Only at the end of Book VI does Maffei address some common European criticisms of the Chinese, focusing on their arrogance and sense of superiority towards all other peoples:

Among these things [i.e., their overly subservient attitude towards their emperor], there are also other things that are far worse, even the mention of which horrifies the well-mannered mind; nonetheless, with an eyebrow having been raised [*elato supercilio*, i.e., arrogantly], the Chinese have contempt for all peoples other than themselves; they think that [all others] are uncivilized, lazy, and barbaric. (*p. 132*)

Significantly, Maffei uses only a few sentences to describe Chinese cultural shortcomings, primarily focusing on their king’s self-proclamations to be the ruler of the world (*dominum mundi*) and the son of heaven (*caeli filium*) as well as the Chinese subservience to their monarch and contempt for other nations. Notably, these criticisms are generic and widely shared beliefs among Europeans; their inclusion seems to be out of obligation rather than an essential part of his ethnography. Further suggesting that these criticisms are not intended to be his main focus: Maffei, immediately after listing these alleged cultural shortcomings, draws attention away from them by turning his focus back to European Christians:

However, clearly nothing is a greater impediment [to the Chinese acceptance of Christian doctrine] than the pattern of everyday life of our Christian compatriots (for whom it would have been proper to shine a light before [or outshine, *praelucere*] pagans towards all justice and chastity), which has not at all been consistent with such a title [i.e., of being Christian] and profession [i.e., of Christian faith]. (*p. 132*)

⁷⁴ For further contextual information, refer to: Michael Schuman, *Superpower Interrupted*, 215.

⁷⁵ In fact, Maffei’s flattery of these ship captains’ “most excellent king” (*optimo Regi*) further reinforces the notion that he was talking about Simão, as the work was dedicated to Philip II of Portugal.

In this way, Maffei condemns the actions of so-called Christian representatives in China for their licentiousness (*licentia*) and greed (*cupiditas*), urging the reader to be more worthy of the Christian title and mission. Furthermore, this denunciation and its focus on the religious offensiveness of European depravity in China suggests that proper Christians, such as the Jesuits or another order of priests, rather than soldiers or sailors, would be better representatives there. Maffei thus attempts to garner support for Jesuit missions eastward.

Further, Maffei leaves out important contextual information about Chinese-Portuguese relations. Central to understanding contemporary tensions between China and Portugal were the numerous naval battles that Portugal and China had fought after the failure of the Tomé Pires mission in the 1520s. In any other history of Portuguese activity in China, these battles would have been presented in full focus, as they catalyzed the collapse of Sino-Portuguese relations for the next half-century. As one particularly lurid account of Portuguese defeat that immediately followed this mission puts it:

Then, in 1522, another Portuguese squadron showed up off the Chinese coast with a commission to forge peaceful relations with the Ming. Unaware of the conflict that preceded them, they blithely sailed into a Chinese onslaught that sank two of their three ships. Those unfortunate Portuguese captured in these engagements endured a horrible end. “Twenty-three individuals were each hacked to pieces, losing their heads, legs and arms,” one surviving eyewitness recounted. “Their genitals were stuffed in their mouths, and the trunk of each body was wrapped around the belly in two chunks.”⁷⁶

Considering Maffei’s thoroughness (and careful omissions of certain details) throughout his work, it would be unreasonable to consider this exclusion to be anything different. Maffei is capable of describing horrific and vivid violence in his writing elsewhere in the *Historiae Indicae*. In a passage from another part of the work translated by Christopher Francese, Maffei describes “a horrific spectacle: disfigured faces, burned out eyes, skin hanging from naked limbs as if flayed to ribbons.”⁷⁷ The description of violence itself, then, is not the issue. It is also highly unlikely that Maffei was unaware of these battles that served as the climax of Chinese antagonism towards the Portuguese. More likely, he was aware of them, but chose to omit them from his account for two reasons.

One reason for Maffei’s silence might be that they exceeded the scope of the ethnography, veering more into the realm of military narrative. Another more significant reason is that he wanted to simplify the reader’s conception of the issue and its solution. Omitting the material leads the reader to infer that China’s distrust of foreigners is not a fundamental, possibly insoluble, disagreement or conflict of interests, but a broken relationship that can be solved by the actions of individuals. Indeed, the most striking part of this section of Book Six of the

⁷⁶ Schuman, *Superpower Interrupted*, 216.

⁷⁷ Christopher Francese, “A Forgotten Masterpiece of Latin Prose: Maffei’s *Historiae Indicae*,” *The Classical Outlook* 94, no. 4 (2019): 192–97, <https://www.jstor.org/stable/10.2307/26865152>.

Historiae Indicae is the language that Maffei uses to conclude it and, in fact, to conclude Book VI as a whole:

Above all in this way, the minds of the Chinese were alienated by the crimes of a few, and inestimable damage was done to the reputation and trustworthiness [of the Europeans]. And following thereafter, all those who came into these places from our world, either on their own or foreign ships, were mistreated and punished [*pessime accepti mulctatique sunt*]:⁷⁸ and for very many years, no shore has been more hostile to the Christian name [*nomini*, alternatively meaning reputation], and the Portuguese name especially. (p. 134)

Noteworthy in Maffei's final sentence of the book on China is the hyperbolic description of Chinese disdain for Christians, especially the Portuguese. Although it is indisputable that the Chinese did not have a positive relationship with the Portuguese, it is also hyperbolic to say that "no shore has been more hostile" to Christianity, considering that the Portuguese had direct access to official avenues of commerce with the Ming government. By exaggerating the effect of European impropriety in China, Maffei lends the situation even more urgency.

In this way, by attributing Chinese hostility to the crimes of a few rogue Portuguese, Maffei reinforces the notion that the primary obstacle to further European activity in China is one of individual character. He implies that the people sent as representatives to China should be those who not just abide by but even exemplify Christian faith and teachings. As a Jesuit, Maffei likely had fellow Jesuit priests — such as Matteo Ricci, with whom Maffei had correspondence, as noted above, and Alessandro Valignano — in mind during the writing of the *Historiae Indicae*.

VIII. Conclusion

A curious reader skimming the final pages of Maffei's ethnography of China might jump to the conclusion that Maffei was, much like many of the other Europeans during that time, thoroughly, even condescendingly, critical of Chinese people, Chinese religion, and Chinese culture. But to anyone reading his work in greater depth, Maffei's appreciation for Chinese culture in all its strengths reveals a more nuanced view of contemporary Chinese culture. In the face of the popular concept of European superiority, his account of China's competence — even excellence — in engineering, art, and social, judicial, and administrative morality, and readiness for Christian proselytization paints a unique and largely favorable picture of Chinese society through a European lens.

Thirteen years after the publication of the *Historiae Indicae*, Maffei's correspondent Matteo Ricci would first enter the Forbidden City in Beijing. Much of Ricci's success has since been attributed to his pioneering use of inculturation — gaining the trust and respect of the Chinese

⁷⁸ Maffei doesn't describe in particular any acts of violence perpetrated against European ships or people, so this phrase could be his way of addressing but glossing over those incidents, although the Latin suggests, perhaps, paying a fine or being imprisoned.

with whom he dealt by learning their culture, language, and customs. While it would be speculative to say that the *Historiae Indicae* contributed to Ricci's success, what Ricci's accomplishments show is that Maffei was right in at least one major respect: by suggesting that a major obstacle to good relations with the Chinese was not anything fundamental to either culture, but rather that bad representatives had ruined the trust that Fernão de Andrade had built many years prior. Written about a decade before Ricci inaugurated a new age of European relations with and study of China, Maffei's still understudied *Historiae Indicae* shows a remarkably generous understanding of Chinese culture, singular in remaining largely unclouded by the common European prejudice against foreign, non-Christian peoples.

Works Cited

- Andretta, Stefano. "MAFFEI, Giampietro - Treccani." Treccani. Accessed January 26, 2024. [https://www.treccani.it/enciclopedia/giampietro-maffei_\(Dizionario-Biografico\)/](https://www.treccani.it/enciclopedia/giampietro-maffei_(Dizionario-Biografico)/).
- Athenaeus, and Douglas Olson. *The Learned Banqueters*. Loeb Classical Library, 2006.
- Cary, Earnest, Herbert Baldwin Foster, and Cassius Dio Cocceianus. *Dio's Roman history. With an English Translation by Earnest Cary, Ph. D., on the Basis of the Version of H.B. Foster. Gr. & Eng.*, 1914.
- "Compassion, Mercy, and Love: Guanyin and the Virgin Mary." Metmuseum.org, 2021. <https://www.metmuseum.org/perspectives/articles/2021/5/virgin-mary-guanyin>.
- Cornelis van Tilburg. *Traffic and Congestion in the Roman Empire*. Routledge, 2007.
- Francese, Christopher. "A Forgotten Masterpiece of Latin Prose: Maffei's *Historiae Indicae*." *The Classical Outlook* 94, no. 4 (2019): 192–97. <https://www.jstor.org/stable/10.2307/26865152>.
- Garnsey, Peter. *Food and Society in Classical Antiquity*. Cambridge, U.K. ; New York: Cambridge University Press, 1999.
- "Great Dionysia." Brown.edu, 2009. https://www.brown.edu/Departments/Joukowsky_Institute/courses/13things/7411.html.
- Hammer, Dean. *A Companion to Greek Democracy and the Roman Republic*. Chichester: Wiley-Blackwell, 2015.
- Hilary, Richard. "The Nepotism of Pope Pius II, 1458-1464." *The Catholic Historical Review* 64, no. 1 (January 1978): 3. <https://www.jstor.org/stable/25020238>.
- Kircher, Athanasius, and Charles D Van. *China Illustrata*. Muskogee, Okla: Indian University Press, Cop, 1987.
- Lach, Donald F. *Asia in the Making of Europe, Volume I*. University of Chicago Press, 2010.
- Langlands, Rebecca. *Sexual Morality in Ancient Rome*. Cambridge University Press, 2009.
- "Ludi Scaenici | Theatrical Performances, Roman Festivals & Religious Rituals | Britannica." www.britannica.com. Accessed January 27, 2024. <https://www.britannica.com/art/ludi-scaenici>.
- Maffei, Giovanni Pietro. *Joannis Petri Maffei Bergomatis E Societate Jesu Historiarum Indicarum Libri XVI*. Google Books. Bernard, 1752. https://books.google.com/books?id=Ts5MAAAAcAAJ&printsec=frontcover&source=gbs_atb#v=onepage&q&f=false.
- Matteo Ricci, and Brendan Gottschall. *Matteo Ricci: Letters from China : A Revised English Translation with Commentary*. Chicago, Illinois: Beijing Center Press, 2019.
- R. Po-Chia Hsia. *A Jesuit in the Forbidden City : Matteo Ricci, 1552-1610*. Oxford: Oxford University Press, 2013.
- Ricci, Matteo. *Lettere*. Quodlibet, 2001.
- Schuman, Michael. *Superpower Interrupted : The Chinese History of the World*. New York: Publicaffairs, 2021.
- Suetonius, Gaius, and John Rolfe. *Suetonius. With an English Translation by J.C. Rolfe Ph. D.*, 1914. <https://archive.org/details/suetonius-loeb/page/n1/mode/2up>.
- The Editors of *Encyclopedia Britannica*. "Nike | Greek Goddess." In *Encyclopædia Britannica*, August 26, 2014. <https://www.britannica.com/topic/Nike-Greek-goddess>.
- Zamora, Margarita. "Abreast of Columbus: Gender and Discovery." *Cultural Critique* 17, no. 17 (1990). <https://doi.org/10.2307/1354142>.

The Effects Of Music-Induced Nostalgia On Healthy Aging By Eliana Sun

Abstract

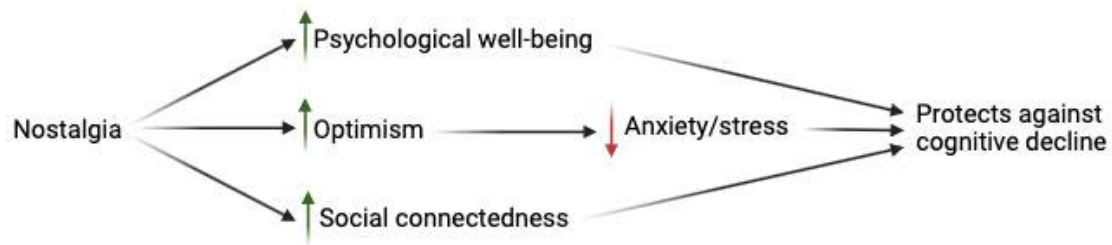
This review paper examines the therapeutic potential of music-induced nostalgia for individuals with cognitive impairments, such as those with Alzheimer's disease and dementia. It investigates how engaging with nostalgic music can yield significant psychological benefits, including enhanced social connectedness, improved psychological well-being, and increased optimism. These benefits are particularly important as they may contribute to delayed or preventative deterioration of cognitive function in affected individuals. By exploring the mechanisms behind music-induced nostalgia and its effects on cognitive health, this paper highlights the need for further research into personalized music interventions. Such interventions are significant in supporting the emotional and cognitive well-being of dementia patients, ultimately enhancing their quality of life and offering a promising avenue for mitigating the progression of cognitive decline.

Introduction

Nostalgic experiences are characterized by recollections of memory often tied to memorable experiences throughout the lifetime of an individual. Historically, nostalgia has been viewed differently. Having been conceptualized as both a neurological disease and psychiatric disorder in the past, it is now considered a bittersweet feeling. Just as the perception of nostalgia has evolved, so has the study of nostalgia in health and disease. Contemporary biomedical and psychological research defines nostalgia as an emotion of longing for memorable moments in an individual's life and implicates the phenomenon across disease conditions such as depression, anxiety, and dementia.

Mounting evidence implicates nostalgia as a key variable in the treatment of dementia, a condition with high unmet medical need (Minyo, and Judge). Current treatment paradigms for dementia focus on pharmacotherapy, palliative care, etc. While a diverse suit of treatment options exists, none appear to halt or reverse the progressive symptoms of dementia. Indeed, at present there are no curative therapies for the treatment of dementia, highlighting the previously mentioned unmet medical need in this space.

In this review article, I posit the notion that nostalgia, induced by music, is a potential therapeutic intervention in the treatment approach for dementia (Graphical Abstract). Through a literature review, I explore herein the potential of nostalgia, induced by music, to elicit several behaviors that contribute to protection against cognitive decline in individuals with dementia.



Created in BioRender.com 

Graphical abstract: Overview of article. Created with BioRender.com.

Nostalgia In Dementia Patients

To determine if nostalgia is a viable treatment option for dementia patients, its preservation in these individuals must first be confirmed. In a study testing the conservation of different properties of nostalgia in individuals with dementia, the memory narratives of 67 participants with varying degrees of dementia were used to determine if nostalgia's social, self-oriented, and existential structure was maintained (2; summarized in Figure 1). The nostalgic group was asked to describe a nostalgic event while the control group was told to describe an ordinary one. Then, both groups were assessed by responding to statements regarding how nostalgic they felt (Figure 1). The narratives were told through face-to-face interviews, with researchers unaware of which group the participant was assigned to (blinded). It was found that all nostalgic narratives had an emotional component, with the memories themselves pertaining to more significant events in life, such as a wedding or graduation. In comparison to ordinary memory narratives, nostalgic narratives were longer, had more expressions of self-esteem and self-continuity, and overall increased positive affect. There was also a trend toward increased expressions of optimism within the nostalgic narratives (Ismail, et al. "The Content of Nostalgic Memories among People Living with Dementia"). From these results, undergoing an experience of nostalgia, especially while recalling specific memories, has a positive influence on overall psychological well-being, through an increase in self-esteem and self-continuity which are both components of an optimistic view of life. Nostalgia is generally understood to hold positive emotional valence and is related to optimistic views on life. Indeed, one of the principal findings of this study is the presence of nostalgia in dementia populations. This importantly highlights a clinical feature in these patients that can be both studied and potentially therapeutically targeted.

These results were complemented by additional reports (El Haj, et al.; Mills, and Coleman). In the latter, a study examining phenomenological reliving in Alzheimer's disease

(AD) patients, the authors (El Haj, et al.) utilized 27 AD participants and 30 cognitively normal participants. In their autobiographical assessment, participants detailedly described an event in their lives, then rated their metacognitive judgments (relying, remembering, sense of realness), component processes (visual and auditory imagery, language, emotion), and narrative properties (importance, rehearsal). The results showed that AD patients had poorer autobiographical recall, phenomenological relying, and visual imagery compared to the control, however, they showed a higher presence in emotion and importance. This indicates the preservation of nostalgia in AD patients, as nostalgic memories are more likely to hold emotional significance in an individual's life (Ismail, et al. "The Content of Nostalgic Memories among People Living with Dementia"). As such, nostalgia and the ability to re-experience a nostalgic memory is still present in the dementia population. Meanwhile, in the first study referenced above (Mills, and Coleman), the researchers interviewed five participants with moderate to severe dementia about their past experiences, and came to the same conclusion that individuals with dementia are still able to recall their life experiences, and their ability to feel emotion and sense of self tangent to this recall is still intact. Furthermore, it is suggested that the emotional content of experiences could have allowed the memories to be preserved, deepening the connection of nostalgia in dementia patients. Together, these reports confirm the presence of nostalgic memories in dementia populations, which in turn enables additional study of the properties of nostalgia on dementia risk, progression, and potential treatment. In particular, the subsequent section of this review paper examines the connection between nostalgia and positive affect and its associated physiological outcomes with potential clinical and therapeutic relevance to dementia.

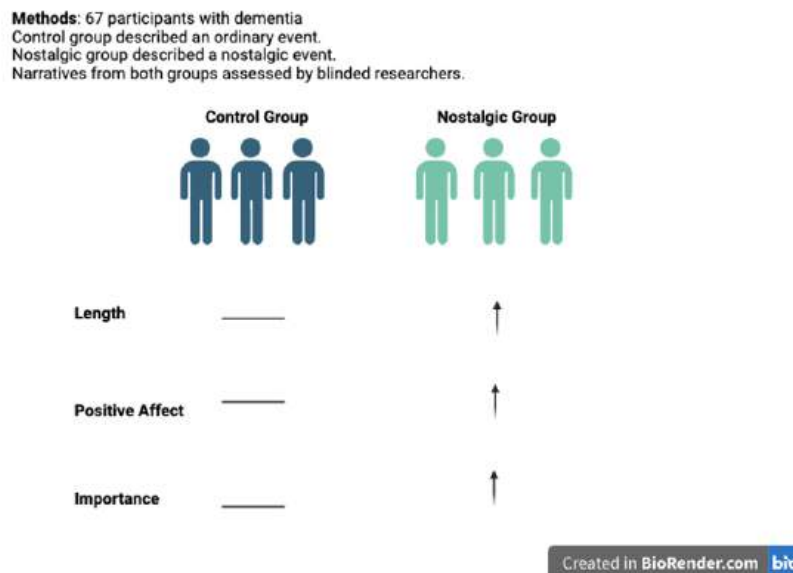


Figure 1: In a study consisting of 67 participants with dementia, the nostalgic group described memories that were longer, had higher positive affect, and had more emotional importance compared to the control group. Created with [BioRender.com](https://www.biorender.com).

Behavioral Outcomes Of Nostalgic Reminiscence

The Impact Of Nostalgia On Social Connectedness In Dementia

Feelings of joy and positive affect are closely associated with nostalgia (Leunissen, et al.). These positive feelings can lead to beneficial behavioral outcomes, such as increased social connectedness. Social connectedness refers to the state of being/feeling supported by an individual's relationships with others. For example, possession of social networks and little perceived social isolation. A lack of social connectedness, including perceived social isolation and loneliness, is a particular concern for older adults. Studies have shown that social isolation and loneliness in adults over 50 can result in adverse psychological outcomes, such as depression and anxiety, as well as physical health consequences like increased obesity, smoking, and premature mortality. This is particularly pronounced during the COVID-19 pandemic (Donovan, and Blazer). Other health outcomes included coronary artery disease, stroke, and suicide. Thus, methods and strategies for improving social connectedness within older adults are necessary to decrease physical health risks and decrease the possibility of neurocognitive disorder.

In the dementia population, social connectedness is a similarly pressing issue because poor social relationships are associated with cognitive decline (Piolatto, et al.), potentially making an individuals' cognitive health situation worse. A study with 9264 participants aged 50 and over revealed the dementia group experienced less social engagement compared to the non-dementia group (Hackett, Steptoe et al. 2019; summarized in Figure 2). Social engagement of the participants was measured through self-reports of meeting friends/relatives and speaking to friends/relatives by phone in a given period. The results established adults who would be diagnosed with dementia already engaged in less social activity two years before the diagnosis, and social engagement continued to decline after the diagnosis. Evidently, as seen through the two different cases of less social engagement prior to dementia diagnosis and less social engagement after dementia diagnosis, increasing social connectedness is protective in both a preventive measure and a mitigating measure. As such, treatment methods that increase social connectedness in dementia patients have the potential to slow cognitive deterioration and improve overall well-being and quality of life of individuals with cognitive impairment.

Nostalgia is one method that promotes increased social connectedness. In the four studies outlined in a paper regarding nostalgia's ability to counteract loneliness (Zhou, et al.), the restorative role of nostalgia is examined and considered as a resource to increase perceptions of social support and protect mental health. In Study 1, 758 migrant children who had moved from rural areas to a city in China were assessed for degree of loneliness with the 10-item UCLA Loneliness Scale, nostalgia proneness with the Southampton Nostalgia Scale, and social support with the 12-item Multidimensional Scale of Perceived Social Support. The results indicated loneliness led to a decrease in perceived social support, but also more strongly encouraged nostalgic engagement which led to increased perceptions of social support. Study 2 utilized 84 participants and assigned them to either the high-loneliness conditions or the low-loneliness condition, where a manipulation was used to make the participants in the high-loneliness conditions agree to more prompts on feelings of loneliness, thus placing them above average on the loneliness scale and decreasing perceived social support. The opposite was true for

participants in the low-loneliness condition. The manipulation led to high-loneliness participants feeling more lonely and more nostalgic, with a positive association between nostalgia and perceptions of social support, which replicates the results from Study 1. 66 participants were utilized in Study 3, where nostalgia was induced by prompting participants in the nostalgia condition to think of a past nostalgic event and reflect, while participants in the control condition were told to think and reflect on an ordinary event. Then, participants rated how nostalgic they were feeling and completed measures of perceived social support. The nostalgic group reported more nostalgia compared to the control condition and higher perceived social support compared to the control condition. The results of this study confirm nostalgia's ability to increase perceived social connectedness. Lastly, in Study 4, 193 factory workers were measured for loneliness, resilience using the 15-item form of the Resilience Scale, and nostalgia. The findings of Studies 1-3 were again replicated; while loneliness reduces perceived social support, it also leads to more nostalgia and therefore, increased perceptions of social support (Figure 2). Study 4 also concluded resilience has no effect on an individual's perceived social support derived from nostalgia, but resilient people are more likely to engage in nostalgia when lonely. The findings from this research emphasize the ability of nostalgic engagement to relieve feelings of loneliness by increasing feelings of social support, information that is highly relevant to the potential strategies that may be utilized to increase social connectedness in patients with dementia, who especially suffer from social isolation. Subsequently, because older adults and dementia patients are prone to isolation and decreased social connectedness, a method which increases the nostalgia they may already feel due to loneliness, perhaps supplementing existing nostalgic engagement, will potentially allow for a barrier against hastened cognitive decline and worsening of the dementia condition through increased social connectedness.

As previously detailed, the preservation of nostalgic structures in the dementia population allows for nostalgic experiences to occur, which in turn ensures that the positive outcomes associated with nostalgia, such as social connectedness, can also be experienced. In a study of 952 participants with a mean age of 69 years, researchers examined the relation between social isolation, loneliness, and cognitive decline (Gardener, et al.). Participants self-reported the frequency of social visits, phone calls, their satisfaction with social visits, number of friends, and degree of loneliness. Participants were then subject to neuropsychological assessments and were diagnosed with mild cognitive impairment and dementia depending on the evaluations. The results indicated those who were socially isolated (self-reported 3 or less friends) or had feelings of loneliness had higher odds of developing mild cognitive impairment and dementia, while those who were socially connected or did not report loneliness did not have higher odds (Figure 2). These results demonstrate how social connectedness, characterized by communicating often and having a sizable social network, can protect against cognitive decline. The association of nostalgia, social connectedness, and cognitive decline suggests that nostalgia can be used as a tool for increasing social connectedness in individuals with dementia. By fostering nostalgic experiences, individuals with dementia can reconnect with cherished memories and relationships, mitigating feelings of loneliness and isolation. This may, in turn, encourage more frequent social

interactions and strengthen social bonds, thereby mitigating feelings of loneliness and isolation and improving cognitive health. Thus, incorporating nostalgic activities and environments into care strategies for the dementia population could play a significant role in enhancing their social connectedness and slowing cognitive decline.

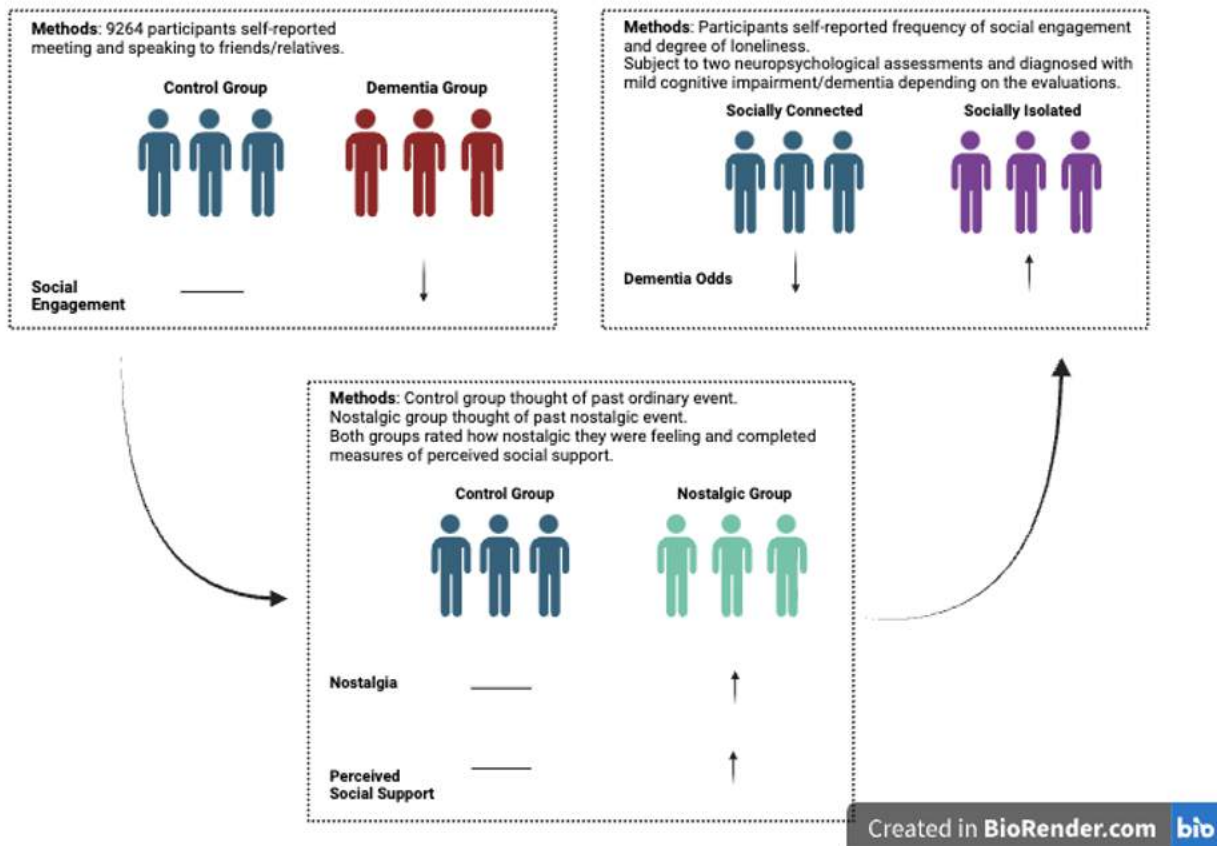


Figure 2: Three studies underline the intersection of nostalgia, social connectedness, and dementia. Dementia patients engage less socially, which worsens mental health. Nostalgic engagement, however, increases social support and therefore decreases odds of dementia.

Created with [BioRender.com](https://www.biorender.com).

The Impact Of Nostalgia On Psychological Well-Being In Dementia

In addition to social connectedness, self-esteem is another factor hindered by dementia that, if increased, has the potential to protect against cognitive decline. Self-esteem can be defined as an individual's perception of their own worth. Aside from outward behaviors associated with dementia such as motor dysfunction, delusion, and appetite/sleep changes, further emotional symptoms include unhappiness, hopelessness, and loss of self-esteem (Cerejeira, et al.). Psychological resources like self-esteem, meaning in life and self-continuity are crucial for maintaining cognitive health (Isene, et al.; Toki, et al.), suggesting their potential therapeutic value in dementia. Consequently, approaches that increase different aspects of

psychological well-being, such as self-esteem, meaning in life, and self-continuity merit consideration in treatment possibilities.

In a meta-analysis by Umar Ismail, et al. , nostalgia is identified as a method to increase psychological well-being. Across 47 experiments with 5043 participants, those in the nostalgic condition were instructed to either reflect on a nostalgic memory or listen to nostalgic music, while the control condition focused on ordinary memories. The results of psychological resources in the nostalgic condition versus the control revealed greater increase of social connectedness (in all 16 studies addressing social connectedness), self-esteem (8 out of 11 concluded in favor of the nostalgic condition, 3 favored neither condition), meaning in life (all 7 studies), and self-continuity (all 6 studies) in the nostalgic condition (Figure 3). From these results, it is evident that nostalgic reminiscence holds implications for improving dementia care since social connectedness, self-esteem, and self-continuity are integral components of an individual's identity and factors impacting mental health.

Similarly, a review by Fleury, et al. further supported nostalgia as a strategy to increase psychological well-being. According to the authors, aging is associated with changes in the brain and heart which hinder regulatory capabilities. However, nostalgia contributes to these emotional, physiological, and behavioral regulatory capabilities, thus making it arguably relevant as an intervention to support healthy aging. Specifically, regarding nostalgia's emotional regulatory capabilities, nostalgia enhances feelings of comfort, safety, and solace in face of hardship. Within dementia patients, it improves psychological well-being: self-esteem, personal growth, meaning in life, and social connectedness. Therefore, nostalgia is associated with augmented psychological well-being, and holds particular significance in the older population, wherein nostalgic reminiscence fosters a sense of safety and leads to healthy aging through increased capability of emotional regulation. Moreover, dementia patients also benefit significantly from nostalgia, as specific areas of psychological well-being that are weakened with the condition can be improved. Enhancing these behaviors by encouraging nostalgic engagement can help protect against cognitive decline.

To concentrate on the impact of nostalgic reminiscence on the psychological well-being of specifically dementia patients, a study reporting three experiments examined the ability of nostalgic reminiscence to benefit not only the general population but also individuals with dementia through the strengthening and increase of psychological resources (Ismail, et al. "Psychological and Mnemonic Benefits of Nostalgia for People with Dementia"). In Experiment 1, 29 participants with mild to moderate cognitive impairment (diagnosed with either probable Alzheimer's disease or a form of dementia) were randomly assigned to either the nostalgia condition or the ordinary (control) condition. Nostalgia was induced by recall of either a nostalgic memory or an ordinary memory, then degree of nostalgia felt was assessed, along with an assessment of psychological resources using the State Functions of Nostalgia Scale (SFNS) which evaluates self-esteem, self-continuity, social connectedness, meaning in life, optimism and positive/negative effect. The results showcased participants in the nostalgia condition achieving higher levels of psychological resources, reiterating nostalgia as an intervention which is

applicable to individuals without cognitive impairment and those with, even more strongly affecting the dementia population perhaps because of the low levels of psychological resources they already experience. Experiment 2 utilized music instead of memory recall to induce nostalgia. 32 participants were again assigned to either the nostalgia condition or the control condition, where a participant in the nostalgia group provided three of their favorite nostalgic songs, then listened to a single selected one. The same song was then provided to a participant in the control group, but it was only considered nostalgic to the participant in the other condition. This experiment resulted in information complementing the previous experiment, where participants in the nostalgia condition reported higher social connectedness, self-esteem, meaning in life, and self-continuity. Finally, in Experiment 3, 50 participants were, again, randomized into either the nostalgia or control condition, and nostalgic induction occurred identically as in Experiment 1. Participants were assessed for anxiety, depression, and cognitive functioning prior to nostalgia induction, then assessed for mood changes using a measure of positive and negative affect. Results of higher levels of social connectedness, meaning in life, self-continuity, and self-esteem from the previous two experiments were again replicated, and additional information about the recall of dementia-related information was revealed. Participants in the nostalgia condition were found to have recalled more dementia-related pieces of information that was read to them during the experiment and demonstrated higher recognition of the statements (Figure 3). The information from Experiment 3 which illustrates the ability of nostalgia to improve even short-term memory recall highlights the promising beneficial outcomes of nostalgic reminiscence, especially regarding protection against further cognitive decline. Furthermore, the ability of nostalgia to boost psychological resources useful for protecting against cognitive decline (self-continuity, meaning in life, self-esteem) was corroborated by all experiments of the study and additional reviews, posing a relevant and optimistic avenue for maintaining cognitive health of individuals with dementia. Collectively, information on nostalgia as a promoter of psychological well-being and memory recall serves as a method which could influence clinical management strategies and basic scientific knowledge of the role of nostalgia in memory consolidation and recall.

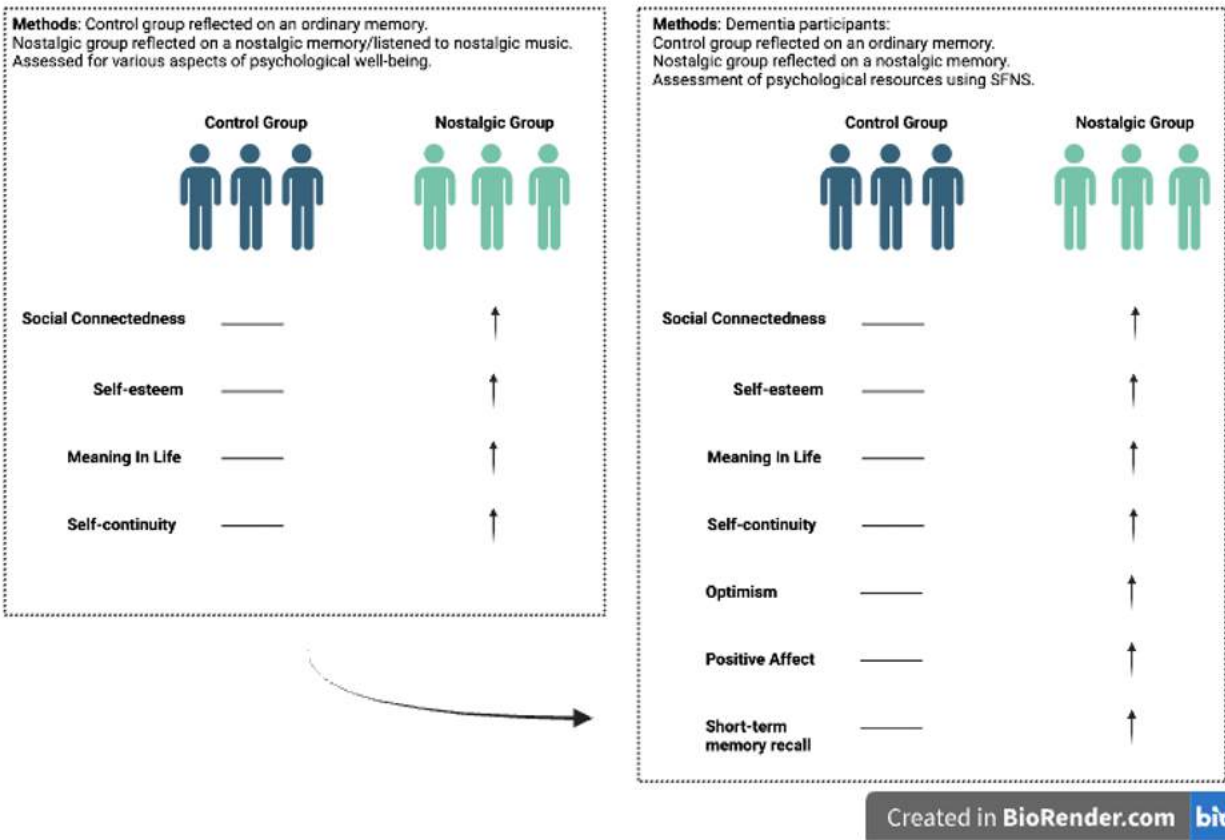


Figure 3: In both cognitively normal and dementia participants, the nostalgic group displayed increased psychological resources. Created with [BioRender.com](https://www.biorender.com).

The Impact Of Nostalgia On Optimism In Dementia

Another behavior often elicited by the experience of nostalgia is optimism. Optimism refers to the tendency to hold a positive outlook on the future, or a mindset emphasizing the favorable side of a situation and is often associated with the positive outcomes of nostalgic reminiscence. Nostalgia is not only associated with present positive affect but also with future optimism. (Wildschut, et al.). A series of studies by Cheung, et al. expands the scope of nostalgia’s impact to the future, examining its ability to make the future seem more positive. Nostalgia’s association with optimism was evaluated in four studies. In Study 1, 102 participants were randomly assigned to nostalgia and control conditions. Participants in the nostalgia condition were instructed to recall a nostalgic past event, while participants in the control condition were instructed to recall an ordinary past event, a procedure of nostalgia induction utilized frequently to discern differences between nostalgic thinking and “ordinary” thinking. After describing their past events, participants completed a nostalgia manipulation check and the Positive and Negative Affect Schedule questionnaire, of which the results reported participants in the nostalgia condition felt more nostalgic and reported more positive affect. Study 1 also established nostalgic narratives had more optimism expressions compared to ordinary narratives, suggesting nostalgia elicits optimism. Study 2 induced nostalgia with the same method as Study

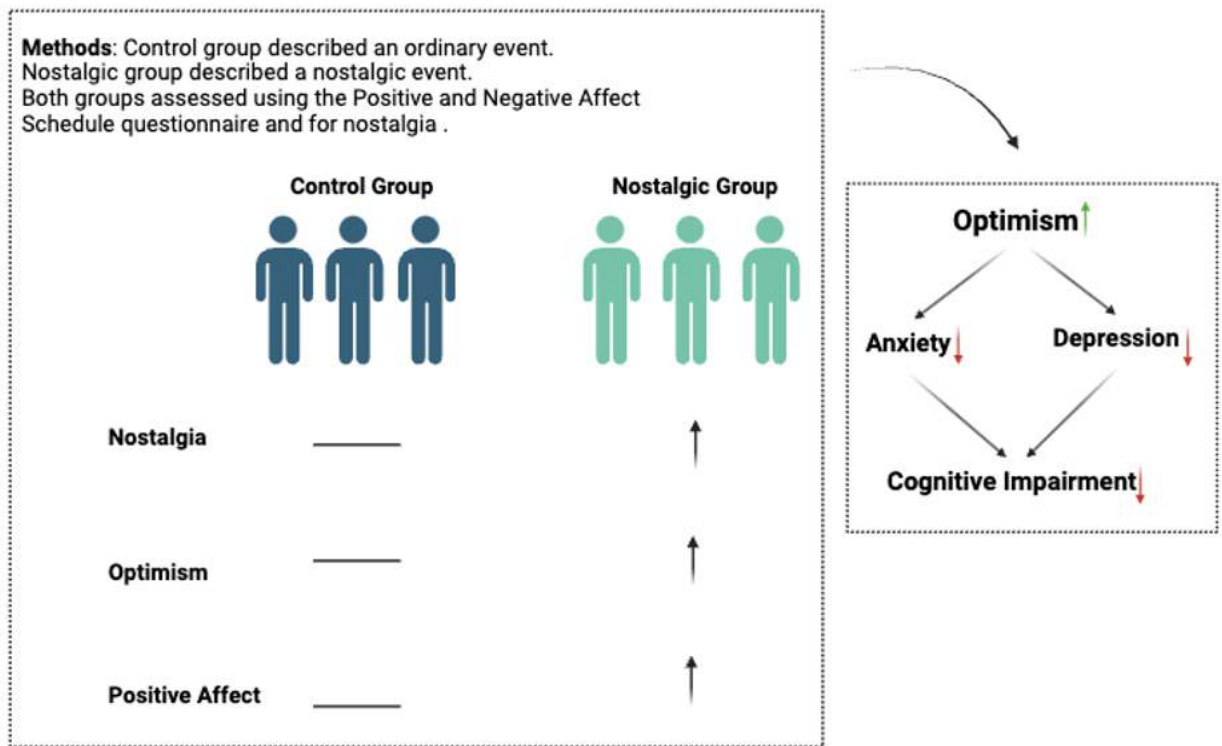
1, using 127 participants who completed measures of nostalgia, positive affect, and optimism. Like Study 1, the nostalgia condition reported more feelings of nostalgia and experienced greater positive affect (Figure 4). Additionally, optimism was expressed more frequently in the nostalgic condition compared to the control condition. These consistent findings underscore optimism as another behavior linked to nostalgia. Engaging in nostalgia evidently boosts both positive affect and optimism. This connection is also relevant to cognitive and mental health in dementia patients, as it can serve as a buffer from symptoms of depression and anxiety common in individuals cognitively impaired.

To grasp the significance of nostalgia and optimism in alleviating depression and anxiety in dementia patients, it is essential to first understand the severity and manifestations of these conditions in individuals with dementia. According to Kitching , depression occurs in about 30% of vascular dementia and Alzheimer's cases, and over 40% in dementia linked to Parkinson's and Huntington's diseases. Mood is commonly negative, and ideas of hopelessness are frequent. In addition, cognitive health may decline further due to depression and its associated emotions. Similarly, anxiety is common in patients with Alzheimer's disease and other forms of dementia such as vascular dementia and frontotemporal dementia. In a group of 191 subjects, caregivers reported anxiety in 30 individuals out of 115 (26.1%) with probable Alzheimer's disease, 22 out of 43 (51.2%) patients with vascular dementia, and 18 out of 33 (54.5%) patients with frontotemporal dementia (Porter, et al.). These numbers highlight the significant prevalence of anxiety and depression in dementia patients, which can severely impact their quality of life and overall well-being.

One promising approach to addressing this issue is fostering a sense of optimism among patients, which has shown potential in alleviating symptoms of anxiety and depression in individuals with dementia. An overview exploring the relation of optimism and mental/physical health relays information of optimism and depressive symptoms/suicidal ideation having an inverse correlation (Conversano, et al.). Thus, the sense of hopelessness commonly felt in cognitively impaired individuals may also be inversely correlated with optimism. In addition, pessimists who had less hope about the future were determined to be at higher risk of depressive and anxiety disorders. Optimism's relationship with depression and anxiety highlights that an increase in optimism is associated with a decrease in depressive symptoms and suicidal thoughts. This suggests that fostering optimism can lead to improved mental health, particularly in relation to depressive disorders. The significance of this information in dementia patients is outlined in (Pellegrino, et al.), where depression is suggested to be a risk factor and a prodrome of cognitive impairment. It can also be a response to cognitive impairment (Muliyala, and Varghese). As such, optimism has a role in protecting against cognitive impairment, since a sense of optimism decreases depressive symptoms which increase risk of conditions like dementia and Alzheimer's. Where anxiety is concerned, it is negatively correlated with optimism (Öcal, et al.) and positively correlated with dementia (Santabárbara, et al.). Accordingly, like depression, optimism can reduce anxiety, a disorder that increases risk of cognitive impairment and decline. Therefore, methods of increasing optimism through nostalgic experiences are potential therapeutic

approaches for the dementia population. Encouraging nostalgic engagement in cognitively impaired individuals is a viable approach to decreasing risk of dementia and advancement of the condition since depression and anxiety are decreased due to the optimism nostalgia elicits. Since depression and anxiety are common in dementia patients, using nostalgic reminiscence as a therapeutic strategy has the potential to improve the overall quality of life for these individuals.

Aside from reducing depression and anxiety, which in turn leads to protection against cognitive decline, optimism can also directly reduce risk of cognitive impairment. In a study by Gawronski, et al. , 4,624 participants aged 65 and over, an age group prone to cognitive impairment, were subject to a four-year period of data collection. Three kinds of assessments were taken: the cognitive impairment assessment, optimism measurement, and covariates measurement. In the cognitive impairment assessment, the modified Telephone Interview for Cognitive Status (TICS-M) was utilized. The TICS-M includes tests of memory, processing speed, naming, and orientation. In the optimism measurement, the six-item Life Orientation Test-Revised (LOT-R) was used, and a higher score was equivalent to higher optimism. Finally, the covariates measurement determined potential confounders linking optimism with cognitive impairment, such as gender, age, race, and depression/anxiety. By the end of the study, 4,065 of the 4,624 participants remained cognitively unimpaired. Results also indicated an inverse association between optimism and risk of cognitive impairment and a dose-response relationship between the two, where low optimism equated to highest risk, moderate optimism equated to somewhat reduced risk and high optimism equated to lowest risk. When accounting for depression and anxiety, the association between optimism and cognitive impairment remained unchanged. The study also corroborates the link of optimism to reduced likelihood of health conditions. Thus, beyond its role in decreasing depression and anxiety, which subsequently reduces the risk of cognitive impairment, optimism also directly protects against cognitive impairment (Figure 4). In both cases, however, nostalgia is an applicable method to increase optimism which will benefit the dementia population and older adults in general by protecting against cognitive impairment and possibly slowing cognitive decline.



Created in BioRender.com 

Figure 4: Nostalgic engagement leads to increased optimism and positive affect, behaviors which decrease anxiety/depression and may prevent cognitive impairment. Created with [BioRender.com](https://www.biorender.com).

Brain Areas

To further confirm the relationship between nostalgia and the behaviors it elicits (social connectedness, psychological well-being, and optimism), the subsequent section of this paper will examine the brain areas that are activated or related to both nostalgia and each of these behaviors. Nostalgia, the tool established to have significance in protecting against cognitive decline, is an emotion composed of multiple components. Authors Yang, et al. determine the brain areas that are associated with nostalgic engagement according to each of the core components which nostalgia encompasses: self-reflection, autobiographical memory, emotion regulation, and reward. The authors posit that brain areas associated with nostalgia are the same brain areas activated when engaging in the four core components. The key brain areas associated with self-reflection are the medial prefrontal cortex (mPFC) and the posterior cingulate cortex (PCC), which are responsible for tasks requiring self-reflection processing and self-consciousness, respectively. Brain areas associated with autobiographical memory processing are the hippocampus, mPFC, and PCC. The hippocampus plays a key role in memory function and the retrieval of autobiographical memory. Emotion regulation is associated with the

anterior cingulate cortex (ACC), which is involved in cognitive control of emotion, and the mPFC. Finally, brain regions associated with the reward network are the striatum, particularly the ventral striatum (VS) which is essential in reward processing, the substantia nigra (SN), the ventral tegmental area (VTA), and the ventromedial prefrontal cortex (vmPFC, which includes the medial orbitofrontal cortex (mOFC)). In summary, the brain areas associated with nostalgic engagement were hypothesized to be the mPFC, PCC, ACC, VS, SN, VTA, and vmPFC (including the mOFC). Another study complements and advances these conclusions (Oba, et al.). 14 participants contributed to a Functional Magnetic Resonance Imaging (fMRI) experiment, wherein visual stimuli were shown with objects and scenes that were emotionally neutral, but could both induce nostalgia and not induce nostalgia. The results of the study indicated brain activity in the hippocampus for activation of memory and in the SN/VTA and VS, reward related areas. In line with the previous report, the authors similarly found nostalgia to be associated with the components of memory retrieval and reward networks, resulting in similar brain areas being activated.

Social connectedness, a behavior that is increased by nostalgic engagement, is associated with several of the same brain areas as nostalgia. A review by Kim, and Sul found lonely individuals experienced less activity of the VS, and had less functional connectivity between the anterior insula and precuneus, areas involved in social cognitive processing. Meanwhile, VS activity increased when seeing the faces of close others, suggesting VS activity is significant in maintaining social connectedness and decreasing perceived social isolation. fMRI studies have also established the mPFC as a potential brain region negatively affected by loneliness. Other fMRI studies have revealed the reward system is activated when an individual is accepted or experiences a socially desirable outcome. Brain areas in the reward system activated due to social acceptance are the vmPFC, ventral anterior cingulate cortex (vACC) and VS (Kim, and Sul). In summary, activity in the VS, mPFC, vmPFC, and vACC are associated with social acceptance and thus, increased social connectedness. From the conclusions of this review, it is evident that nostalgic engagement and social connectedness involve multiple of the same brain areas, such as the mPFC, vmPFC, VS, and ACC, which taken together are regions involved in all four components of nostalgia established in Yang, et al. . Similar brain areas between nostalgia and social connectedness reaffirm the connection between the two behaviors and draws attention to their positive relationship. An increase in nostalgic reminiscence, therefore, appears to be associated with increased activity in the same brain areas involved in social acceptance and connectedness.

The brain areas involved in psychological well-being also correlate with brain activity during nostalgic reminiscence. Regarding self-esteem specifically, Agroskin, et al. utilized 48 participants who went through MRI scans and answered questionnaires that included a measure of self-esteem. Results from the MRI and self-esteem measure showed a positive relation between trait self-esteem and gray matter volume (GMV) in the ACC, right lateral prefrontal cortex (rLPFC), hypothalamus, and right hippocampus. These regions are known to play a role in emotion regulation and autobiographical memory, overlapping the areas relevant to both

nostalgia and social connectedness. Since self-esteem is also a behavior that aids in threat management, increasing activity and maintaining GMV in these brain areas serves to both increase self-esteem and reduce sense of threat. Indeed, the results of the study suggested an association between reduced GMV of the ACC and an inclination to ruminate, which refers to the tendency to dwell on negative feelings. Self-continuity, another aspect of psychological well-being, involves similar brain areas of the vmPFC and medial prefrontal regions ranging from the mPFC to the rostral anterior cingulate cortex (rACC) (C. Sedikides, et al.). The same brain areas are seen to be active/have higher GMV in psychological well-being as social connectedness and nostalgia, reinforcing the connection these behaviors have with nostalgic engagement.

Optimism, another behavior elicited by nostalgia, is associated with similar brain regions. In a study by Dolcos, et al. , 61 healthy participants underwent MRI scans and completed personality measures assessing trait optimism using the Life Orientation Test, trait anxiety using the State-Trait Anxiety Inventory-Trait, depression symptoms with the Beck Depression Inventory, and positive/negative affect with the Positive and Negative Affect Schedule-Trait. Increased trait optimism was found to be linked with increased GMV in the orbitofrontal cortex (OFC) and decreased anxiety, decreased negative affect, and less stressful experiences. The OFC is an area of the PFC that is involved in reward-related processing. The left rACC was also positively correlated with optimism. These results are supplemented by a literature review establishing the ACC and inferior frontal gyrus (IFG) as areas linked to optimism. The IFG is another region located in the PFC, and is involved in language processing and possibly empathy processing (Liakakis, et al.).

As seen through the preceding studies, the brain areas involved in nostalgia are also involved in the behaviors nostalgia elicits: social connectedness, psychological well-being (self-esteem and self-continuity) and optimism. The overlapping brain areas reaffirm nostalgia's ability to increase mechanisms beneficial to dementia patients. The PFC, ACC, VS, and hippocampus along with other regions involved in reward processing and emotion regulation are commonly seen to have increased volume or higher activity when engaging in nostalgia, and those same regions demonstrate increased volume and higher activity in nostalgia's byproducts (Figure 5). Consequently, the association between these behaviors is strengthened, enabling an investigation into the correlation between brain regions affected by dementia/Alzheimer's disease and the potential viability of nostalgia as a neurologically backed treatment.

Alzheimer's disease and dementia both involve loss of memory and cognitive ability, skills which the hippocampus is responsible for. Raji, et al. establish the anterior hippocampal/parahippocampal regions and precuneus as areas where gray matter atrophy occurs in AD. The hippocampus is involved in many of the behaviors mentioned before, such as nostalgia and self-esteem. This proves problematic as it is difficult to protect against cognitive decline using nostalgia and self-esteem in patients where gray matter damage has already occurred, possibly preventing the benefits that nostalgia and self-esteem can bring. Since gray matter atrophy cannot be reversed, the ability to increase self-esteem and experience nostalgia

might seem impaired in AD patients. However, as previously established, AD patients are indeed capable of nostalgic reminiscence, and stimulating existing neurons in the hippocampus to increase activity may be a possible approach to increasing psychological well-being. Thus, although GMV in the hippocampus is decreased in AD, increasing neuronal activity through nostalgic engagement is a potential avenue to protect against further cognitive decline. Similarly, targeted stimulation of neurons in other brain areas affected by AD/dementia may maintain the function of nostalgia as a method to preserve cognitive ability. As for other regions involved in nostalgia and its associated behaviors, dementia only affects the inferior PFC regions (Salat, et al.), which don't include the mPFC or vmPFC which are relevant to nostalgia. Therefore, although atrophy of specific brain areas also involved in nostalgia is present in AD, engaging in nostalgic reminiscence remains a possible option to protect against further cognitive decline.

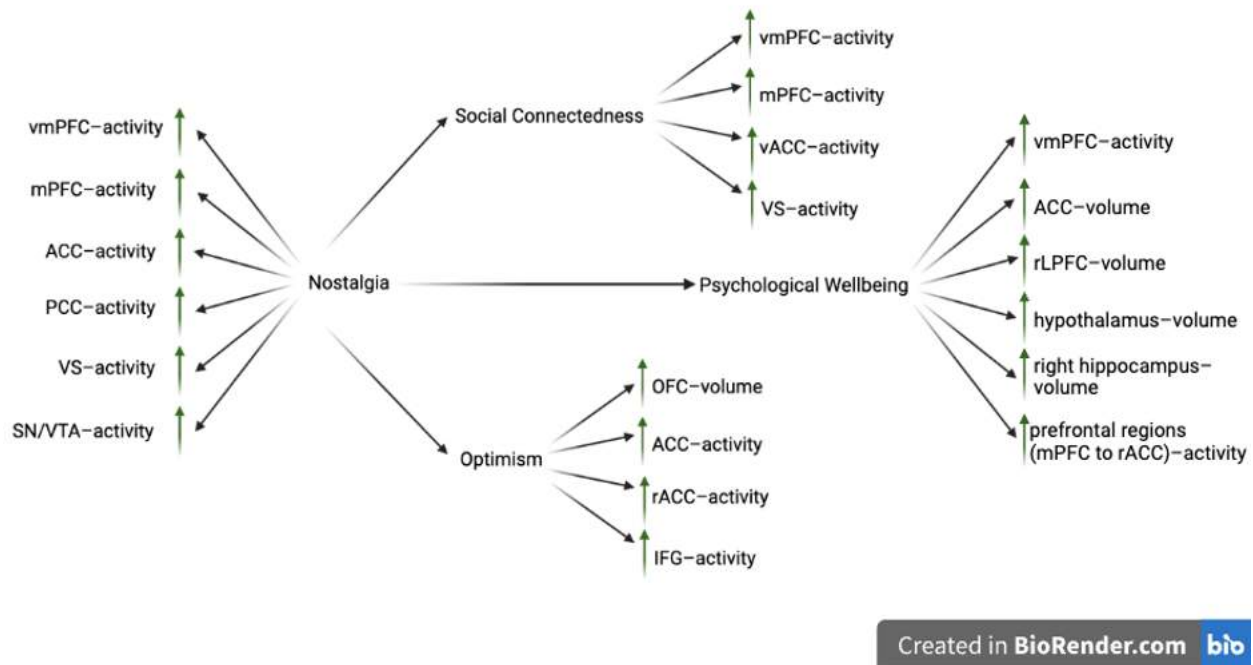


Figure 5: Nostalgia, social connectedness, psychological well-being, optimism, and their associated brain areas. Created with BioRender.com.

Music As An Intervention

Knowing the positive impact nostalgia has on individuals with cognitive impairment, it is important to identify effective methods for inducing nostalgia. Throughout this paper, nostalgia has most frequently been induced through memory recall or narration, however, nostalgia can also be evoked through scent (Reid, et al.), movies (Rasmussen, et al.), objects, and music.

Music is a frequently studied approach to inducing nostalgia. Its psychological benefits also correlate with many behaviors caused by nostalgic reminiscence. A literature review by Constantine Sedikides, et al. outlines these benefits, specifically when nostalgia is evoked

through music. It is first established that music is a powerful source of nostalgia, especially because of the “reminiscence bump”, a period in an individual’s life, usually the teenage to early adult years, which are formative in one’s music preferences for life. Nostalgic music was found to increase self-esteem, meaning in life, optimism, and social connectedness through triggering memories of loved ones. In the dementia population, the review stated brain areas associated with musical memory remained undamaged in dementia and may be used to alleviate discomfort and increase general psychological well-being. This information is complemented by several previously referenced studies. Studies 3 and 4 in the study by Cheung, et al. utilize music as the method to induce nostalgia and participants. The findings of Study 3 showed that listening to nostalgic songs increased optimism and self-esteem more than listening to songs without nostalgic significance. Study 4 utilized lyrics of nostalgic songs instead of inducing nostalgia with music by audio but found similar results of higher levels of social connectedness, self-esteem, and optimism in the nostalgic condition. These results not only reaffirm music’s ability to evoke nostalgia, but also demonstrate music’s role in the behaviors nostalgia elicits. Another previously referenced review (Yang, et al.) notes that the hippocampus is involved in music-evoked emotions, and the mPFC was engaged in response to autobiographically salient music pieces that triggered personal memories. Oba, et al. supplemented these results by drawing to attention an association of music-evoked nostalgia and the activation of the hippocampus, vmPFC, and VS, a simultaneous reaction of the reward and memory systems. These are the same brain regions relevant to nostalgia, social connectedness, and psychological well-being.

Music as a method of eliciting nostalgia has also been shown to be highly effective, even in comparison to other approaches such as pictures and other stimuli. Authors Kaiser, and Berntsen examined the differing characteristics of music-evoked autobiographical memories through a literature review, demonstrating music is capable of evoking autobiographical memories and is more effective than silence in AD, behavioral variant frontotemporal dementia (bv-FTD), brain damage, and depression. The memories were specific, positive (except for the dementia group, which indicated an equal number of positive and negative memories) and retrieved quickly. In AD patients specifically, fewer autobiographical memories were retrieved after being shown pictures compared to after listening to familiar music, making music a more effective tool in AD to reap the therapeutic benefits of nostalgia. In addition, specificity of the memories was increased when retrieved through music. In patients with bv-FTD, frequency and specificity of music-evoked autobiographical memories were initially worse compared to memories evoked by pictures, however, re-exposure marked an improvement in frequency and specificity. The results of these studies suggest a significant potential for music to evoke nostalgia and improvement in autobiographical memory recall, especially in cases of AD and dementia. In another study testing the effects of combined nostalgic music stimulation and sensory activities on patients with mild cognitive impairment, 49 participants were divided into experimental and control groups (Lee, et al.). The experimental group underwent sessions of five-sense activities (activities that targeted use of the five senses) associated with nostalgic

music, each lasting 2 hours and occurring 2 times per week for 5 weeks. The control group did not receive any stimulation. Results of this study indicated an increase in cognitive function and a decrease in geriatric depression in the experimental group. Such an outcome draws attention to music's ability to reactivate or stimulate neuronal activity and connections in people with forms of cognitive impairment, causing reminiscence of past moments and enhancing memory and cognitive function.

Evidently, incorporating nostalgic music into the treatment of dementia patients, especially when activating all five senses, encourages improvement of cognition. There are additional strategies that can be used while utilizing nostalgic music to maximize therapeutic benefit. Rao, et al. demonstrated the importance of considering the reminiscence bump in creating a treatment approach for dementia patients. The reminiscence bump is a period in one's life, 10-30 years old, which elderly people tend to revisit most often (Munawar, et al.). This is a result of the age range 10-30 holding memories of important events and being most formative in shaping different preferences, music being one of them (Constantine Sedikides, et al.). According to (Rao, et al.), music from the reminiscence bump period is most readily recalled. This information is corroborated in Kaiser, and Berntsen, where music from the reminiscence bump period of AD participants evoked more autobiographical memories. Thus, a potential approach to protecting against cognitive decline through music may be to create playlists that specifically include songs from the respective reminiscence bump periods of different patients to maximize impact of the nostalgic experience.

Conclusion

The exploration of nostalgia, particularly when induced by music, as a therapeutic intervention for dementia, presents a promising avenue in advancing understanding about neurodegenerative disorders. Because dementia patients retain the ability to experience nostalgia, the behaviors which it evokes, those behaviors being increased social connectedness, psychological well-being, and optimism, may also have a positive effect on these individuals' health. An improvement of these behaviors is significant given their ability to protect against cognitive decline and other conditions such as depression and anxiety. Thus, nostalgia is important to consider when finding ways to prevent dementia worsening. In addition, several brain regions involved in the experience of nostalgia overlap with those associated with social connectedness, psychological well-being, and optimism, further strengthening the connection between nostalgia and its positive effects. Moreover, although brain atrophy from dementia or AD cannot be reversed, stimulation of certain areas through nostalgic reminiscence may have the potential to decrease rapid worsening of the condition.

Music is a commonly used method for evoking nostalgia, often employed in studies to explore the effects of nostalgic reminiscence in both cognitively normal individuals and dementia patients. In many cases, participants are instructed to listen to specific songs that trigger nostalgic memories, a strategy that has consistently proven effective. Therefore, it is important to study ways in which the effects of music-induced nostalgia can be maximized to

provide fuller protection against cognitive decline. Creating personalized playlists according to an individual's reminiscence bump period is a promising strategy, but further exploration is needed to identify additional methods that could enhance the therapeutic benefits. More research on the intersection of music, nostalgia, and dementia is essential to developing effective methods for improving the quality of life for dementia patients.

Works Cited

- Agroskin, D. et al. "The Self-Liking Brain: A Vbm Study on the Structural Substrate of Self-Esteem." *PLoS One*, vol. 9, no. 1, 2014, p. e86430, doi:10.1371/journal.pone.0086430.
- Cerejeira, J. et al. "Behavioral and Psychological Symptoms of Dementia." *Front Neurol*, vol. 3, 2012, p. 73, doi:10.3389/fneur.2012.00073.
- Cheung, W. Y. et al. "Back to the Future: Nostalgia Increases Optimism." *Pers Soc Psychol Bull*, vol. 39, no. 11, 2013, pp. 1484-96, Medline, doi:10.1177/0146167213499187.
- Conversano, C. et al. "Optimism and Its Impact on Mental and Physical Well-Being." *Clin Pract Epidemiol Ment Health*, vol. 6, 2010, pp. 25-9, doi:10.2174/1745017901006010025.
- Dolcos, S. et al. "Optimism and the Brain: Trait Optimism Mediates the Protective Role of the Orbitofrontal Cortex Gray Matter Volume against Anxiety." *Soc Cogn Affect Neurosci*, vol. 11, no. 2, 2016, pp. 263-71, Medline, doi:10.1093/scan/nsv106.
- Donovan, N. J. and D. Blazer. "Social Isolation and Loneliness in Older Adults: Review and Commentary of a National Academies Report." *Am J Geriatr Psychiatry*, vol. 28, no. 12, 2020, pp. 1233-44, doi:10.1016/j.jagp.2020.08.005.
- El Haj, M. et al. "Phenomenological Reliving and Visual Imagery during Autobiographical Recall in Alzheimer's Disease." *J Alzheimers Dis*, vol. 52, no. 2, 2016, pp. 421-31, Medline, doi:10.3233/JAD-151122.
- Fleury, J. et al. "Feeling Safe and Nostalgia in Healthy Aging." *Front Psychol*, vol. 13, 2022, p. 843051, PubMed-not-MEDLINE, doi:10.3389/fpsyg.2022.843051.
- Gardener, H. et al. "Social Connectivity Is Related to Mild Cognitive Impairment and Dementia." *J Alzheimers Dis*, vol. 84, no. 4, 2021, pp. 1811-20, Medline, doi:10.3233/JAD-210519.
- Gawronski, K. A. et al. "Dispositional Optimism and Incidence of Cognitive Impairment in Older Adults." *Psychosom Med*, vol. 78, no. 7, 2016, pp. 819-28, Medline, doi:10.1097/PSY.0000000000000345.
- Hackett, R. A. et al. "Social Engagement before and after Dementia Diagnosis in the English Longitudinal Study of Ageing." *PLoS One*, vol. 14, no. 8, 2019, p. e0220195, doi:10.1371/journal.pone.0220195.
- Isene, T. A. et al. "Meaning in Life for Patients with Severe Dementia: A Qualitative Study of Healthcare Professionals' Interpretations." *Front Psychol*, vol. 12, 2021, p. 701353, doi:10.3389/fpsyg.2021.701353.
- Ismail, S. et al. "Psychological and Mnemonic Benefits of Nostalgia for People with Dementia." *J Alzheimers Dis*, vol. 65, no. 4, 2018, pp. 1327-44, Medline, doi:10.3233/JAD-180075.
- Ismail, S. et al. "The Content of Nostalgic Memories among People Living with Dementia." *Int J Aging Hum Dev*, vol. 94, no. 4, 2022, pp. 436-58, Medline, doi:10.1177/00914150211024185.
- Kaiser, A. P. and D. Berntsen. "The Cognitive Characteristics of Music-Evoked Autobiographical Memories: Evidence from a Systematic Review of Clinical

- Investigations." *Wiley Interdiscip Rev Cogn Sci*, vol. 14, no. 3, 2023, p. e1627, Medline, doi:10.1002/wcs.1627.
- Kim, M. J. and S. Sul. "On the Relationship between the Social Brain, Social Connectedness, and Wellbeing." *Front Psychiatry*, vol. 14, 2023, p. 1112438, doi:10.3389/fpsy.2023.1112438.
- Kitching, D. "Depression in Dementia." *Aust Prescr*, vol. 38, no. 6, 2015, pp. 209-211, doi:10.18773/austprescr.2015.071.
- Lee, Pai-Lin et al. *Nostalgic Music Stimulation and Five-Sense Activities: A Cognitive Enhancement Intervention for Mild Cognitive Impairment*. 2024.
- Leunissen, Joost et al. "The Hedonic Character of Nostalgia: An Integrative Data Analysis." *Emotion Review*, vol. 13, no. 2, 2021, pp. 139-56, doi:10.1177/1754073920950455.
- Liakakis, G. et al. "Diversity of the Inferior Frontal Gyrus—a Meta-Analysis of Neuroimaging Studies." *Behavioural Brain Research*, vol. 225, no. 1, 2011, pp. 341-47, doi:https://doi.org/10.1016/j.bbr.2011.06.022.
- Mills, M. A. and P. G. Coleman. "Nostalgic Memories in Dementia--a Case Study." *Int J Aging Hum Dev*, vol. 38, no. 3, 1994, pp. 203-19, Medline, doi:10.2190/NCAJ-0G0L-VTQ4-V1L8.
- Minyo, M. J. and K. S. Judge. "Perceived Unmet Need and Need-Related Distress of People Living with Dementia." *Gerontol Geriatr Med*, vol. 8, 2022, p. 23337214221092886, doi:10.1177/23337214221092886.
- Muliyala, K. P. and M. Varghese. "The Complex Relationship between Depression and Dementia." *Ann Indian Acad Neurol*, vol. 13, no. Suppl 2, 2010, pp. S69-73, doi:10.4103/0972-2327.74248.
- Munawar, K. et al. "Understanding the Reminiscence Bump: A Systematic Review." *PLoS One*, vol. 13, no. 12, 2018, p. e0208595, doi:10.1371/journal.pone.0208595.
- Oba, K. et al. "Memory and Reward Systems Coproduce 'Nostalgic' Experiences in the Brain." *Soc Cogn Affect Neurosci*, vol. 11, no. 7, 2016, pp. 1069-77, Medline, doi:10.1093/scan/nsv073.
- Öcal, E. E. et al. "Relationship between Mental Disorders and Optimism in a Community-Based Sample of Adults." *Behav Sci (Basel)*, vol. 12, no. 2, 2022, doi:10.3390/bs12020052.
- Pellegrino, L. D. et al. "Depression in Cognitive Impairment." *Curr Psychiatry Rep*, vol. 15, no. 9, 2013, p. 384, doi:10.1007/s11920-013-0384-1.
- Piolatto, M. et al. "The Effect of Social Relationships on Cognitive Decline in Older Adults: An Updated Systematic Review and Meta-Analysis of Longitudinal Cohort Studies." *BMC Public Health*, vol. 22, no. 1, 2022, p. 278, Medline, doi:10.1186/s12889-022-12567-5.
- Porter, V. R. et al. "Frequency and Characteristics of Anxiety among Patients with Alzheimer's Disease and Related Dementias." *J Neuropsychiatry Clin Neurosci*, vol. 15, no. 2, 2003, pp. 180-6, doi:10.1176/jnp.15.2.180.
- Raji, C. A. et al. "Age, Alzheimer Disease, and Brain Structure." *Neurology*, vol. 73, no. 22, 2009, pp. 1899-905, doi:10.1212/WNL.0b013e3181c3f293.

- Rao, C. B. et al. "A Focus on the Reminiscence Bump to Personalize Music Playlists for Dementia." *J Multidiscip Healthc*, vol. 14, 2021, pp. 2195-204, PubMed-not-MEDLINE, doi:10.2147/JMDH.S312725.
- Rasmussen, Katrine W. et al. "Using Nostalgia Films to Stimulate Spontaneous Autobiographical Remembering in Alzheimer's Disease." *Journal of Applied Research in Memory and Cognition*, vol. 10, no. 3, 2021, pp. 400-11, doi:https://doi.org/10.1016/j.jarmac.2020.11.001.
- Reid, C. A. et al. "Scent-Evoked Nostalgia." *Memory*, vol. 23, no. 2, 2015, pp. 157-66, doi:10.1080/09658211.2013.876048.
- Salat, David H. et al. "Selective Preservation and Degeneration within the Prefrontal Cortex in Aging and Alzheimer Disease." *Archives of Neurology*, vol. 58, no. 9, 2001, pp. 1403-08, doi:10.1001/archneur.58.9.1403.
- Santabàrbara, J. et al. "Does Anxiety Increase the Risk of All-Cause Dementia? An Updated Meta-Analysis of Prospective Cohort Studies." *J Clin Med*, vol. 9, no. 6, 2020, doi:10.3390/jcm9061791.
- Sedikides, C. et al. "Self-Continuity." *Annu Rev Psychol*, vol. 74, 2023, pp. 333-61, doi:10.1146/annurev-psych-032420-032236.
- Sedikides, Constantine et al. "The Psychological Benefits of Music-Evoked Nostalgia." *Psychology of Music*, vol. 50, no. 6, 2021, pp. 2044-62, doi:10.1177/03057356211064641.
- Toki, H. et al. "Maintaining Continuity of Self as Perceived by People in the Early Stages of Dementia: A Qualitative Study." *Florence Nightingale J Nurs*, vol. 31, no. 1, 2023, pp. 56-61, doi:10.5152/fnjn.2023.22115.
- Umar Ismail, S. et al. "Nostalgia as a Psychological Resource for People with Dementia: A Systematic Review and Meta-Analysis of Evidence of Effectiveness from Experimental Studies." *Dementia (London)*, vol. 19, no. 2, 2020, pp. 330-51, doi:10.1177/1471301218774909.
- Wildschut, T. et al. "Nostalgia: Content, Triggers, Functions." *J Pers Soc Psychol*, vol. 91, no. 5, 2006, pp. 975-93, Medline, doi:10.1037/0022-3514.91.5.975.
- Yang, Z. et al. "Patterns of Brain Activity Associated with Nostalgia: A Social-Cognitive Neuroscience Perspective." *Soc Cogn Affect Neurosci*, vol. 17, no. 12, 2022, pp. 1131-44, Medline, doi:10.1093/scan/nsac036.
- Zhou, X. et al. "Counteracting Loneliness: On the Restorative Function of Nostalgia." *Psychol Sci*, vol. 19, no. 10, 2008, pp. 1023-9, Medline, doi:10.1111/j.1467-9280.2008.02194.x.

Novel Applications of Elastin-like Polypeptides (ELPs) in the Treatment of Pathogenic Free-Living Amoeba By Avni Kalra

Abstract

Primary amoebic meningoencephalitis (PAM) is a rare and acute yet fulminant infection caused by the amoeba *Naegleria fowleri*. PAM is characterized by headaches, fever, nausea, and stiff neck. Although rare, PAM is fatal, with a mortality rate of 98% and causes death within two weeks of exposure. Granulomatous Amoebic Encephalitis (GAE) is an opportunistic infection with non-specific symptoms such as fevers, headaches, nausea, and vomiting. Similar to PAM, it has a staggering mortality rate of 97-98%. There are several key factors involved in the high mortality rate for both indications, including the ineffectiveness of common treatments such as amphotericin B, fluconazole, azithromycin, and Rifampin alongside poor penetration of the blood-brain barrier (BBB). Elastin-like polypeptides (ELPs) are biopolymeric nanoparticles that mimic the properties of natural elastin, a key component of the extracellular matrix found in connective tissue. ELPs are specifically characterized by their biocompatibility, targeted and controlled release, phase change behavior, and the ability to encapsulate multiple drugs. While ELPs have been extensively researched in the context of various diseases, their potential in treating PAM and GAE remains an unexplored area of interest. This paper therefore focuses on possible approaches in which ELPs might be leveraged to increase the efficacy of existing treatments for PAM and GAE. By imagining how ELP nanomedicines could be applied for novel therapeutic strategies against PAM and GAE, we hope to inspire future translational avenues for this rare disease to improve patient outcomes.

Primary Amoebic Meningoencephalitis

Primary amoebic meningoencephalitis, more commonly known as PAM, is a rare but fulminating amoebic infection caused by the free-living amoeba (FLA), *Naegleria fowleri*. *N. fowleri* is indigenous to freshwater environments such as freshwater habitats, still water lakes, rivers, and other aqueous bodies (Grace et al.). Although PAM is associated with numerous neurological manifestations, the avenue through which *N. fowleri* accesses the central nervous system (CNS) is singular: the nasal cavity. A person generally contracts *N. fowleri* by participating in recreational water activities, such as swimming or diving in which water inadvertently enters the nasal cavity (Güémez and García).

The amoeba has a predilection for warm freshwater environments which are ideal for its growth and reproduction. The nasal cavity, to a certain degree, replicates the environment *N. fowleri* is found in, creating ideal conditions for the amoeba to thrive (Tillery et al.). *N. fowleri* initiates its pathogenesis by breaching the nasal mucosa and cribriform plate. Subsequently, it utilizes the olfactory nerve bundle to make its way directly to the CNS where it corrodes brain tissue. This method of infection is highly beneficial for *N. fowleri* from an evolutionary standpoint due to the anatomical proximity of the olfactory nerve bundles to the CNS, thereby

exacerbating the speed at which the infection progresses (Grace et al.; Herman et al.; Visvesvara et al.).

N. fowleri enters the nasal cavity in trophozoite form, which is the amoeba's active infectious state as opposed to a cyst form — the dormant form of the amoeba adopted when environmental conditions are not suitable ("Naegleria"). This trophozoite stage possesses food cups enabling the amoeba to destroy tissue in the CNS (Rodríguez-Mera et al.; JOHN et al.). In tandem with the food cups for tissue destruction, *N. fowleri* releases cytolytic molecules which exacerbate nerve destruction (Chen et al.; Marciano-Cabral and Cabral). These molecular invasions result in a constellation of neurological symptoms such as severe headaches, nausea, seizures, hallucinations and photophobia (Grace et al.).

Prior to understanding treatment for alleviating symptoms of PAM, it is first imperative to understand the body's innate immune responses to the amoebal assault. The primary line of defense against *N. fowleri* is the complement immune system — a part of the innate immune system that enhances the ability of antibodies to attack the pathogen's cell membrane (Jamerson et al.). The complement system includes macrophages and neutrophils. Macrophages and neutrophils together orchestrate an immune response wherein neutrophils pinch and engulf the *N. fowleri* trophozoite, and macrophages release non-oxidative mediators such as TNF- α 6 (Oh et al.; Moseman). In conjunction with the complement immune system, the mucosal epithelial tissues have innate responses to the pathogen. Collectively, they form a mucosal cell lining which acts as a barrier between the host's nasal cavity and all structures located above it (Linden et al.). Additionally, the mucosal epithelial cells secrete chemical defensive compounds such as mucins, antibodies, defensins, protegrins, collectins, cathelicidins, lysozyme, histatins, and nitric oxide (Cervantes-Sandoval et al.; Marciano-Cabral and Cabral). Despite this detailed and extensive immune response, the human body's immune response often proves insufficient to halt *N. fowleri* infection and the rapid onset of PAM. Consequently, it is not surprising that PAM remains an insurmountable infection with a mortality rate of 97% (Vugia et al.)

Diagnosis

Due to its rapid progression and rarity, a timely diagnosis of PAM remains a formidable obstacle. In fact, most diagnoses of PAM occur post-mortem, and autopsies typically reveal herniation in the frontal lobe alongside loss of gray matter in that region. PAM is officially diagnosed using a lumbar puncture for cerebrospinal fluid (CSF) analysis (Pana et al.). A definitive diagnosis can be made upon an observation of motile *N. fowleri* trophozoites on centrifuged CSF. PAM can also be diagnosed using laboratory testing for *Naegleria fowleri* nucleic acid in CSF, biopsy, tissue specimens, or *N. fowleri* antigen in CSF (Güémez and García).

Current Treatments

Current treatments rely upon a timely diagnosis, broad spectrum antifungal drugs, and therapeutic hypothermia to manage the inflammation. According to the CDC, PAM is treated with a combination of drugs that will be briefly described below:

Amphotericin B

Amphotericin B (AmB) is a polyene antifungal that also has an amoebicidal effect. It functions by binding to the ergosterols in the cell membrane. This causes pores in the cellular membrane which facilitate permeability. The porous fungal cellular membrane subsequently causes lysis and breaks down the cell structure of the amoeba (Pugh and Levy; Noor and Preuss). Amphotericin B, prescribed intravenously or intrathecally, has been used in all North American cases of PAM (Güemez and García). However, a high concentration of AmB is needed to reach the minimal inhibitory concentration (MIC); the MIC is the lowest dose that will inhibit growth of the microorganism (Andrews). A high concentration of AmB is required to reach the MIC and kill *N. fowleri* in the CNS because AmB demonstrates poor penetration of the BBB, is insoluble in aqueous solutions, and exhibits dose-limiting side effects including renal toxicity, anemia, chills, nausea, fever, vomiting, and headaches (Rajendran et al.; Grace et al).

Azithromycin

Azithromycin is a macrolide antibiotic, and it works by inhibiting bacterial protein synthesis to prevent the transit of aminoacyl-tRNA and the growing protein through the ribosome (Pugh and Levy). In the case of PAM, a study by Goswick and Brenner discovered the potential synergistic effects of azithromycin and amphotericin B. Although azithromycin has a MIC 123 times higher than amphotericin B, it still demonstrates greater in vivo effects than amphotericin B due to its unique pharmacokinetic profile, which include a long-elimination half-life and high tissue accumulation levels (Goswick and Brenner).

Fluconazole

Fluconazole is an antifungal agent, and just like Azithromycin, has shown synergistic effects with amphotericin B in the treatment of PAM. It does so by inhibiting ergosterol synthesis (Pugh and Levy). Fluconazole's synergistic effects can be attributed to its ability to enhance the bactericidal effects of neutrophils. Beyond its primary antifungal function, fluconazole appears to have immunomodulatory effects by assisting neutrophils in the immune response (Zervos et al.).

Rifampin

Rifampin is typically used to treat bacterial infections but has also been administered in a select few cases of PAM. This is because it does not reach a sufficient concentration in the CNS at standard doses. Although it reaches an adequate concentration in the CSF, which bathes the CNS, compartmental concentrations in the CNS vary (Güemez and García). Another concern with Rifampin is that it induces certain liver enzymes. These liver enzymes hinder the

pharmacokinetics of other drugs during combination therapy. As an example, rifampin increases fluconazole's clearance rate and ultimately reduces the half-life of fluconazole (Ayudhya et al.)

Miltefosine is an antineoplastic agent employed primarily for the treatment of breast cancer and leishmaniasis (Pugh and Levy). Its uses have also been explored for FLA. By 2013, the CDC reported the use of miltefosine for 26 cases of PAM (Alli et al.) It is categorized as a broad-spectrum phospholipid antimicrobial agent related to the signaling molecule lysophosphatidylcholine. Additionally, the phospholipid in miltefosine possesses an attached alkyl phosphocholine. The molecule is amphiphilic, having a polar phosphocholine head region and an aliphatic tail. It exists in a zwitterionic form with a permanently charged quaternary ammonium ion and anionic phosphate (Grace et al). This is relevant because lysophosphatidylcholine's permanently charged nature is responsible for its poor CNS penetration and results in a high MIC (Roy et al.; Grace et al).

Granulomatous Amoebic Encephalitis

Granulomatous amoebic encephalitis, better known as GAE, is an opportunistic infection with a staggering mortality rate of 97%-98%. Unlike PAM, GAE is caused by three different pathogens: *Acanthamoeba* spp, *Balamuthia mandrillaris*, and *Sappinia pedata*. GAE primarily affects immunocompromised individuals (Siripurapu et al.; Parija et al.). It is microscopically detected through a lung, sinus, brain tissue, or skin biopsy conducted postmortem. Diagnosing GAE can be challenging due to its rarity and nonspecific symptoms, which include headaches, fever, nausea, vomiting, and neurological deficits (Keane et al.). A definitive diagnosis involves cerebrospinal fluid analysis or neuroimaging studies such as MRI or CT scans. The optimal treatment for GAE involves polymicrobial microbial therapy coupled with the resection of lesions to better control the infection (Da Rocha-Azevedo et al.).

***Acanthamoeba* spp.**

Being the primary etiological factor contributing to GAE, *Acanthamoeba* spp. are one of the most prevalent environmental protozoa because they are found in sea water, tap water, swimming pools, natural thermal water, soil, and dust (Kalra et al.). *Acanthamoeba* are both free living and parasitic. Like *N. fowleri*, they can appear in both the actively feeding and dividing trophozoite stage or the dormant cyst form (Wang et al.). *Acanthamoeba*'s amoeboid locomotion can be attributed to acanthopodia, which are spiny surface structures, alongside the formation of hyaline pseudopodia. They possess a central nucleus along with a nucleolus and a single pulsating vacuole (Siddiqui and Khan).

There are direct and indirect factors contributing to the pathogenicity of *Acanthamoeba*; direct factors include phagocytosis and the ability to produce pre-forming toxins like acanthaporin which exhibit cytotoxicity for human neuronal cells and contribute to neural tissue damage (Siddiqui and Khan; Duggal et al.). Indirect factors, by contrast, don't directly contribute to virulence but can nevertheless affect pathogenicity. Some indirect factors include chemotaxis, the ability to interact with and/or form biofilms, the propensity to engage in encystation, and

possible interactions with bacterial endosymbionts contributing to *Acanthamoeba*'s pathogenesis (Cursons et al.).

Acanthamoeba exhibit multiple avenues of entry, the most prominent of which targets the nasal cavity. Infection occurs following inhalation of air or aspiration of water containing the trophozoite form of the amoeba (Kot et al.; Visvesvara et al.). The trophozoite subsequently makes its way to the CNS through the nasal mucosa and endothelium of the brain's capillaries (Chalmers). Secondly, trophozoites can also gain entry through ulcerated skin or oral mucosa, and once in the body, they enter the bloodstream and disseminate to different areas such as the (Ellison et al.).

Balamuthia Mandrillaris

B. mandrillaris is also considered free living and exhibits both cyst and trophozoite stages, allowing it to divide through binary fission (Bhosale and Parija). Its ecological niche is not widely known; however, it is closely related to *Acanthamoeba* based on RNA sequencing and phylogenetic analysis (Stothard et al.; Phan et al.; Cope et al.). *B. mandrillaris* was originally thought to infect the brain using the olfactory nerve bundles like *N. fowleri*. Histopathologic findings, though, do not show olfactory lobe involvement in the manner seen with *N. fowleri* (Cope et al.). Instead, *B. mandrillaris* enters the body through breaks in the skin or the respiratory tract by the inhalation of cysts. It primarily attacks two body systems: the brain, specifically the CNS, and the skin (Lee et al.).

Sappinia Pedata

S. pedata is another FLA responsible for causing GAE. Similar to *N. fowleri*, *Acanthamoeba* spp., and *B. mandrillaris*, *S. pedata* exists in both cyst and trophozoite forms. Its mechanism of infection is unknown. However, based on *Acanthamoeba* and *B. mandrillaris*'s mechanisms of infection, it is thought to be through the nasopharynx or may be introduced into the bloodstream (Da Rocha-Azevedo et al.). One person so far has been reported to have contracted GAE due to *S. pedata*. Due to this, generalized symptoms are unknown. However, this patient had a previous sinus infection and subsequently experienced nausea, vomiting, photophobia, blurry vision, and a loss of consciousness (CDC - DPDx - Free Living Amebic Infections; Qvarnstrom et al.). The patient was subsequently treated with azithromycin, pentamidine, itraconazole, and flucytosine (Da Rocha-Azevedo et al.; Qvarnstrom et al.).

Treatment

Taravoud, Fechtali-Moute, et al discussed the efficacy of drugs used in the treatment of GAE based on both in vitro and in vivo studies. Two Cotrimoxazole were used. Trimethoprim and Sulfamethoxazole specifically. Cotrimoxazole was the most frequently used drug for GAE treatment. It demonstrated a relatively modest in vivo efficiency, only 47% of patients treated with Cotrimoxazole survived. In vitro studies demonstrated that Cotrimoxazole did not exhibit amoebicidal activity below 100 milligrams/ml (Taravoud et al.).

Amphotericin B is another frequently used drug for the treatment of GAE. However, in vitro studies demonstrate a low efficacy and show a natural resilience of *Acanthamoeba* spp. to amphotericin B. Only 23% of patients treated with amphotericin B survived (Taravaud, Loiseau, et al.).

Rifampicin is frequently used in GAE treatment, but it does not show any anti-acanthamoebal activity in vitro. Nevertheless, it has lipophilic properties which allow it to penetrate the blood-brain barrier and contributes to its efficacy in vivo (Elsheikha et al.).

Among the azole-based drugs (e.g., fluconazole, ketoconazole, voriconazole, itraconazole), ketoconazole was the most efficient azole with an EDU of 46%. The most effective combination GAE treatment aligning with the recommendations from the Infectious Disease Society of America (IDSA) involves co-administration of rifampicin, cotrimoxazole, and ketoconazole. It is recognized and successful in 83% of cases (Taravaud et al.).

A similar treatment has been used for GAE and may vary due to the pathogen causing it. For example, in a survival case of GAE caused by *Balamuthia Mandrillaris*, the patient was treated with sulfamethoxazole, azithromycin, flucytosine, and amphotericin B (Chen et al.). In another surviving case of GAE caused by *Balamuthia Mandrillaris*, a non-conventional treatment of nitroxoline was used. Nitroxoline is primarily used to treat Urinary Tract Infections (UTI). Nitroxoline had demonstrated amoebicidal activity against *B. mandrillaris* in vitro. This patient's combination treatment involved nitroxoline, miltefosine, azithromycin, albendazole, fluconazole, and dose reduced flucytosine (Spottiswoode et al.).

Nanoparticles as a Novel Therapeutic Strategy

Nanoparticles typically range in size from one to 100 nanometers (nm) and have garnered significant attention for their versatile applications in the diagnosis, treatment, and prevention of various diseases (Joseph et al.). Their small size and large surface area to volume ratio particularly enable efficient drug delivery by: **a)** encapsulating therapeutic agents such as drugs; **b)** protecting the therapeutic payload from degradation; and **c)** enhancing their bioavailability (Elumalai et al.; Sultana et al.).

Nanoparticles can also be designed to exhibit specific behaviors including controlled release and active targeting of specific tissues and/or cells (Elumalai et al.). This essentially optimizes their therapeutic efficacy across various diseases such as cancer, infectious diseases, and neurological disorders (Zottel et al.; Tobin and Brenner).

Nanoparticles can also be effectively utilized within the context of combating FLA the conjugation of nanoparticles with drugs specifically holds significant promise for treating diseases caused by FLA by more effectively addressing the challenges seen with more conventional treatment methods (Siddiqui et al.; Fan et al.). By leveraging the unique properties of nanoparticles including their customizable surface drug conjugations, ability to improve drug solubility, enhanced stability, and potential for targeted delivery, nanoparticles can overcome current limitations associated with standard-of-care treatments (Cheng et al.; Mitchell et al.). The

following is a survey of the literature highlighting various nanoparticle-based approaches to treating amoebal infections:

Metronidazole conjugated magnetic nanoparticles loaded with amphotericin B

A study by Abdelnasir, S., Anwar, A, et.all introduces metronidazole-modified iron oxide nanoparticles loaded with amphotericin B as a therapeutic avenue for combating infections due to FLA. It focuses specifically on *A. Castellani*, a potential cause of GAE.

The use of metronidazole conjugated magnetic nanoparticles provides multifaceted advantages. These nanoparticles demonstrate excellent drug entrapment efficiency, and they ensured the effective delivery of both metronidazole and amphotericin B. Moreover, their biocompatibility and minimal hemolytic activity render them suitable drug carrier candidates. The *in vitro* experiments demonstrated potent synergistic effects alongside dose-dependent amoebicidal and cytotoxic activities against both *A. Castellani* trophozoites and cysts.

These drug-laden nanoparticles outperformed the control groups (individual drugs alone and nanoparticles alone). Additionally, the incorporation of iron oxide nanoparticles introduced magnetic properties, effectively enabling enhanced imaging sensitivity via MRI or magnetic resonance imaging for multifunctional theragnostic applications (Abdelnasir et al.).

Gold Nanoparticles

A study by Mungroo, et.all focuses on the therapeutic potential of curcumin, a bioactive small molecule compound derived from turmeric. Known for its diverse biological properties, curcumin was specifically explored for its amoebicidal effects. Curcumin was conjugated with gold nanoparticles as a treatment strategy against infections caused by two FLA *B. mandrillaris* and *N. fowleri*.

The synthesis of curcumin and gold nanoparticles was confirmed through dynamic light scattering (DLS), a technique used to measure the size distribution of particles in a solution. It indicated an average particle size of 53 nanometers. Curcumin exhibited substantial activity against both the amoeba with concentration-dependent effects showcasing an AC50 of 172 μM for *B. mandrillaris* and 74 μM for *N. fowleri*. An AC50 of 172 μM and 74 μM for *B. mandrillaris* and *N. fowleri* respectively represents the concentrations at which the inhibitory effect of amoebic growth is at 50% of its maximum. Once curcumin was conjugated with gold nanoparticles to enhance its amoebicidal activity, the resulting nanoparticle showed a remarkable increase of up to 78% in amoebicidal activity against *B. mandrillaris* and 69% against *N. fowleri*.

The observed enhancement in amoebicidal activity post-conjugation was attributed to the biological activity of gold nanoparticles as they have been reported to induce reactive oxygen species (ROS) formation; this, in turn, leads to apoptosis and influences many cellular processes. The intricate interplay between curcumin and gold nanoparticles therefore holds promise for improving the treatment of infections caused by these amoebae (Mungroo et al.).

Elastin-like polypeptide (ELP) nanoparticles

Elastin-like polypeptide (ELP) nanoparticles are biopolymers derived from the structural protein elastin. Elastin is a polymeric extracellular matrix protein integral to the complex macromolecular network which provides a structural framework outside of cells within tissues and organs alongside contributing to the stretchable nature of vertebrate tissues (Lima et al.). ELP nanoparticles are inspired by tropoelastin, the soluble precursor of elastin that contains hydrophobic motifs; a hydrophobic motif is a specific pattern of amino acids (e.g., valine, isoleucine) within a protein that exhibits hydrophobic properties (Despanie et al.). Specifically, the hydrophobic motifs in ELP nanoparticles are (Valine-Proline-Glycine-Xaa-Glycine)_n, where Xaa is any amino acid and n specifies the number of times this pentapeptide motif repeats in the polymer chain (Jolinde van Strien et al.). ELPs uniquely exhibit temperature-sensitive behavior by transitioning from a soluble state below their transition temperature (T_i) into a cloudy coacervate comprised of insoluble microparticles beyond their T_i; unlike most polymers, however, this process is reversible, which permits novel applications as a ‘smart’ biomaterial within the drug delivery field (Bidwell; Guo et al.).

1) Passive Targeting

The passive targeting capability of ELP nanoparticles could potentially prove invaluable in the treatment of PAM and GAE. Considering the predilection of these diseases for specific anatomical sites such as the CNS, ELPs can be engineered to carry drugs (e.g, amphotericin B, miltefosine) targeting both diseases.

Specifically, ELP nanoparticles loaded with amphotericin B might exploit the thermoresponsive behavior of ELPs to accumulate within the CNS, where GAE and PAM manifest. This passive delivery approach would nevertheless ensure that a high concentration of amphotericin B or miltefosine reaches the infection site, thereby enhancing therapeutic efficacy against the amoeba. In other words, this method could lower the minimum inhibitory concentrations required and therefore remove the obstacle of dose-limiting toxicities during administration of these drugs.

2) Controlled Release

ELP nanoparticles potentially open the possibility of leveraging sophisticated controlled release strategies suitable for novel treatments targeting both PAM and GAE. In the case of GAE, where prolonged therapeutic intervention is required, ELPs encapsulating miltefosine with thin cholesterol layers or hydrogels present a unique method for sustained drug release. This approach would ensure that encapsulated miltefosine is released gradually over time to maintain a consistent therapeutic concentration within the affected tissues, thereby improving efficacy against the amoebic pathogens causing GAE relative to existing treatment options.

3) Active targeting

Active targeting using ELP nanoparticles might prove pivotal in addressing the challenges of treating both GAE and PAM. By fusing an ELP with cell penetrating peptides

(CPPs), for instance, the drug delivery system gains the ability to efficiently traverse the blood brain barrier (BBB). For example, ELPs loaded with amphotericin B, when modified with cell penetrating peptides, can effectively cross the blood brain barrier thus ensuring that the drug reaches the CNS..

4) Multifunctional approaches

The multifunctional attributes of ELP nanoparticles, particularly in micellar structures with coronal modifications, could play a critical role in optimizing treatments against PAM and GAE. For instance, designing micellar ELPs capable of encapsulating miltefosine within a core while cell penetrating peptides are displayed on the corona could theoretically present a novel means of simultaneously achieving passive and active targeting to ensure that the drug is delivered directly to the site of amoebal infection. The synergy of these elements within a putative ELP nanoparticle design would contribute significantly to potentiating a multifaceted approach required for the effective treatment of both diseases.

Possible ELP nanoparticle design

ELPs have several tailorable properties such as coacervate formation at a transition temperature (T_t), an amphiphilic nature, and — most importantly — the ability to provide surface modifications (e.g., ligands, peptides, or other targeting molecules) that can be utilized for the treatment of FLA and their respective indications.

An ELP nanoparticle, for example, can be tailored to enhance its stability and efficacy. In the case of PAM, *N. fowleri* has a G-protein coupled receptor (GPCR) on its amoebal surface (Chen and Moseman). Specialized ligands targeting the GPCR on *N. fowleri* can be covalently bonded to ELP nanoparticles to enhance the ELP nanoparticles' targeted delivery.

Several methods can be used to develop or find ligands specific to *N. fowleri* GPCRs. Molecular docking work involves evaluating factors such as binding affinities, hydrophobicity, and hydrogen bonding; pharmacophore modeling, by contrast, works by identifying common structural features and chemical properties among ligands with similar pharmacophoric elements.

Additionally, a quantity structure-activity relationship (QSAR) analysis might be leveraged to identify potential ligands by correlating the chemical structure of a ligand under investigation with its biological activity. Bioinformatics and homology modeling can also be employed to identify potential ligands as well. Bioinformatics analysis, for example, could be performed on the known sequence of the GPCR while homology modeling would subsequently be used to predict that GPCR's three-dimensional structure; this would be helpful in isolating the specific structures targeted by any potential ligands identified.

Lastly, high-throughput screening (HTS) — involving the screening of hundreds to thousands of compounds — would be employed to identify those exhibiting a strong affinity for the GPCR being targeted.

Apart from specialized targeting ligands, there are several surface modifications that can be made to enhance the properties of the ELP nanoparticle including the addition of cationic

polymers covalently bonded to the ELP nanoparticle. Cationic polymers might potentially enhance the affinity of the ELP nanoparticles when targeting negatively charged amoebal cell surfaces based on physicochemical principles; such improved cellular adhesion and uptake might possibly enhance therapeutic efficiency.

Conclusion

ELP nanoparticles can be leveraged as an advanced yet adaptable drug delivery system for the design of novel therapeutics at the intersection of neurology and infectious diseases. ELP nanoparticles' biocompatibility, temperature responsive behavior, and capacity for finely tuned drug release kinetics therefore render them optimal candidates for the treatment of two high-mortality illnesses: PAM and GAE. According to Stahl and Olson, as surface water temperatures increase due to anthropogenic climate change, it is likely that *N. fowleri* will become a more significant threat to human health, and this might easily be extrapolated to other FLA. Therefore, this paper's emphasis on developing ELP-based nanoparticles as prophylactic measures against pathogenic FLA offer new avenues of investigation for future investigators.

Works Cited

- Grace, Eddie, et al. "Naegleria fowleri: Pathogenesis, Diagnosis, and Treatment Options." *Antimicrobial Agents and Chemotherapy*, vol. 59, no. 11, Nov. 2015, pp. 6677–81. <https://doi.org/10.1128/aac.01293-15>.
- Güemez, Andrea, and Elisa García. "Primary Amoebic Meningoencephalitis by Naegleria Fowleri: Pathogenesis and Treatments." *Biomolecules*, vol. 11, no. 9, Sept. 2021, p. 1320. <https://doi.org/10.3390/biom11091320>.
- Tillery, Logan, et al. "Naegleria Fowleri: Protein Structures to Facilitate Drug Discovery for the Deadly, Pathogenic Free-living Amoeba." *PloS One*, vol. 16, no. 3, Mar. 2021, p. e0241738. <https://doi.org/10.1371/journal.pone.0241738>.
- Herman, Emily K., et al. "Genomics and Transcriptomics Yields a System-level View of the Biology of the Pathogen Naegleria Fowleri." *BMC Biology*, vol. 19, no. 1, July 2021, <https://doi.org/10.1186/s12915-021-01078-1>.
- Visvesvara, Govinda S., et al. "Pathogenic and Opportunistic Free-living Amoebae: Acanthamoeba Spp., Balamuthia Mandrillaris, Naegleria Fowleri, and Sappinia Diploidea." *FEMS Immunology & Medical Microbiology*, vol. 50, no. 1, June 2007, pp. 1–26. <https://doi.org/10.1111/j.1574-695x.2007.00232.x>.
- "Naegleria." *PubMed*, 1 Jan. 2024, pubmed.ncbi.nlm.nih.gov/30571068.
- Rodríguez-Mera, Itzel Berenice, et al. "Role of Cathepsin B of Naegleria Fowleri During Primary Amebic Meningoencephalitis." *Parasitology Research*, vol. 121, no. 11, Sept. 2022, pp. 3287–303. <https://doi.org/10.1007/s00436-022-07660-y>.
- JOHN, D. T., Jr., et al. "Sucker-Like Structures on the Pathogenic Amoeba Naegleria Fowleri." *APPLIED AND ENVIRONMENTAL MICROBIOLOGY*, by American Society for Microbiology, vol. 47, no. 1, American Society for Microbiology, pp. 12–14. www.ncbi.nlm.nih.gov/pmc/articles/PMC239603/pdf/aem00158-0030.pdf.
- Chen, Ching-Wen, and E. Ashley Moseman. "Pro-inflammatory Cytokine Responses to Naegleria Fowleri Infection." *Frontiers in Tropical Diseases*, vol. 3, Jan. 2023, <https://doi.org/10.3389/fitd.2022.1082334>.
- Marciano-Cabral, Francine, and Guy A. Cabral. "The Immune Response to Naegleria Fowleri Amebae and Pathogenesis of Infection." *FEMS Immunology and Medical Microbiology*, vol. 51, no. 2, Nov. 2007, pp. 243–59. <https://doi.org/10.1111/j.1574-695x.2007.00332.x>.
- Jamerson, Melissa, et al. "Pathogenic Naegleria Fowleri and Non-pathogenic Naegleria Lovaniensis Exhibit Differential Adhesion to, and Invasion of, Extracellular Matrix Proteins." *Microbiology*, vol. 158, no. 3, Mar. 2012, pp. 791–803. <https://doi.org/10.1099/mic.0.055020-0>.
- Oh, Y. -h., et al. "Cytopathic Changes and Pro-inflammatory Cytokines Induced by Naegleria Fowleri Trophozoites in Rat Microglial Cells and Protective Effects of an anti-Nfa1 Antibody." *Parasite Immunology*, vol. 27, no. 12, Oct. 2005, pp. 453–59. <https://doi.org/10.1111/j.1365-3024.2005.00799.x>.

- Moseman, E. Ashley. “Battling Brain-eating Amoeba: Enigmas Surrounding Immunity to *Naegleria Fowleri*.” *PLOS Pathogens*, vol. 16, no. 4, Apr. 2020, p. e1008406. <https://doi.org/10.1371/journal.ppat.1008406>.
- Linden, S. K., et al. “Mucins in the Mucosal Barrier to Infection.” *Mucosal Immunology*, vol. 1, no. 3, May 2008, pp. 183–97. <https://doi.org/10.1038/mi.2008.5>.
- Cervantes-Sandoval, Isaac, et al. “*Naegleria Fowleri* Induces MUC5AC and Pro-inflammatory Cytokines in Human Epithelial Cells via ROS Production and EGFR Activation.” *Microbiology*, vol. 155, no. 11, Nov. 2009, pp. 3739–47. <https://doi.org/10.1099/mic.0.030635-0>.
- Vugia, Duc J., et al. “Notes From the Field: Fatal *Naegleria Fowleri* Meningoencephalitis After Swimming in Hot Spring Water — California, 2018.” *MMWR Morbidity and Mortality Weekly Report*, vol. 68, no. 36, Sept. 2019, pp. 793–94. <https://doi.org/10.15585/mmwr.mm6836a3>.
- Pana, Angela, et al. “Amebic Meningoencephalitis.” *StatPearls - NCBI Bookshelf*, 21 Jan. 2023, www.ncbi.nlm.nih.gov/books/NBK430754.
- Pugh, J. Jeffrey, and Rebecca A. Levy. “*Naegleria Fowleri*: Diagnosis, Pathophysiology of Brain Inflammation, and Antimicrobial Treatments.” *ACS Chemical Neuroscience*, vol. 7, no. 9, Aug. 2016, pp. 1178–79. <https://doi.org/10.1021/acscemneuro.6b00232>.
- Noor, Asif, and Charles V. Preuss. “Amphotericin B.” *StatPearls - NCBI Bookshelf*, 28 Feb. 2024, www.ncbi.nlm.nih.gov/books/NBK482327.
- Andrews, Jennifer M. “Determination of Minimum Inhibitory Concentrations.” *Journal of Antimicrobial Chemotherapy*, vol. 48, no. suppl_1, July 2001, pp. 5–16. https://doi.org/10.1093/jac/48.suppl_1.5.
- Rajendran, Kavitha, et al. “Brain-Eating Amoebae: Silver Nanoparticle Conjugation Enhanced Efficacy of Anti-Amoebic Drugs Against *Naegleria Fowleri*.” *ACS Chemical Neuroscience*, vol. 8, no. 12, Dec. 2017, pp. 2626–30. <https://doi.org/10.1021/acscemneuro.7b00430>.
- Goswick, Shannon M., and George M. Brenner. “Activities of Azithromycin and Amphotericin B Against *Naegleria Fowleri* in Vitro and in a Mouse Model of Primary Amebic Meningoencephalitis.” *Antimicrobial Agents and Chemotherapy*, vol. 47, no. 2, Feb. 2003, pp. 524–28. <https://doi.org/10.1128/aac.47.2.524-528.2003>.
- Zervos, Emmanuel E., et al. “Fluconazole Increases Bactericidal Activity of Neutrophils.” *Journal of Trauma and Acute Care Surgery*, vol. 41, no. 1, July 1996, pp. 10–14. <https://doi.org/10.1097/00005373-199607000-00003>.
- Ayudhya, Duangchit Panomvana Na, et al. “Effect of Rifampicin on the Pharmacokinetics of Fluconazole in Patients With AIDS.” *Clinical Pharmacokinetics*, vol. 43, no. 11, Jan. 2004, pp. 725–32. <https://doi.org/10.2165/00003088-200443110-00003>.
- Alli, Ammar, et al. “Miltefosine: A Miracle Drug for Meningoencephalitis Caused by Free-Living Amoebas.” *Curēus*, Mar. 2021, <https://doi.org/10.7759/cureus.13698>.
- Roy, Sharon L., et al. “Assessment of Blood–brain Barrier Penetration of Miltefosine Used to Treat a Fatal Case of Granulomatous Amebic Encephalitis Possibly Caused by an Unusual *Balamuthia mandrillaris* Strain.” *Parasitology Research*, vol. 114, no. 12, Sept. 2015, pp. 4431–39. <https://doi.org/10.1007/s00436-015-4684-8>.

- Siripurapu, Govinda, et al. "Successful Management of post-COVID-19 Acanthamoebic Encephalitis." *International Journal of Infectious Diseases*, vol. 110, Sept. 2021, pp. 226–28. <https://doi.org/10.1016/j.ijid.2021.07.046>.
- Parija, SubhashChandra, et al. "Management of Granulomatous Amebic Encephalitis: Laboratory Diagnosis and Treatment." *Tropical Parasitology*, vol. 5, no. 1, Jan. 2015, p. 23. <https://doi.org/10.4103/2229-5070.149889>.
- Keane, Niamh A., et al. "A Surviving Case of Acanthamoeba Granulomatous Amebic Encephalitis in a Hematopoietic Stem Cell Transplant Recipient." *the American Journal of Case Reports*, vol. 21, June 2020, <https://doi.org/10.12659/ajcr.923219>.
- Da Rocha-Azevedo, Bruno, et al. "Diagnosis of Infections Caused by Pathogenic Free-Living Amoebae." *Interdisciplinary Perspectives on Infectious Diseases*, vol. 2009, Jan. 2009, pp. 1–14. <https://doi.org/10.1155/2009/251406>.
- Kalra, Sonali K., et al. "Acanthamoeba and Its Pathogenic Role in Granulomatous Amebic Encephalitis." *Experimental Parasitology*, vol. 208, Jan. 2020, p. 107788. <https://doi.org/10.1016/j.exppara.2019.107788>.
- Wang, Yuehua, et al. "Biological Characteristics and Pathogenicity of Acanthamoeba." *Frontiers in Microbiology*, vol. 14, Apr. 2023, <https://doi.org/10.3389/fmicb.2023.1147077>.
- Siddiqui, Ruqaiyyah, and Naveed Ahmed Khan. "Biology and Pathogenesis of Acanthamoeba." *Parasites & Vectors*, vol. 5, no. 1, Jan. 2012, <https://doi.org/10.1186/1756-3305-5-6>.
- Duggal, Shalini Dewan, et al. "Role of Acanthamoeba in Granulomatous Encephalitis: A Review." 22 Jan. 2018, www.scitechnol.com/peer-review/role-of-acanthamoeba-in-granulomatous-encephalitis-a-review-JszK.php?article_id=7095.
- Cursons, R. T., et al. "Immunity to Pathogenic Free-living Amoebae: Role of Humoral Antibody." *Infection and Immunity*, vol. 29, no. 2, Aug. 1980, pp. 401–07. <https://doi.org/10.1128/iai.29.2.401-407.1980>.
- Kot, Karolina, et al. "Immunopathogenicity of Acanthamoeba Spp. In the Brain and Lungs." *International Journal of Molecular Sciences*, vol. 22, no. 3, Jan. 2021, p. 1261. <https://doi.org/10.3390/ijms22031261>.
- Visvesvara, Govinda S., et al. "Pathogenic and Opportunistic Free-living Amoebae: Acanthamoebaspp., Balamuthia Mandrillaris, Naegleria Fowleri, and Sappinia Diploidea." *FEMS Immunology & Medical Microbiology*, vol. 50, no. 1, June 2007, pp. 1–26. <https://doi.org/10.1111/j.1574-695x.2007.00232.x>.
- Chalmers, Rachel M. "Acanthamoeba." Elsevier eBooks, 2014, pp. 263–76. <https://doi.org/10.1016/b978-0-12-415846-7.00014-7>.
- Ellison, David, et al. "Parasitic Infections." Elsevier eBooks, 2013, pp. 403–24. <https://doi.org/10.1016/b978-0-7234-3515-0.00018-0>.
- Bhosale, Namrata K., and Subhash Chandra Parija. "Balamuthia Mandrillaris: An Opportunistic, Free-living Ameba - an Updated Review." *PubMed*, vol. 11, no. 2, Nov. 2021, pp. 78–88. https://doi.org/10.4103/tp.tp_36_21.

- Stothard, D. R., et al. “The Evolutionary History of the Genus *Acanthamoeba* and the Identification of Eight New 18S rRNA Gene Sequence Types.” *Journal of Eukaryotic Microbiology*, vol. 45, no. 1, Jan. 1998, pp. 45–54. <https://doi.org/10.1111/j.1550-7408.1998.tb05068.x>.
- Phan, Isabelle Q., et al. “The Transcriptome of *Balamuthia Mandrillaris* Trophozoites for Structure-guided Drug Design.” *Scientific Reports*, vol. 11, no. 1, Nov. 2021, <https://doi.org/10.1038/s41598-021-99903-8>.
- Cope, Jennifer R., et al. “The Epidemiology and Clinical Features of *Balamuthia Mandrillaris* Disease in the United States, 1974–2016.” *Clinical Infectious Diseases*, vol. 68, no. 11, Sept. 2018, pp. 1815–22. <https://doi.org/10.1093/cid/ciy813>.
- Lee, Daniel C., et al. “*Acanthamoeba* Spp. And *Balamuthia Mandrillaris* Leading to Fatal Granulomatous Amebic Encephalitis.” *Forensic Science, Medicine and Pathology*, vol. 16, no. 1, Nov. 2019, pp. 171–76. <https://doi.org/10.1007/s12024-019-00202-6>.
- Da Rocha-Azevedo, Bruno, et al. “Diagnosis of Infections Caused by Pathogenic Free-Living Amoebae.” *Interdisciplinary Perspectives on Infectious Diseases*, vol. 2009, Jan. 2009, pp. 1–14. <https://doi.org/10.1155/2009/251406>.
- CDC - DPDx - Free Living Amebic Infections. www.cdc.gov/dpdx/freelivingamebic/index.html.
- Qvarnstrom, Yvonne, et al. “Molecular Confirmation of *Sappinia pedata* as a Causative Agent of Amebic Encephalitis.” *The Journal of Infectious Diseases*, vol. 199, no. 8, Apr. 2009, pp. 1139–42. <https://doi.org/10.1086/597473>.
- Taravaud, Alexandre, et al. “Drugs Used for the Treatment of Cerebral and Disseminated Infections Caused by Free-living Amoebae.” *Clinical and Translational Science*, vol. 14, no. 3, Mar. 2021, pp. 791–805. <https://doi.org/10.1111/cts.12955>.
- Taravaud, Alexandre, Philippe M. Loiseau, et al. “In Vitro Evaluation of Antimicrobial Agents on *Acanthamoeba* Sp. And Evidence of a Natural Resilience to Amphotericin B.” *International Journal for Parasitology, Drugs and Drug Resistance*, vol. 7, no. 3, Dec. 2017, pp. 328–36. <https://doi.org/10.1016/j.ijpddr.2017.09.002>.
- Elsheikha, Hany M., et al. “Drug Discovery Against *Acanthamoeba* Infections: Present Knowledge and Unmet Needs.” *Pathogens*, vol. 9, no. 5, May 2020, p. 405. <https://doi.org/10.3390/pathogens9050405>.
- Chen, Xiang-Ting, et al. “Pathogenic Free-living Amebic Encephalitis From 48 Cases in China: A Systematic Review.” *Frontiers in Neurology*, vol. 14, Feb. 2023, <https://doi.org/10.3389/fneur.2023.1100785>.
- Spottiswoode, Natasha, et al. “Successful Treatment of *Balamuthia Mandrillaris* Granulomatous Amebic Encephalitis With Nitroxoline.” *Emerging Infectious Diseases*, vol. 29, no. 1, Jan. 2023, pp. 197–201. <https://doi.org/10.3201/eid2901.221531>.
- Joseph, Tomy, et al. “Nanoparticles: Taking a Unique Position in Medicine.” *Nanomaterials*, vol. 13, no. 3, Jan. 2023, p. 574. <https://doi.org/10.3390/nano13030574>.
- Elumalai, Karthikeyan, et al. “Review of the Efficacy of Nanoparticle-based Drug Delivery Systems for Cancer Treatment.” *Biomedical Technology*, vol. 5, Mar. 2024, pp. 109–22. <https://doi.org/10.1016/j.bmt.2023.09.001>.

- Sultana, Afreen, et al. "Nano-based Drug Delivery Systems: Conventional Drug Delivery Routes, Recent Developments and Future Prospects." *Medicine in Drug Discovery*, vol. 15, Sept. 2022, p. 100134. <https://doi.org/10.1016/j.medidd.2022.100134>.
- Elumalai, K., Srinivasan, S., & Shanmugam, A. (2024). Review of the efficacy of nanoparticle-based drug delivery systems for cancer treatment. *Biomedical Technology*, 5, 109–122. <https://doi.org/10.1016/j.bmt.2023.09.001>
- Zottel, Alja, et al. "Nanotechnology Meets Oncology: Nanomaterials in Brain Cancer Research, Diagnosis and Therapy." *Materials*, vol. 12, no. 10, May 2019, p. 1588. <https://doi.org/10.3390/ma12101588>.
- Tobin, Ellis, and Sara Brenner. "Nanotechnology Fundamentals Applied to Clinical Infectious Diseases and Public Health." *Open Forum Infectious Diseases*, vol. 8, no. 12, Nov. 2021, <https://doi.org/10.1093/ofid/ofab583>.
- Siddiqui, Ruqaiyyah, et al. "Zinc Oxide Nanoconjugates Against Brain-Eating Amoebae." *Antibiotics*, vol. 11, no. 10, Sept. 2022, p. 1281. <https://doi.org/10.3390/antibiotics11101281>.
- Fan, XueMei, et al. "Encephalomyelomeningitis Caused by Balamuthia Mandrillaris: A Case Report and Literature Review." *Infection and Drug Resistance*, vol. Volume 16, Feb. 2023, pp. 727–33. <https://doi.org/10.2147/idr.s400692>.
- Cheng, Xiaoxiao, et al. "Advances in Nanomaterial-based Targeted Drug Delivery Systems." *Frontiers in Bioengineering and Biotechnology*, vol. 11, Apr. 2023, <https://doi.org/10.3389/fbioe.2023.1177151>.
- Mitchell, Michael J., et al. "Engineering Precision Nanoparticles for Drug Delivery." *Nature Reviews. Drug Discover/Nature Reviews. Drug Discovery*, vol. 20, no. 2, Dec. 2020, pp. 101–24. <https://doi.org/10.1038/s41573-020-0090-8>.
- Abdelnasir, Sumayah, et al. "Metronidazole Conjugated Magnetic Nanoparticles Loaded With Amphotericin B Exhibited Potent Effects Against Pathogenic Acanthamoeba Castellanii Belonging to the T4 Genotype." *AMB Express*, vol. 10, no. 1, July 2020, <https://doi.org/10.1186/s13568-020-01061-z>.
- Mungroo, Mohammad Ridwane, et al. "Gold-Conjugated Curcumin as a Novel Therapeutic Agent Against Brain-Eating Amoebae." *ACS Omega*, vol. 5, no. 21, May 2020, pp. 12467–75. <https://doi.org/10.1021/acsomega.0c01305>.
- Lima, Leticia Ferreira, et al. "Elastin-like Polypeptides in Development of Nanomaterials for Application in the Medical Field." *Frontiers in Nanotechnology*, vol. 4, Apr. 2022, <https://doi.org/10.3389/fnano.2022.874790>.
- Despanie, Jordan, et al. "Elastin-like Polypeptides: Therapeutic Applications for an Emerging Class of Nanomedicines." *Journal of Controlled Release*, vol. 240, Oct. 2016, pp. 93–108, <https://doi.org/10.1016/j.jconrel.2015.11.010>.
- Jolinde van Strien, et al. "Elastin-like Polypeptide-Based Micelles as a Promising Platform in Nanomedicine." *Journal of Controlled Release*, vol. 353, Elsevier BV, Jan. 2023, pp. 713–26, <https://doi.org/10.1016/j.jconrel.2022.12.033>. Accessed 28 Apr. 2023.

- Bidwell, Gene L. “Novel Protein Therapeutics Created Using the Elastin-Like Polypeptide Platform.” *Physiology*, vol. 36, no. 6, Nov. 2021, pp. 367–81.
<https://doi.org/10.1152/physiol.00026.2021>.
- Guo, Yingshu, et al. “The Construction of Elastin-like Polypeptides and Their Applications in Drug Delivery System and Tissue Repair.” *Journal of Nanobiotechnology*, vol. 21, no. 1, Nov. 2023, <https://doi.org/10.1186/s12951-023-02184-8>.
- Chen, Ching-Wen, and E. Ashley Moseman. “Pro-inflammatory Cytokine Responses to *Naegleria Fowleri* Infection.” *Frontiers in Tropical Diseases*, vol. 3, Jan. 2023, <https://doi.org/10.3389/fitd.2022.1082334>.

Empirically-Derived Applications of Music Therapy Across Lifespans By Chloe Yoon

Abstract

Music therapy in one form or another has been a part of many cultures for thousands of years, changing many times over the centuries. This paper explores the practical applications of music therapy techniques for children, adolescents, and the elderly, highlighting the methods used and therapeutic benefits. I will start by discussing the history of music therapy and its applications throughout different cultures. Then I will discuss different types of treatments that can be applied to each age group along with their rationale. Finally, I will discuss future applications of music therapy and where additional studies are needed.

Introduction

For eons, all humans across various cultures and backgrounds have used music as a common form of therapy and expression. One of the first groups known to heal with sound were Australians, who used an instrument called the yidaki, also known as the didgeridoo. The yidaki was believed to have assisted in healing broken bones and illnesses of every kind for at least 40,000 years (Nicholls, 2017). Traditional Chinese medicine, which is still actively used today, initially involved music rather than acupuncture or traditional medicine (Marshall, 2020). The Chinese character for 'medicine' combines the characters for 'music' and 'grass,' suggesting that they believed in the important role of music in regulating life harmony and improving health. The roots of music therapy can also be found in ancient Hindu mythology, Sanskrit, and local folk traditions (Gangopadhyay & Prasad, 2019). More recently, music therapy became popular in the United States in the aftermath of World War I and II (Spencer, 2013). Musicians would travel to hospitals to play music for soldiers suffering both physically and mentally during and after the two world wars. The military currently utilizes music therapy in the recovery of military service members in Army hospitals (LeDoux, 2016).

Music therapy incorporates a wide range of methods. But generally, they may be divided into two types: active and passive music therapy. Active music therapy occurs when a patient consciously takes part in creating music through improvisation, songwriting, or musical instrument playing. Improvisation music therapy involves patients' expression of their emotions and thoughts without words through free musical improvisation techniques. Such methods may be particularly effective for patients with communication difficulties (Bruscia, 1998). Passive music therapy occurs when the patient listens to music. Music in such therapy is specially chosen for the patient and would be suitable for the client's therapeutic goals. One of the most popular techniques for passive music therapy is guided imagery in music, which helps clients listen to music and imagine peaceful images at the same time (Bonny & Savary, 1973). Therapists commonly apply other psychosocial methods, such as cognitive-behavioral techniques, which are supposed to intensify the positive effect of music therapy.

Music therapy is effective in treating depression, anxiety, and even post-traumatic stress in the mental health profession. It offers non-verbal expression to process traumatic experiences.

Listening to and actively participating in music can relieve symptoms of depression and anxiety by promoting relaxation and emotional release (Erkkilä *et al.*, 2011). Music therapy has been found to improve the functioning of motor, communication, and cognitive responses in neurological disorders such as Parkinson's, stroke, and dementia. For example, rhythmic sound stimulation is useful for the restoration of walking and coordination in people affected by Parkinson's disease (Thaut, 2015). There have also been great benefits associated with music therapy for children diagnosed on the autism spectrum in terms of developing social skills, communication, and sensory regulation improvements. Music therapy is employed on a large scale, and its benefits documented in the autism spectrum are due to the inherent quality of music in that it has structured yet flexible principles, thus making it an optimal vehicle for such children (Geretsegger *et al.*, 2014).

Music therapy aids in pain management and physical rehabilitation by alleviating pain, reducing stress, and enhancing mood. Surgical and cancer patients have described music therapy as being used to lower their pain intensity and enhance their general health-related quality of life (Cepeda *et al.*, 2006). In end-of-life care, music therapy comforts, reduces anxiety, and supports emotional and spiritual wellness. It helps patients and their families to cope with sorrow and loss; therefore, a sense of peace is attained during furious moments (Hilliard, 2003).

Although considerable evidence showing the effectiveness of music therapy exists, skepticism about the validity of music as a mode of treatment persists, with concerns about its scientific basis and practical applications. Music is primarily used today for purposes of entertainment; hence, it is difficult to think that it can bring about such radical changes in a patient's medical condition. The uncommon use of music therapy may also be due to the relatively recent establishment of music therapy as an educational program. The first academic program in music therapy was established in 1944 at Michigan State University, and music therapy as a profession first became accessible through the American Music Therapy Association (AMTA) in 1988, which strives to advocate for and educate music therapists (L'Etoile, 2000).

This paper addresses the development and application of active and passive music therapy across different age groups. Specifically, I will discuss how music therapy is applied to children, adults, and geriatric populations, as well as the science that supports these applications. I will review neuropsychological and medical research that supports these practices. Finally, I will discuss the potential future of music therapy, including increasing accessibility, occupations, and the current state of research.

The Role of Music in Childhood Development

From the moment they enter the world, with ears attuned to the melodies of life, children embark on a musical journey that shapes their very existence until the cusp of adolescence. Although active music therapy intervention is limited for young children, music still plays a significant broader role in the development of their brains. The fantastical tunes that weave their way around nursery rhymes and the jaunty beats that support playground games are just a couple of simple ways music is intertwined with the very fabric of childhood. In the first two years of

life alone, the brain sees quite a lot of change (Lally & Valentine-French, 2019). Branch-like structures called dendrites multiply, collecting information from other brain cells. Myelin, a fatty coating around nerve fibers, develops, improving coordination and thinking skills by strengthening neural connections. While the brain grows quickly, different areas mature at different rates. Motor areas develop before sensory areas, and the prefrontal cortex, responsible for emotions and planning, is one of the last to mature. As the prefrontal cortex develops, children gain better emotional control, planning abilities, and judgment.

Rhythmic abilities obtained through listening to music are also related to cognitive motor processing in children (Frischen *et al.*, 2022). From a very young age, children try to recognize patterns within their physical and social surroundings, using them as a foundation for making predictions. Predictive processes are essential for both rhythm and cognitive functions, serving as bridges between them. The underlying neurological mechanisms involved in rhythm processing and cognitive development include the auditory cortex, prefrontal cortex, and cerebellum. These regions play critical roles in processing rhythmic information and supporting cognitive functions. Rhythm processing skills contribute to the development of cognitive abilities such as attention, memory, language, and executive function in children. The ability to perceive and predict rhythmic patterns serves as a foundation for the development of higher-level cognitive skills.

Frequent music exposure can enhance early language learning in kids since the brain's music and language domains partially overlap; that is, the same brain regions are processed by music as well as speech (Peretz *et al.*, 2015). These areas of the brain include the auditory cortex, Broca's area, Wernicke's area, and the motor area. The areas of the brain in the domains that deal with music and those that handle speech processes include the auditory cortex and the emotion processing areas in the frontal and temporal areas.

Music has also been shown to enhance literacy skills in young children. For example, Gordon and colleagues conducted a study in 2015 at Nashville and Stanford University in which children were exposed to different components of music such as rhythm, clapping, singing, musical notation, and instruments, suggesting that music training improves phonological awareness, particularly in rhyming skills. This aligns with previous studies showing that musicians tend to have better phonological awareness skills. However, the effectiveness of music training on rhyming skills depends on the number of training hours, with at least 40 hours needed for improvement. Gordon and colleagues' work also explores other phonological outcomes and reading fluency but finds inconclusive results, indicating the need for further investigation. Despite limitations in study designs and consistency of outcomes, the findings suggest potential benefits of music training in improving reading-related skills, especially phonological awareness. However, more research is needed to understand the underlying mechanisms and to establish a solid link between music education and literacy skills.

Music Therapy for Adolescents

Adolescence is the time between childhood and adulthood when many teens experiment with identity and face strong emotions. Since it is a developing phase that is incomplete in itself,

it is found that they are vulnerable and prone to risks that can turn out to be dangerous for them. Music or sounds also play a significant role in the lives of adolescents because they are central to the formation of their identity, emotions, and social life. When brains mature during this period of teenagers, early-maturing primary sensory and motor regions become vast bags of gray matter and lead brain development compared to the later-maturing prefrontal cortex that controls decision-making and higher-level operations like abstract reasoning (Konrad *et al.*, 2013). The maturation of the limbic system of the brain responsible for emotions and the reward system takes place before the maturation of the prefrontal cortex, which controls decision-making and higher-level operations.

It is also at this age that teenagers are experiencing surges of hormones that can result in changes in their moods and complex emotions (Peper & Dahl, 2015). This results in a high turnout of teenagers contracting depression and other mental disorders at this age. While adolescence is a time of great confusion and conflict for teenagers, it is also an overwhelming period for expressing feelings and emotions. Thus, music is a strong medium through which teenagers can express their feelings, connect with their age-mates, and attempt to make sense of their evolving identities. From playlists on their smartphones to concert experiences shared with friends, music is present in nearly all aspects of teenagers' lives, playing a considerable role in their mood regulation, self-perception, and social integration.

A study conducted in 2004 by the International Foundation for Music Research (IFMR) suggests that regular music lessons correlate to refinement in general education and even improvement in intelligence (Schellenberg, 2004). 144 students were randomly assigned to four groups: two received music (keyboard or voice) lessons, and two were in control groups (drama lessons or no lessons at all). Lessons were held at the Royal Conservatory of Music, with follow-up testing a year later. All groups showed increased IQ, likely influenced by starting school. However, music groups had greater increases in IQ than control groups. Teenagers who took music lessons showed larger improvements compared to controls in academic skills, as measured by the Kaufman Test of Educational Achievement (KTEA).

Emotion-regulating Improvisational Music Therapy (ERIMT) has been shown to reduce depressive symptoms in young adult students (Aalbers *et al.*, 2020). In ERIMT sessions, individuals make music under the guidance of a therapist. Through playing different musical instruments and using their voice, participants have a vehicle to express emotions, thoughts, and experiences that, at times, cannot be expressed in simple verbal language. Through the music, the participant explores and processes feelings in the nurturing and safe environment the therapist provides. As the session goes on, the therapist might give prompts or other interventions, urging the participant to look deeper into their feelings or helping the group communicate or connect better. In this musical conversation, participants experiment, improvise, and explore different emotions, leading to release, insight, and much more perception into one's social growth and well-being. The therapist mirrors the emotions back and guides the music-making process. The intent is to create a safe, exploratory environment for emotional growth. ERIMT can be helpful

for the treatment of a wide variety of mental health conditions, including anxiety, depression, trauma, and addiction, in individual and group settings.

Depression is a major mental health issue, especially among young adults, so a study in 2020 was conducted using a multiple-case study approach with various methods, analyzing eleven female students. Results showed that nine out of eleven students experienced decreased depressive symptoms after ERIMT, lasting up to four weeks. All students reported improved emotional regulation abilities, and these results remained consistent after four weeks. This study suggests that ERIMT could help reduce depressive symptoms and improve emotional well-being among young adult students.

Music therapy has also been shown to be effective in treating students enrolled in special education programs (Rickson & McFerran, 2024). In this context, music therapy diverges from music teaching by prioritizing the development of non-musical skills and utilizing music for emotional and personal growth rather than technical improvement. Assessments by music therapists track changes in motor skills, communication, socialization, cognition, and musical responses. Various interventions are used, such as singing, songwriting, improvisation, music listening, computer-based activities, lyric analysis, and guided movement to music (American Music Therapy Association, 2021).

A meta-analysis of eleven studies found that music therapy provides significant benefits for children with behavioral or developmental disorders (Gold *et al.*, 2004). Music therapy motivates and is theorized to help offset short attention spans often present in children with developmental delays and disabilities, enabling them to reach their potential. Music therapy is also an intervention frequently used with children diagnosed with autism spectrum disorder (ASD) and has been supported by several quantitative studies. Music therapy was revealed to be much more effective than placebo treatment or standard care alone in improving the communicative abilities of children with ASD.

It was also noted that varied music therapy approaches yield the most success in lowering aggression and hostility rates, while more fixed behavioral interventions show inconsistent results (Montello & Coons, 2003). Montello and Coons also observed enhanced attention, motivation, and reduced hostility among students with emotional, learning, and behavioral disorders post-music therapy. However, some students exhibited temporary behavioral regressions after sessions, particularly those with attention-deficit/hyperactivity disorder (ADHD), suggesting the need for more structured therapy. Rickson (2006) explored the impact of instructional and improvisational music therapy on adolescent boys with ADHD, noting a significant reduction in ADHD symptoms in the classroom.

Music Therapy for the Geriatric Population

The geriatric population comprises adults aged 75 years and older who are afflicted with various socio-economic and biological problems stemming from physical, cognitive, and emotional well-being that are oftentimes exacerbated by age-related ailments or chronic

conditions. Of the available therapeutic interventions, music therapy has emerged as a very promising approach to meeting the multidimensional needs of older adults.

For instance, Parkinson's disease (PD) is usually diagnosed in geriatric patients, and a musical strategy that could be acted upon is called rhythmic auditory stimulation. Rhythmic auditory stimulation (RAS) is the intervention providing auditory cues to enhance gait and balance in individuals suffering from Parkinson's disease or other movement disorders through facilitation of temporal processing in the brain utilizing the application of a fixed beat, which can either be of a metronome or musical nature (Pando-Naude *et al.*, 2021).

RAS is designed to complement, not replace, conventional drug treatments for PD. Although there is no standardized method for delivering RAS, it typically includes elements such as rhythmicity, motor synchronization, customization, and the use of gait kinematics as outcome measures. The intervention involves a healthcare professional providing auditory cues with beats per minute (bpm) adjusted to the individual's baseline gait, measured in steps per minute (spm). Features of RAS, such as the beat tempo or rhythm complexity, can be modified from baseline, often resulting in short-term improvements in gait kinematics, including speed, cadence, and step length (Pando-Naude *et al.*, 2024). This therapy is generally conducted in a supervised setting, like a physical therapy clinic or hospital, and generally lasts 30 to 60 minutes, two to three times per week, for several weeks or months. Studies have demonstrated its effectiveness in decreasing both the frequency of freezing episodes and the occurrence of falls in patients.

Music therapy also offers promising avenues for addressing various symptoms experienced by stroke patients (Xu *et al.*, 2022). For instance, engaging with lyrics and singing has been shown to enhance language comprehension and speech fluency in patients. In addition, dysphagia, or swallowing difficulties, is experienced by stroke patients, especially elderly ones. Dysphagia must be treated in a timely manner since it may cause psychological problems, such as fear of food and drinks, anxieties, and even depression disorders. It has been noted that singing in music therapy is one of the most substantial oral motor control treatments for stroke patients because the specific group showed significant changes in laryngeal elevation, laryngeal vestibule closure, respiration, and swallow function. Furthermore, music therapy aids in mitigating motor dysfunction associated with stroke by promoting functional and structural flexibility within the nervous system post-injury. Activities like playing musical instruments and following rhythmic cues can also enhance the frequency and fluency of finger movements.

Dementia stands as the leading cause of disability among the elderly worldwide (Apurv Shirsat *et al.*, 2023). This cognitive impairment primarily manifests as memory loss, language difficulties, and executive function challenges, significantly impacting an individual's behavior and emotions and affecting their overall well-being and quality of life. One common approach to addressing the mental and behavioral symptoms of dementia patients is through personalized music playlists in healthcare settings. Music is often integrated into daily routines as a supplementary therapy alongside pharmacological treatments, aligning with national dementia care strategies across various countries. Tailoring music selections to resonate with each

individual's preferences and personal significance enhances the therapeutic effects of music for dementia patients.

The findings of Shirlyne Vianna Moreira *et al.* (2018) revealed that, compared to passive activities like listening to music or watching television, short sessions of group singing led by a music therapist had a more significant positive impact on the quality of life among individuals with dementia. Older adults participating in interactive music interventions, which involved personalized music selection and engaging activities such as clapping and dancing, showed greater improvements in psychological and behavioral well-being. Scott Kaiser from the Pacific Neuroscience Institute suggests that recalling music or engaging in singing may activate brain regions unaffected by neurodegenerative diseases like Alzheimer's and dementia (Garrison, 2021). Consequently, the perception of music, associated emotions, and memory may be preserved beyond the point at which other memory and cognitive functions decline.

In a recent education systematic review, Lam *et al.* (2020) conducted a review of music therapy effects among people with dementia. The critical review evaluated 82 studies investigating the benefits of music therapy for cognitive functions, behavioral and psychological symptoms, and overall quality of life in patients with this clinical condition. It was inferred from the results that while some studies showed possible improvements in memory, cognition, daily functioning, and quality of life in patients suffering from dementia, the area in which music therapy showed the most positive impact was the reduction of behavioral and psychological symptoms such as depression, anxiety, and agitation—very common symptoms for patients with Alzheimer's and dementia. A more recent review in 2021 went a step further to establish the effectiveness of music therapy for various conditions, including dementia. Results from the review indicated that music therapy reduced mood-related symptoms, including depression and anxiety, but it also improved memory in patients with mild Alzheimer's disease.

Conclusion

Music plays a critical role in the cognitive, emotional, and social development of human beings from early childhood through late adulthood. The influence of music on the brain and its effect on changing behavior begins from early childhood to the stage of being an adult. In younger children, music plays a crucial role in shaping neural pathways, accelerating language acquisition, and honing cognitive motor processing abilities. Nursery rhymes, rhythmic activities, and musical play create rich, engaging environments that support the rapid growth of the brain, particularly in areas associated with coordination, sensory processing, and emotion regulation.

As children grow, their ability to recognize and predict rhythmic patterns becomes intertwined with broader cognitive skills. These skills underpin essential functions such as attention, memory, language development, and executive function. Studies have shown that regular exposure to music can significantly improve phonological awareness and other literacy skills, highlighting the overlapping brain regions involved in processing music and language. However, while the benefits of music training are evident, further research is needed to fully

understand the mechanisms behind these improvements and to establish stronger links between music education and academic performance.

In adolescence, music still plays an important role in emotional expression, the formation of identity, and social connections. The brain changes occurring at this time, together with the hormonal changes, mean that teenagers are well placed to pick up on, respond to, and interact with the emotional and social content of music. For this reason, music therapy has been seen to help adolescents manage some aspects of mental health, such as depression and anxiety. Techniques such as ERIMT offer safe spaces in which to explore and grow emotionally to improve regulation and general well-being.

In adulthood, especially in the elderly, music therapy has therapeutic effects on different physical, cognitive, and emotional states. In the elderly population, specifically where neurodegenerative conditions such as Parkinson's disease are concerned, RAS can improve motor functions and decrease symptoms. Stroke patients benefit in terms of efficient recovery of speech and motor functions and from a reduction in the psychological impact of dysphagia. In patients with dementia, customized music often significantly increases their quality of life by reducing behavioral and psychological symptoms, indicating that music has long-term effects on the process of memory as well as on emotions. Moreover, there are many emerging directions for adults.

Emerging Applications for Veterans

Today, music therapy is commonly used for patients with post-traumatic stress disorder (PTSD), particularly among veterans, often resulting from their involvement in wartime experiences. Treating this disorder is essential because individuals with PTSD are sensitive to dopamine reception, increasing their vulnerability to substance use disorders (Landis-Shack *et al.*, 2022). Music provides a healthier avenue for dopamine release, potentially facilitating the development of new neural pathways beneficial for PTSD recovery. Passive music therapy enables listeners to momentarily disconnect from their surroundings, offering a soothing effect on PTSD sufferers. It is commonly complemented with cognitive behavior therapy, assisting patients in identifying and challenging negative thoughts while acquiring practical self-help strategies.

Engaging in active group music therapy can foster a sense of social interaction by collectively tapping into the connection between rhythm and song lyrics. For example, a collaborative songwriting program conducted by *SongwritingWith:Soldiers* in 2018 with veterans suffering from PTSD yielded several positive outcomes, including a 25% decrease in depressive symptoms, a 77% rise in feelings of optimism and hope, and an 83% increase in engagement with creative activities. Moreover, connecting deeply with the lyrics of a pertinent song can help patients reinterpret trauma by expressing emotions through music.

Future Directions

Overall, the extensive body of research underscores the significant role of music therapy in enhancing various aspects of human development and health. From the tender beginnings of infancy to the challenges of old age, music acts as a potent catalyst for enhancing cognitive abilities, nurturing emotional well-being, and fostering social connections. Future research should continue to explore the mechanisms through which music exerts its effects, optimize therapeutic interventions, and expand the understanding of music's potential to improve health outcomes.

For example, patients experiencing memory loss often exhibit an intriguing phenomenon wherein they retain the ability to recall their favorite songs from their youth despite losing memories of loved ones and fundamental life skills (Moreira *et al.*, 2023). Why do individuals with dementia retain the ability to remember their favorite songs? This retention may be attributed to the unique way musical memory is stored and accessed in the brain. Musical memory is the form of memory entailed in musical experiences such as playing music or even listening to it (Jancke, 2008). The auditory cortex, motor cortex, and premotor cortex are different centers involved in forming musical memory. As these associated areas of the brain appear to be less affected by cognitive decline, this could elucidate why individuals with dementia are more apt to recall musical memories compared to other types of memories.

Music therapy shows great promise for the future, with technological advancements and ongoing research highlighting its numerous benefits. Wearable devices and sensors will offer real-time feedback for therapists to monitor patients' physical responses to music and adjust therapy accordingly. Soon, personalized programs of music therapy that are tailor-made through data will be used to build the best intervention. This will make treatment highly believable and tailor-made to emotional and psychological needs.

Additionally, the use of virtual reality (VR) in music therapy could transform the field significantly. VR can provide immersive experiences that enhance therapeutic engagement, especially for patients dealing with anxiety, PTSD, or other mental health issues (Boeldt *et al.*, 2019). VR technology is innovative and beneficial because it increases the availability of exposure therapy as well as its effectiveness as it permits individualized and controlled treatment that is easier for therapists to apply. Virtual reality (VR) technology is making significant strides in the field of music therapy, with programs like Musical Neglect Training (MNT) paving the way for innovative therapeutic approaches. A recent article in *Psychology Today* (Quirk, 2024) discusses how music therapy and virtual reality can aid post-stroke recovery, particularly for patients with visuospatial neglect (VSN). Visuospatial neglect (VSN) affects a person's ability to perceive objects on one side, often resulting in mobility and safety challenges. MNT uses instruments to engage neglected areas, while VR provides immersive experiences that boost brain activity and patient involvement. Results suggest that combining these approaches can significantly aid recovery in stroke patients. Merging these treatments has shown promising preliminary outcomes toward engagement and rehabilitation.

A pivotal aspect shaping the future of music therapy lies in the heightened emphasis on interdisciplinary collaboration, marking a significant evolution in the field. I foresee music

therapists teaming up with other professional groups—perhaps psychologists, neurologists, or occupational therapists—for the formulation of all-inclusive treatment plans. Collaboration in this manner helps to better manage the all-round needs of the patients. Cognitive-behavioral techniques and music therapy can be combined to enhance motor rehabilitation exercises. Such interdisciplinary synergy extends the boundaries of music therapy use and enriches its application, making this modality a very flexible therapy tool in the sphere of medical practice.

In conclusion, music therapy is expected to gain wider acceptance and integration into standard healthcare practices. With increasing research backing its efficacy across a spectrum of conditions, from neurological disorders like Parkinson's disease and stroke to mental health challenges such as depression and autism, music therapy is poised to gain broader acceptance within the medical community and secure inclusion in insurance coverage. Greater recognition will make music therapy a common complementary treatment, improve access for all, and establish its role in comprehensive patient care. Moreover, the implementation of VR in music therapy is expected to bring great transformation to music therapy.

Works Cited

- Aalbers, Sonja, et al. "Efficacy of Emotion-Regulating Improvisational Music Therapy to Reduce Depressive Symptoms in Young Adult Students: A Multiple-Case Study Design." *The Arts in Psychotherapy*, vol. 71, Elsevier BV, Nov. 2020, pp. 101720–20, <https://doi.org/10.1016/j.aip.2020.101720>.
- Anita Chen Marshall. "Traditional Chinese Medicine and Clinical Pharmacology." Springer eBooks, Springer Nature, Jan. 2020, pp. 455–82, https://doi.org/10.1007/978-3-319-68864-0_60.
- "APA PsycNet." *Apa.org*, 2024, psycnet.apa.org/record/1974-01018-000.
- Apurv Shirsat, et al. "Music Therapy in the Treatment of Dementia: A Review Article." *Cureus*, Cureus, Inc., Mar. 2023, <https://doi.org/10.7759/cureus.36954>.
- Boeldt, Debra, et al. "Using Virtual Reality Exposure Therapy to Enhance Treatment of Anxiety Disorders: Identifying Areas of Clinical Adoption and Potential Obstacles." *Frontiers in Psychiatry*, vol. 10, Frontiers Media, Oct. 2019, <https://doi.org/10.3389/fpsy.2019.00773>.
- Christine Judith Nicholls. "Friday Essay: The Remarkable Yidaki (and No, It's Not a 'Didge')." *The Conversation*, 6 Apr. 2017, theconversation.com/friday-essay-the-remarkable-yidaki-and-no-its-not-a-didge-74169.
- Dahl, Ronald E. "The Teenage Brain: Surging Hormones—Brain-Behavior Interactions during Puberty - Jiska S. Peper, Ronald E. Dahl, 2013." *Current Directions in Psychological Science*, 2016, journals.sagepub.com/doi/10.1177/0963721412473755.
- Deutscher. "Brain Development during Adolescence (21.06.2013)." *Deutsches Rzteblatt*, 2024, www.aerzteblatt.de/int/archive/article/141135.
- E. Glenn Schellenberg. "Music Lessons Enhance IQ - E. Glenn Schellenberg, 2004." *Psychological Science*, 2016, journals.sagepub.com/doi/10.1111/j.0956-7976.2004.00711.x.
- Frischen, Ulrike, et al. "The Relation between Rhythm Processing and Cognitive Abilities during Child Development: The Role of Prediction." *Frontiers in Psychology*, vol. 13, Frontiers Media, Sept. 2022, <https://doi.org/10.3389/fpsyg.2022.920513>.
- Gangopadhyay, Abirlal, and R. Prasad. "Therapeutic Elements of Music in Ancient India: A Brief Review in *Bṛhatrayī*." *DELETED*, vol. 57, no. 2, Springer Science+Business Media, June 2022, pp. 78–86, <https://doi.org/10.1007/s43539-022-00043-3>.
- Garrison, Amelia. "How Music Prevents Cognitive Decline | Pacific Neuroscience Institute." *Pacific Neuroscience Institute*, 21 Dec. 2021, www.pacificneuroscienceinstitute.org/blog/alzheimers-disease/how-music-prevents-cognitive-decline/.
- Geretsegger, Monika, et al. "Music Therapy for People with Autism Spectrum Disorder." *Cochrane Library*, vol. 2016, no. 3, Elsevier BV, June 2014, <https://doi.org/10.1002/14651858.cd004381.pub3>.
- Gold, Christian, et al. "Effects of Music Therapy for Children and Adolescents with Psychopathology: A Meta-Analysis." *Journal of Child Psychology and Psychiatry*, vol.

- 45, no. 6, Wiley, July 2004, pp. 1054–63,
<https://doi.org/10.1111/j.1469-7610.2004.t01-1-00298.x>.
- Gordon, Reyna L., et al. “Does Music Training Enhance Literacy Skills? A Meta-Analysis.” *Frontiers in Psychology*, vol. 6, Frontiers Media, Dec. 2015,
<https://doi.org/10.3389/fpsyg.2015.01777>.
- Hei Long Lam, et al. “Effects of Music Therapy on Patients with Dementia—a Systematic Review.” *Geriatrics*, vol. 5, no. 4, Multidisciplinary Digital Publishing Institute, Sept. 2020, pp. 62–62, <https://doi.org/10.3390/geriatrics5040062>.
- Hilliard, R. E. “The Effects of Music Therapy on the Quality and Length of Life of People Diagnosed with Terminal Cancer.” *Journal of Music Therapy*, vol. 40, no. 2, Oxford University Press, June 2003, pp. 113–37, <https://doi.org/10.1093/jmt/40.2.113>.
- Jaakko Erkkilä, et al. “Individual Music Therapy for Depression: Randomised Controlled Trial.” *The British Journal of Psychiatry*, vol. 199, no. 2, Cambridge University Press, Aug. 2011, pp. 132–39, <https://doi.org/10.1192/bjp.bp.110.085431>.
- Lally, Martha, and Suzanne Valentine-French. *Lifespan Development: A Psychological Perspective Second Edition*. 2019, dept.clcillinois.edu/psy/LifespanDevelopment.pdf.
- Landis-Shack, Nora, et al. “Music Therapy for Posttraumatic Stress in Adults: A Theoretical Review.” *Psychomusicology Music Mind and Brain*, vol. 27, no. 4, American Psychological Association, Dec. 2017, pp. 334–42, <https://doi.org/10.1037/pmu0000192>.
- Lutz Jäncke. “Music, Memory and Emotion.” *Journal of Biology*, vol. 7, no. 6, BioMed Central, Jan. 2008, pp. 21–21, <https://doi.org/10.1186/jbiol82>.
- M Soledad Cepeda, et al. “Music for Pain Relief.” PubMed, National Institutes of Health, Apr. 2006, <https://doi.org/10.1002/14651858.cd004843.pub2>.
- Montello, L., and EE Coons. “Effects of Active versus Passive Group Music Therapy on Preadolescents with Emotional, Learning, and Behavioral Disorders.” *Journal of Music Therapy*, vol. 35, no. 1, Oxford University Press, Mar. 1998, pp. 49–67,
<https://doi.org/10.1093/jmt/35.1.49>.
- “Musicians Bring Healing to Wounded Warriors.” www.army.mil, 8 Sept. 2015,
www.army.mil/article/155034/musicians_bring_healing_to_wounded_warriors.
- “Neural Overlap in Processing Music and Speech | Philosophical Transactions of the Royal Society B: Biological Sciences.” *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2015, royalsocietypublishing.org/doi/10.1098/rstb.2014.0090.
- Pando-Naude, Victor, et al. “An ALE Meta-Analytic Review of Top-down and Bottom-up Processing of Music in the Brain.” *Scientific Reports*, vol. 11, no. 1, Nature Portfolio, Oct. 2021, <https://doi.org/10.1038/s41598-021-00139-3>.
- . “Rhythmic Auditory Stimulation for Motor Rehabilitation in Parkinson’s Disease.” *Cochrane Library*, vol. 2024, no. 2, Elsevier BV, Feb. 2024,
<https://doi.org/10.1002/14651858.cd015759>.
- Rickson, Daphne J., and Katrina McFerran. “Music Therapy in Special Education: Where Are We Now?.” *Kairaranga*, vol. 8, no. 1, New Zealand Ministry of Education. Available

- from: Massey University. Private Bag 11 222; Palmerston North 4442, New Zealand. Tel: +64-6-351-3396; Fax: +64-6-351-3472; email: kairaranga@massey.ac.nz; Web site: http://www.massey.ac.nz/massey/learning/departments/school-curriculum-pedagogy/kairaranga/kairaranga_home.cfm, 2024, pp. 40–47, eric.ed.gov/?id=EJ914615.
- S. de l'Etoile. "The History of the Undergraduate Curriculum in Music Therapy." *Journal of Music Therapy*, vol. 37, no. 1, Oxford University Press, Mar. 2000, pp. 51–71, <https://doi.org/10.1093/jmt/37.1.51>.
- Shirlene Vianna Moreira, et al. "Music Therapy Enhances Episodic Memory in Alzheimer's and Mixed Dementia: A Double-Blind Randomized Controlled Trial." *Healthcare*, vol. 11, no. 22, Multidisciplinary Digital Publishing Institute, Nov. 2023, pp. 2912–12, <https://doi.org/10.3390/healthcare11222912>.
- Spencer, Jenna. A Historical Review of Music Therapy and the Department of Veterans Affairs. 2013, digitalcommons.molloy.edu/cgi/viewcontent.cgi?article=1011&context=etd.
- Thaut, Michael H., et al. "Neurobiological Foundations of Neurologic Music Therapy: Rhythmic Entrainment and the Motor System." *Frontiers in Psychology*, vol. 5, Frontiers Media, Feb. 2015, <https://doi.org/10.3389/fpsyg.2014.01185>.
- Wheeler, B. L. "Bruscia, K. E., Ed. (1998). *The Dynamics of Music Psychotherapy*. Gilsum, NH: Barcelona Publishers. 584 Pages. ISBN 1-891278-05-3. \$45.00." *Music Therapy Perspectives*, vol. 17, no. 2, Oxford University Press, Jan. 1999, pp. 104–5, <https://doi.org/10.1093/mtp/17.2.104>.
- . "Grocke, D., & Wigram, T. (2007). *Receptive Methods in Music Therapy: Techniques and Clinical Applications for Music Therapy Clinicians, Educators and Students*. London and Philadelphia: Jessica Kingsley. 271 Pages. ISBN 1-84310-413-X. \$34.95." *Music Therapy Perspectives*, vol. 25, no. 2, Oxford University Press, Jan. 2007, pp. 127–29, <https://doi.org/10.1093/mtp/25.2.127>.
- Xu, Chengyan, et al. "Potential Benefits of Music Therapy on Stroke Rehabilitation." *Oxidative Medicine and Cellular Longevity*, vol. 2022, Hindawi Publishing Corporation, June 2022, pp. 1–11, <https://doi.org/10.1155/2022/9386095>.

Figure Skating Training and Balance: A Quantitative and Comparative Analysis

By Maytal Zasler

Abstract

The positive effects of physical exercise on brain neuroplasticity have been recognized for many years. Figure skating training (FST) is correlated with beneficial changes in cerebral components that control motor skills and balance development. The areas of the brain affected by the repetitive motor skills and balance training required for figure skating, such as the cerebellum as well as visual and vestibular cortices, have also been linked with improvements in higher-level coordination and balance function. The objective of this study was to examine the effects of FST on objectively measured, higher-level balance ability in adolescent females compared with a control group of adolescent females who did not participate in regular FST. The Balance Error Scoring System (BESS) was utilized as an objective and portable static stability assessment tool for this study. The BESS requires both floor and foam pad examinations. Twenty female figure skaters and twenty female non-figure skaters were tested with the BESS during the study period. BESS scores were collected for both floor and foam pad performance and statistically analyzed. Study results from the BESS assessments indicated that regular FST was positively associated with better performance on measures of balance. These results have significant implications for understanding the underlying consequences associated with specific types of regular exercise training on brain neuroplasticity. The figure skaters were the Independent Variable and the non-figure skater sample acted as the Control. The Dependent Variable observed was the corresponding BESS score of each participant. The results of this study suggest that prolonged and consistent figure skating training (2+ years, 4+ hours per week) improves balance, with the figure skaters demonstrating an average of 10.8 errors and the non-figure skaters demonstrating an average of 14.85 errors. ($t(19) = 3.59, p = .00047, \text{Cohen's } d = 0.99$).

Introduction

The brain, an intricate network of critical connections controlling our thoughts, emotions, and movements remains one of the last frontiers of human scientific exploration. The brain can be positively influenced through many different kinds of experiences including the practice of repetitive motor skills. Internal as well as external stimuli have been shown to have the potential to produce a reorganization of not only brain structure but also connectivity and function (Cramer, 2011). The cerebellum, located in the posterior lower portion of the brain, is a major control center for movement and balance. Long-term FST and balance training have been proven to increase cerebellar volume, as well as enhance cortical brain thickness in other areas associated with balance function including the visual and vestibular areas of the brain (Rogge, Röder, & Zech, 2018; Zhang, Liu, & Liu, 2021). Prior research on the neuroplastic effects of repetitive motor behaviors in the context of FST would suggest that longer-term skaters might have better overall balance skills than non-skaters given enhanced brain connections and volume

resulting from such training. It is well established that regular physical exercise and repetitive motor behaviors enhance neuroplastic change in the brain (Hotting & Röder, 2013).

The connection between better balance and neuroplastic changes in both the structure and volume of critical brain structures is supported by neuroscientific research. Figure skating is an exercise in motor skill development, and the repetitive nature of the sport is traced to notable improvements in these motor skills (Zhang et al., 2021). A 2021 study with a sample size of 19 high-level figure skaters and 15 non-figure skaters found larger amounts of gray matter in the figure skaters, primarily in the cerebellum (Zhang et al., 2021). Given that gray matter tissue is essential to neurological function and the maintenance of fine motor skills, the results of the experiment described above provide foundational evidence for the significance and purpose of this study. In other words, figure skating and the brain are inherently connected.

A peer-reviewed study on the connection between physical activity and brain health states that physical activity is “a non-pharmacologic therapy for different pathological affections as well as for the maintenance of general health status,” (Di Liegro, 2019, para. 5) suggesting exercise to be a determinative and likely epigenetic factor for health. Epigenetics is the study of phenotype, or gene expression, variation caused by environmental factors. Epigenetic factors do not directly alter an organism’s DNA sequence but rather the way its genetic code is expressed. This same peer-reviewed study found that “epigenetic mechanisms induced by [physical activity] can build up an “epigenetic memory” that affects long-term brain plasticity, neurogenesis, and function” (Di Liegro, 2019, para. 14). Since figure skating training is a form of physical activity and, therefore, an epigenetic factor, then athletes with prolonged experience would likely demonstrate enhancement of specific gene activity driving neuroplastic change.

According to PubMed Central research in the biomedical category, “exercise is a subsection of physical activity that is structured, premeditated, and cyclic and has as a final or an intermediate objective the progress or physical fitness maintenance related to both skills- or health-related aspects” (Plaza-Diaz, 2022, para. 2). Further supporting the connection between figure skating training and epigenetic activity, skating closely adheres to the following stated standards for traditional exercise. It is repetitive and cyclic, meaning that prolonged time on the ice could be translated to increased epigenetic activity, and therefore improved neuroplasticity, or “proper neuronal signaling” (Mansuy, 2011, para. 12). Epigenetics have been found to specifically manifest in differentiated cellular functions in the brain as epigenetic activity targets individual cells. It is proven that “while neuronal and epithelial cells in a given person contain exactly the same DNA, epigenetic processes can modulate gene activity in each cell type differently, and thus activate or silence specific cellular functions or features” (Mansuy, 2011, para. 8). Research which shows that physical activity, such as figure skating, is tied to positive epigenetic variation and expression suggests that such activity may also promote neuroplastic change.

This causal relationship can also be observed in brain scans which demonstrated greater volume in the right cerebellar hemisphere and vermian lobules VI-VII in speed skaters versus non-speed skaters (Park, 2012). According to this study, these regions of the brain are associated

with balance and coordination, meaning that the intricate motor skills required for speed skating are fortified with repetitive motor movements and balance skill practice. Studies done in rodent populations support the connection between increased volume of brain cortices and more complex balance ability and performance. For example, a rodent study that observed and analyzed brain changes after an 8-day rotarod, or motor coordination, training period found increased volume of the cerebellum, frontal cortices, and amygdala (Surgent, 2020, para. 4). These findings suggest that regimented motor skill exercise, like FST, promote neuroplastic changes in brain regions associated with movement control and balance.

Statement of the Problem

A gap exists in the biomedical literature that fails to directly correlate prolonged figure skating training with improved higher-level balance function through objective/quantitative standardized balance testing methods. Although studies have shown volumetric differences in the cerebellum between ice-skating athletes and non-ice-skating athletes, these studies did not objectively assess actual balance skills. To model the implied functional consequences of these neuroplastic changes on balance skills associated with FST, prior studies have conducted brain scans such as three-dimensional magnetic resonance imaging. Direct and quantitative tests of balance like BESS have not been used to demonstrate this connection. This study aimed to establish a positive correlation between FST and improved balance by using an objective/quantitative assessment of balance/movement control.

Research Questions

The literature reviewed for this study suggests that longer-term FST and other motor activities increase neuroplasticity in brain regions associated with balance function. As a consequence, such activity should positively impact functional balance skills and be measurable. The following hypotheses are based on established literature in this area of neuroscience:

1. Does FST improve higher-level balance function in adolescent females (14-18 years old)?
2. Will figure skaters have lower or higher Balance Error Scoring System (BESS) scores than non-figure skaters?

Description of Variables

Table 1

Description of Variables

Variable Type	Variable
Independent	Number of Years Skated
Dependent	Total BESS Scores for Figure Skaters
Control	Total BESS Scores for Non-figure Skaters
Constants	Same balance pad, test conducted barefoot with eyes closed and hands on hips

Hypotheses

The examined literature supports the premise that FST and other repetitive skating motor tasks are associated with neuroplastic changes in areas of the brain associated with balance function. Such enhancements to participants' balance function should be demonstrated by objective measures of balance such as the BESS when compared with BESS scores of a control group of non-figure skaters. The null hypothesis is that there is no effect of FST on balance function for adolescent females. The alternative hypothesis is that adolescent females with at least two years of figure skating training will have better balance scores on the BESS when compared to the BESS scores of an age-matched control group of adolescent females with no FST experience.

Methodology

Institutional Review Board

This study was reviewed and approved by the Institutional Review Board (IRB).

Consent

This study investigated 20 figure skaters and 20 non-figure skaters. All participants were required to sign a consent form (see Figure 2) prior to participation. The form was signed by the participant and the parent/guardian if the subject was under the age of 18. The primary investigator reviewed the study as well as the associated risks with each subject.

Testing Protocol

The figure skaters tested were screened to ensure they had a minimum of two years of consistent experience before participating in the study. For this study, "consistent experience" is defined as 4+ hours on the ice a week. As part of the intake, the participant's age and years of skating experience were documented.

The BESS was conducted for each participant following the stipulated protocol. All tests were carried out on a hard flat surface with shoes removed (socks could be worn). The BESS contains three subsections of testing, repeated twice, that are each performed for twenty seconds with the subjects' eyes closed and hands on iliac crests/lower hips. A stopwatch was needed to accurately time each exercise. The first round of testing was done on a hard, flat surface and the second series of tests was done on an aerobic foam block placed on a hard, flat surface. A 41 by 50 cm Airex Balance Pad was used for this experiment. This pad was sourced from a concussion clinic that uses the Airex to perform concussion testing on patients. The purpose of the balance pad was to increase the difficulty of the exercises by creating a surface that fluctuates and deforms with added weight. A spotter was present to stand by the participant while they performed each trial to decrease fall risk, especially on the foam pad. Exercise #1: Two-foot stance with feet facing forward shoulder-length apart. Exercise #2: One-foot, non-dominant stance with free leg held up at a 90-degree angle. Exercise #3: Tandem stance with the tip of the non-dominant foot touching the heel of the dominant foot.

Data Collection

The BESS was scored out of sixty total points as directed, where higher scores indicated more errors made during testing in deviation from the original exercise position. The criteria for an error are listed in Figure 1 below as per the University of North Carolina Chapel Hill’s Sports Medicine Research Laboratory (University of North Carolina Sports Medicine Lab). Error counting began as soon as the participant achieved the initial position. Every error made was counted as one point toward the participant’s total score.

Figure 1

Score Card		
Balance Error Scoring System (BESS) (Guskiewicz)		
Balance Error Scoring System – Types of Errors 1. Hands lifted off iliac crest 2. Opening eyes 3. Step, stumble, or fall 4. Moving hip into > 30 degrees abduction 5. Lifting forefoot or heel 6. Remaining out of test position >5 sec The BESS is calculated by adding one error point for each error during the 6 20-second tests.	SCORE CARD: (# errors)	FIRM Surface
	Double Leg Stance (feet together)	
	Single Leg Stance (non-dominant foot)	
	Tandem Stance (non-dom foot in back)	
	Total Scores:	
	BESS TOTAL:	

Figure 1 was developed by researchers and clinicians at the University of North Carolina at Chapel Hill Sports Medicine Research Laboratory.

Figure 2

A copy of the consent form used for this study is below in Figure 2.

Effects of Figure Skating on Higher-Level Balance Function in Adolescent Females

Contact: maytal_zasler25@collegiate-va.org

By signing below, I give consent for my child to participate in this High School research study focusing on the effects of figure skating on higher-level balance. My child will complete the BESS Test (Balance Error Scoring System) in which they will carry out three stand-still exercises with their eyes closed for twenty seconds each — two feet on the ground, one leg stance (non-dominant), tandem stance (dominant in front of non-dominant) — once on firm ground and once on a 3-inch foam pad. There will be a spotter present during the test to reduce fall risk. Twenty figure skaters and twenty non-figure skaters will be tested.

Participant identities are kept confidential and no personally identifying information will be used in the final report. The study includes a pre-screening survey with the following questions: Age? Years of figure skating experience (if applicable)? Prior concussion(s) or ankle injury? *Participants must be in an age range of 14-18 years old, and figure skaters must have a minimum of two years of consistent skating experience (4+ hours per week). This study offers no financial compensation and poses a low physical risk to all screened participants.

PARENT or SKATER 18 YEARS OLD: I have read and understood the parameters of this research study and give consent for my child to participate.

Date: _____

Print Name: _____

Signature: _____

CHILD: I have read and understood the parameters of this research study and give my consent to participate.

Date: _____

Print Name: _____

Signature: _____

Results

Data Analysis

Table 2 shows raw data collected for each participant according to age and BESS score. The Independent sample's *t*-test was used to determine if there were any significant differences in mean BESS scores between skaters and non-skaters. Table 3 shows means and standard deviations. The 20 participants who were classified as skaters ($M = 10.8$, $SD = 3.71$) compared to the 20 participants who were classified as non-skaters ($M = 14.85$, $SD = 4.44$) demonstrated higher levels of balance ability according to the BESS. Results of the *t*-test support these findings with $t(19) = 3.59$, $p = .00047$. Cohen's $d = 0.99$, and suggest a large effect size. The null hypothesis for the present study was rejected.

Table 2
Raw Scores

Age	Total BESS Scores	
	Skaters	Non-Skaters
14	17	17
15	17	13
15	6	13
15	13	9
15	13	8
15	9	12
15	8	13
15	12	15
15	13	17
16	9	20
16	7	7
16	5	5
17	6	6
17	7	7
17	9	9
17	12	12
17	10	10
18	11	20
18	16	16
18	16	16

Table 3
Descriptive Statistics

	Skaters	Non-Skaters
Mean BESS Score	10.8	14.85
Standard Deviation	3.71	4.44

Discussion of Results
Threats to Validity

This study does not take into full consideration all variables such as other sports activity participation by the non-skater sample. The study did not account for how other physical activity and training experiences might have affected the control group’s balance abilities as measured by the BESS. Other factors such as prior injury or current illness may have affected the results as

well. If the participant had a recent ankle, knee, or concussive brain injury, then their scores may have been adversely affected beyond any impact due to non-participation in FST. Hearing problems or ear infections could have also potentially interfered with balance function and resulted in lower BESS scores.

Limitations

Since this study only assessed female figure skaters and non-figure skaters, it is not possible to generalize the results to male skaters or to adolescents involved in other activities that may impact neuroplasticity and balance function. Also posing a limitation to this study, the research was conducted in different locations with different factors and distractions present that may have affected participants during testing. Because the BESS was not conducted in the same controlled environment for each test, the data may have been impacted. Additionally, due to the small population of figure skaters in the local geographic area where this study was conducted, and the fact that the population was exclusively female, the sample for the Independent Variable was limited.

Recommendations

For future studies examining this issue, it would be beneficial to conduct each test in the same distraction-free environment to not disrupt the participant during testing. If the testing environment is a constant, then the data would be more comparable along the lines of effective and accurate research. It would also be helpful to have a larger sample size, with more varied years of skating experience across both male and female subjects and a broader age range to allow for sub-analysis, so that more conclusive data could be collected. For example, does the number of years of skating experience have an effect on the skaters' balance ability and BESS scores? Another recommendation to improve any future replication efforts would be to include information on prior lower extremity injuries and/or concussive brain injury as both could impact BESS outcomes. With this information present, the study becomes more specific and better able to explain why some scores may be lower than others within a sample. This data can help identify a causal relationship between irregular weakness and lower BESS scores.

Conclusion

The null hypothesis that there is no effect of FST on balance function for adolescent females was rejected by the findings of this study. The neuroscientific literature addressing the positive impacts of repetitive training and motor practice in the context of FST was upheld by the results of the current study. Adolescent females with at least two years of figure skating training had better balance scores on the BESS when compared to the BESS scores of an age-matched control group of adolescent females with no FST experience.

Works Cited

Peer-Reviewed

1. Cramer, S.C., Sur, M., Dobhin, B.H., et al. (2011). Harnessing Neuroplasticity for Clinical Applications. *Brain*. 134 (Pt 6): 1591-1609. <https://pubmed.ncbi.nlm.nih.gov/21482550/>.
2. Hötting, K., Röder, B. (2013). Beneficial Effects of Physical Activity on Neuroplasticity and Cognition. *Neurosci Biobehav Rev*. 37 (9 Pt B): 2243-57. <https://pubmed.ncbi.nlm.nih.gov/23623982/>.
3. Di Liegro, C.M., Schiera, G., Proia, P., et al. (2019). Physical Activity and Brain Health. *Genes*. 10 (9): 720 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6770965/>.
4. Mansuy, I.M., Mohanna, S. (2011). Epigenetics and the Human Brain: Where Nurture Meets Nature. *Cerebrum*. Epub 2011 May 25. <https://pubmed.ncbi.nlm.nih.gov/23447777/>.
5. Park, I.S., Lee, N.J., Kim, T.Y., et al. (2012). Volumetric Analysis of Cerebellum in Short-Track Speed Skaters. *Cerebellum*. 11 (4): 925-30. <https://pubmed.ncbi.nlm.nih.gov/22351379/>.
6. Plaza-Diaz, J., Izquierdo, D., Torres-Martos, A., et al. (2022). Impact of Physical Activity and Exercise on the Epigenome in Skeletal Muscle and Effects on Systemic Metabolism. *Biomedicines*. 10 (1): 126. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8773693/>.
7. Rogge, A.K., Rödder, B., Zech, A., et al. (2018). Exercise-induced Neuroplasticity: Balance Training Increases Cortical Thickness in Visual and Vestibular Cortical Regions. *Neuroimage*. 1:179:471-479. <https://pubmed.ncbi.nlm.nih.gov/29959048/>.
8. Sargent, O.J., Dadalko, O.I., Pickett, K.A., et al. (2020). Balance and the Brain: A Review of Structural Brain Correlates of Postural Balance and Balance Training in Humans. *Gate Posture*. 71:245-252. <https://pubmed.ncbi.nlm.nih.gov/31082657/>.
9. Zhang, K., Liu, Y., Liu, J., et al. (2021). Detecting Structural and Functional Neuroplasticity in Elite Ice-Skating Athletes. *Hum Mov Sci*. Epub 2021 Apr 30. [Epubhttps://pubmed.ncbi.nlm.nih.gov/33940321/](https://pubmed.ncbi.nlm.nih.gov/33940321/).

Non-Peer Reviewed

1. University of North Carolina at Chapel Hill Sports Medicine Research Laboratory. Balance Error Scoring System (BESS). https://atriumhealth.org/documents/carolinasrehab/bess_manual_.pdf.

Formation of Children Consumer Behaviour in The Present and Future

By Wenjing Zhou

Abstract:

Past papers have individually highlighted the role of observational learning, parenting styles, Kahenman's system 2 thinking, and self control in influencing behaviour (Melztoff, Tzschentke 626-637, Baumrind 56-95, Kahneman 31-37). However, there remains a gap in understanding how these pre-established aspects can together contribute to the formation of consumer patterns among children and shape the trajectories of their future consumption.

This literature review bridges the gap between these theories and children's consumption behavior through a synthesis of these theories of parenting, cognition, observational learning and a case study on online consumer behavior. The paper only addresses children under the age of 8 who are at a critical developmental stage where foundational habits begin forming.

The first part of the review investigates how observational learning of parental consumptions establishes children's consumer behaviour through a case study on children's online consumptions. The second part investigates how Baurmind's theory of parenting style can further modify children's consumer patterns, directing their future development of rational consumptions. This is reasoned by the synthesized relationship between parenting styles with Kahneman's system 2 thinking and self control.

Through the pre-established idea of children's observational learning and a case study, the first part of this paper finds that parental consumer behavior shapes children in developing similar consumption patterns and choices.

The second part of study links the theory of parenting styles, system 2 thinking, and self control to reveal how authoritarian and authoritative parenting direct children's future development of rational consumptions. In contrast, permissive and neglective parenting direct the development of irrational consumption.

These behaviors then influence future market trends, which firms can predict and leverage for strategic planning.

Introduction

Consumer behavior comprises the actions and decisions that people or households make when they choose, buy, use, and dispose of a product or service (Brooks). The study of consumer behaviour aids firms in understanding drives for consumer spending patterns, which allows them to adjust marketing strategies and policies quickly to enhance resource allocation and maximize profit. These behaviours can easily change in response to social trends, technological advancements, political changes, etc. For instance, through technological advancements, trading has transformed from the barter system, to cash exchanges, and now becoming the world of digital transactions. Nowadays, people would much prefer purchasing from stores that accept card payments, than cash. These changes in consumer behavior reshape the market, making it essential for firms to redevelop strategies to take advantage of these shifts.

In the future, as older generations age, the children, the younger generations will be the ones driving upcoming trends and advancements (Velasquez). This highlights the importance of studying the children's present consumer behaviour formation and patterns to predict their future purchasing tendencies.

Previous papers have highlighted the role of observational learning, parenting styles, development of Kahneman's system 2 thinking, self control individually in establishing and modifying behaviours. However, no paper has yet tied these established theories and ideas together, pointing out their collective effort in building and shaping children's consumer behaviour. This paper thus explores how parental influences shape children's current consumer behaviour and direct their future development of rational/irrational consumptions.

Building on previous scholarly papers and concepts, this study focuses specifically on children under the age of 8, a critical developmental stage where foundational habits begin to form. The first part of the paper examines how parental consumer behavior shapes children's current consumption desires. Based on the case study of children's online purchasing behavior, the study found that parental consumer behavior influences children in developing similar consumption patterns and choices. The reason behind this is that children learn observationally. Specifically, two observational learnings are discussed: Piaget's deferred imitation - the process of performing a previously observed action at a later time - and Lorenz's imprinting effect - a form of rapid, supposedly irreversible learning that results from exposure to an object during a specific period during early life (Meltzoff, Tzschentke 626-637).

By linking pre-existing literature of observational learnings -deferred imitation and imprinting effect- with the case study, the first part of the paper finds that children initially develop similar consumption patterns or desires similar to their parents' through early age observations. However, not all imitated purchases are relevant or appropriate for children, so parents continue to shape their children's established consumption patterns by regulating their current purchases. Understanding the importance of predicting children's future consumption, the second part of the study examines how parenting styles contribute to children's current purchasing decisions, thereby influencing the trajectory of their consumption patterns toward either rational or irrational consumer behaviors in the future.

To examine this, the second part of the study synthesised characteristics of Baumrind's different parenting style to the development of Kahneman's system 2 thinking and self control. The finding is that authoritarian and authoritative parenting have a greater probability of influencing children's future development of rational consumer behaviour. Permissive and neglectful parenting, on the other hand, has a greater possibility of influencing children's development of irrational consumer behaviours in the future.

Although the limitations of this study include some reliance on secondary sources rather than primary ones, it still provides robust concepts, theories and data-driven insights into development of children's consumer behaviour.

The findings of this study can help firms understand future market trends by predicting children's future irrational or rational consumer behavior based on different parenting cultures across

demographics. This understanding enables firms to plan marketing strategies ahead of time, tailoring their approaches to better align with anticipated consumer behaviors. Thus, they can make more informed decisions regarding marketing strategies and inventory management, leading to greater profitability.

1. How Do Parental Consumer Behaviour Shape Children Consumption Desires?

Case 1: Children's Adoption of E-Purchasing Behaviour

The study explores the relationship between children and adult consumer behaviours through the case of children's adoption of E-purchasing behaviour. At ages between 2-7 years old, some children have demonstrated the ability to make intentional or unintentional online purchases. Three children's online purchases from the news were studied. The first was a six year old girl from Quebec, who with the help of voice commands made a \$2,100 purchase on Amazon (Lofaro). The second evidence presents a toddler who cleared out her/his mom's online shopping cart, with \$2,000 worth of items (Burke). The last piece illustrates a £1,200 spending on online games by a seven year old boy (Norris).

All 3 instances share a commonality: parents and family members had performed past online purchases around the children. This repeated exposure to online shopping activities suggested that children likely observed this behaviour, and imitated it (Bandura 3-4). Thus, it is plausible that children's online consumption is an observational learning of the parents' action of purchasing online. Specifically, children imitate their observations of the parents through deferred imitation - the process of performing a previously observed action at a later time - and imprinting effect - a form of rapid, supposedly irreversible learning that results from exposure to an object during a specific period during early life (Meltzoff, Tzschentke 626-637).

Children's Consumer Learning From Parents Through Deferred Imitation and Imprinting

How does the learning mechanism of deferred imitation and imprinting specifically apply to children? According to Piaget's cognitive development theory, the toddler is likely to have been in the sensorimotor stage or at the beginning of the preoperational stage, which are the initial phases of cognitive development (1-11). Out of most cognitive processes at sensorimotor stage, only deferred imitation directly relates to the learning of physical movements and replication of behaviour. Most other common processes such as object performance and animism are mostly mental works on perception of objects, provoking simple behaviours such as searching or playing with a teddy bear. Moreover, deferred imitation allows for the replication of more complex behaviors that involve multiple steps and actions, such as the case's online purchasing behaviour. That's because it does not necessarily require an understanding of the behaviour that is replicated. Thus, toddlers can encode, retrieve from memory, and present actions without fully grasping the underlying concepts, which is not the case for many other cognitive and physical processes at this age (Meltzoff).

Connecting this idea to the case of children's online purchases, parent's past behavior regarding online purchases were likely exposed to children, who observed this behavior closely.

Given the similarities in observed behavior and the complexity involved, deferred imitation is likely the most appealing process for understanding the development of toddlers' online purchasing behaviors.

Furthermore, the ability to engage in deferred imitation continues to develop as children grow (Jones). Deferred imitation thus appeals to older children in evidence 1 and 3 who exhibit similar online purchases as well.

On the other hand, the similar behaviours of E-purchases between children and parents can also be attributed to the children's learning of parental consumption through the imprinting effect. The difference between deferred imitation and the imprinting effect however, is that while deferred imitation can replicate anyone's behaviors, imprinting involves behavioral learning specifically from caregivers to whom children are attached. By forming this early attachment to their caregivers - parents - the children form social and emotional bonds which support their learning of behaviours and characteristics from their parents (Tzschentke 626-637). Thus, as parents have shown the E-purchasing behaviour before, children likely observed and learnt the behaviour from them through imprinting effect.

In summary, through the learning processes of imprinting and deferred imitation, children absorb and replicate parental consumer behaviors, which significantly influences their propensity to select and consume similar products.

2. How do parenting styles further shape children's consumer behaviour and direct the development of future rational/ irrational consumption patterns?

The first part of the study concluded parental purchases influence their children in developing similar consumption desires and patterns. However, not all of the products they imitate are relevant or appropriate for children to use. Thus, parents not only influence which products their children gravitate towards but also regulate these purchases through their parenting styles. This regulation facilitates the development of self control and system 2 thinking through practices like enforcing rules, offering guidance, and providing positive reinforcement. Parents help children learn to manage their impulses and make thoughtful decisions (Zeinali). The development of self control and system 2 thinking continues shaping the children's consumer behaviours/patterns, especially in directing future rational purchases - consumptions that maximise private benefits (Ganti). Likewise, parenting styles that neglect or give in to children's request will fail to facilitate their development of self control, directing children's future irrational behaviours - consumptions without logical reasons or clear thinking, such as impulse purchases, pestering (Wang).

Types of Parenting Styles

Different styles of parenting contributes to the development of self control which continues to modify the children's consumer behaviours, preventing irrational consumption from occurring in the future. There are four classic types of parenting styles: authoritative, authoritarian, neglectful, and permissive (Baumrind 56-95). Baumrind states that the styles are measured by two aspects: demandingness, which refers to the claims that parents make on children to become integrated into

the family and community through their maturity expectations, supervision, disciplinary efforts, and willingness to confront a disputative child, and responsiveness, which refers to the extent to which parents intentionally foster individuality, self-regulation, and self-assertion by being attuned, supportive, and acquiescent to children's special needs and demands. She states that authoritarian parents, who are demanding but not responsive, impose strict expectations on their children and expect high levels of obedience. Thus, they stand their ground and strictly reject their children's unnecessary consumer requests. Baumrind also indicates that permissive parents, who are responsive but not demanding, are very lenient and impose low levels of responsibilities or expectations on their children, allowing them to do whatever they want. Thus, they let children purchase whatever they want despite the products being inappropriate for young ages. She also states that neglectful parents, who are neither responsive nor demanding, provide minimal parental attention and care to their children, with no regulation of the children's behavior. Thus, similar to permissive parents, neglectful parents continue shopping, allowing children to purchase according to their will. Finally, Baumrind indicates that authoritative parents, who are both demanding and responsive, are nurturing and supportive but still have expectations that are appropriately enforced with consequences. They attempt to control their children's behavior by explaining rules, discussing, and reasoning. Thus, they likely reject children's inappropriate purchases but provide an explanation for their actions.

The Development of Self Control and System 2 under Different Parenting Styles

There is a commonality between authoritative and authoritarian parenting is that they are both demanding and thus would likely reject inappropriate or unnecessary consumer requests from children. Experiencing rejection to consumer requests allows the children to learn the concept of delayed gratification, the ability to wait for something they want, which requires the use of self-control, the ability to manage one's impulses, emotions, and behaviors to achieve long-term goals, to resist immediate impulses (Psychology Today). By consistently practicing delayed gratification, individuals enhance their self-regulation skills, which, according to Daniel Kahneman, is essentially the strengthening of our system 2 - a slower, more deliberate, and analytical mode of thinking that requires conscious effort. The practice of delayed gratifying activities or in other words system 2 activities strengthen self control because each repetition allows the regulation by System 2 to become more automatic, reducing the cognitive load required to maintain self-control in performing those activities (Kahneman 31-37). Therefore, self control can be more easily executed and maintained.

This is supported by another study that recorded children's responses to rejected requests. When a purchase request is refused, while some children may feel annoyed and disappointed, they eventually resign, even after pestering. As one seven year old girl noted: "I wouldn't really mind but, when I was smaller, I didn't actually understand but, now that I'm older, I understand that I have a brother and a sister and they need some things as well. I just understand that I can't get everything I need" (Lawlor). This cognitive shift from not understanding to understanding the parents' decisions shows an improved ability to delay gratification by recognizing the broader

context and long-term benefits, which in this case is family needs and fairness. Moreover, from annoyance and confusion to resignation, those emotions also illustrate the ability to maintain one's desires.

Self-control are effortful cognitive processes, but nature of the brain tends to conserve energy by defaulting to the easiest or most familiar way of thinking, avoiding self control processes as much as possible (Kahneman 31-37). This is why external reinforcers, such as authoritative and authoritarian parenting, are needed to strengthen and develop self-control. In turn, it also alternatively explains why neglectful and permissive parenting, who wouldn't reject or respond to children's requisition, cannot facilitate the development/reinforcement of self control and system 2 thinking.

Varying Development of Self Control Direct Children's Future Consumer Behaviour

Now that the study has established which and how parenting styles strengthen children's self control and system 2 thinking, it now examines how the two factors subsequently direct children's development of rational or irrational consumer behaviours in the future.

Self control/regulation is a critical aspect of personal discipline and psychological well-being, enabling individuals to make decisions that align with their values and objectives (Cuncic). In the context of consumer behaviour and behavioural economics, it plays a critical role in combating impulses or according to Kahneman, most thinking from system 1 - fast, automatic, and intuitive mode of thinking that operates effortlessly and quickly, allowing system 2 to dominate. Thus, it impedes irrational consumptions such as impulse buying, pestering from occurring when they become financially independent in the future by shaping it into children's consumer patterns.

Differing Parenting Styles Direct Children's Future Consumer Behaviour

The second part of the study asks: How do parenting styles further shape children's consumer behaviour and direct the development of future rational or irrational consumption patterns? It can be concluded that authoritative and authoritarian parenting facilitate the development of self control and system 2, likely directing more of children's rational consumer behaviours in the future. In contrast, neglectful and permissive parenting fails to reinforce the development of self control and system 2, likely directing more of children's irrational consumer behaviour in the future.

Discussion:

In sum, in the first part of the study, using the concept of children's observational learning and a case study, the research finds that parental consumer behavior influences children to develop similar consumption patterns and choices.

In the second part of the study, the paper links parenting styles with System 2 thinking and self-control, concluding that authoritarian and authoritative parenting guide children toward the future development of rational consumption habits. Conversely, permissive and neglectful parenting are associated with the future development of irrational consumption behaviors.

This finding helps firms predict marketing strategies that can take advantage of and make early informed decisions based on understanding children's future consumer behavior according to varying parenting cultures across different demographics. For example, in China, where most parentings are more strict, ranging between authoritarian and authoritative, children's consumption would be more rational when they become financially independent in the future. Firms should emphasise marketing the practicality instead of the design aesthetics of the products to attract Chinese consumers in the future. This early understanding and planning of consumer patterns could boost consumption from the Chinese population and increase the firm's profitability, highlighting the importance of this study.

This is just one example of how the study's findings can be practically applied, and there are many more implications waiting to be explored.

Works Cited

- Blakeley, Sasha. "Deferred Imitation & Child Development | What is Piaget's Deferred Imitation? - Lesson." Study.com, <https://study.com/academy/lesson/what-is-deferred-imitation-definition-examples.html>. Accessed 2 August 2024.
- Burke, Minyvonne. "Toddler clears out mom's online shopping cart, orders \$2000 worth of items from Walmart." NBC News, 23 January 2022, <https://www.nbcnews.com/news/us-news/toddler-clears-moms-online-shopping-cart-orders-2k-items-walmart-rcna13172>. Accessed 2 August 2024.
- "Child Development and Early Learning - Transforming the Workforce for Children Birth Through Age 8." NCBI, <https://www.ncbi.nlm.nih.gov/books/NBK310550/>. Accessed 9 August 2024.
- Cuncic, Arlin. "Self-Regulation: How to Develop and Practice It." Verywell Mind, 5 May 2023, <https://www.verywellmind.com/how-you-can-practice-self-regulation-4163536>. Accessed 30 July 2024.
- Jones, Susan S. "The development of imitation in infancy - PMC." NCBI, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2865075/>. Accessed 8 August 2024.
- Kahneman, Daniel. Thinking, fast and slow. Farrar, Straus and Giroux, 2011, pp. 31-37.
- Lawlor, Margaret Anne. "The changing face of pester power." RTE, 14 December 2018, <https://www.rte.ie/brainstorm/2018/1213/1016954-the-changing-face-of-pester-power/>. Accessed 30 July 2024.
- Lofaro, Joe. "Don't tell mom: Quebec girl, 6, secretly racks up \$2,100 Amazon bill." CTV, 20 January 2023, <https://montreal.ctvnews.ca/don-t-tell-mom-quebec-girl-6-secretly-racks-up-2-100-amazon-bill-1.6239278?cache=%2F7.318878>. Accessed 2 August 2024.
- Norris, Sophie. "Boy, 7, spends £1200 on online games, including £800 on virtual cat food." The Mirror, 16 February 2021, <https://www.mirror.co.uk/news/uk-news/boy-7-accidentally-spends-1200-23508222>. Accessed 2 August 2024.
- Piaget, Jean. "The theory of stages in cognitive development." Measurement and piaget, McGraw Hill, 1971, pp. 1-11.
- Ganti, Akhilesh. "Rational Choice Theory: What It Is in Economics, With Examples." Investopedia, 29 June 2024, <https://www.investopedia.com/terms/r/rational-choice-theory.asp>. Accessed 9 August 2024.
- Wang, Jianhua. "Research on the Irrational Behavior of Consumers' Safe Consumption and Its Influencing Factors." NCBI, 6 December 2018, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6313737/>. Accessed 9 August 2024.
- Romer, Daniel. "Adolescent Risk Taking, Impulsivity, and Brain Development: Implications for Prevention." NCBI, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3445337/>. Accessed 30 July 2024.

- “Self-Control.” Psychology Today, <https://www.psychologytoday.com/ca/basics/self-control>. Accessed 30 July 2024.
- Zeinali, Ali. “The mediational pathway among parenting styles, attachment styles and self-regulation with addiction susceptibility of adolescents.” NCBI, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3430035/>. Accessed 9 August 2024.
- Brooks, Ruth. “What is consumer behaviour?” University of Lincoln, 28 September 2022, <https://online.lincoln.ac.uk/what-is-consumer-behaviour/>. Accessed 28 August 2024.
- Tzschentke, B., and A. Plagemann. "Imprinting and Critical Periods in Early Development." *World's Poultry Science Journal*, vol. 62, no. 4, 2006, pp. 626–637.
- Baumrind, Diana. The Influence of Parenting Style on Adolescent Competence and Substance Use. vol. 11, *The Journal of Early Adolescence*, 1991. SageJournals, <https://doi.org/10.1177/02724316911111004>.
- Meltzoff, Andrew N. “Infant Imitation After a 1-Week Delay: Long-Term Memory for Novel Acts and Multiple Stimuli.” NCBI, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4137879/>. Accessed 28 August 2024.
- Bandura, Albert. https://www.asecib.ase.ro/mps/Bandura_SocialLearningTheory.pdf. Accessed 28 August 2024.

Takeaways for Chinese Policymakers: Promoting Social Connectedness of Autistic Children and Young Adults in China By Xiaoxi Jia

Abstract

People with autistic spectrum disorder (ASD) tend to face various difficulties in their lives, such as social isolation and a lack of friendship, social network centrality, and social involvement. This condition fits in the current Chinese context. Improving social connectedness for the autistic population can mitigate their current challenges by helping develop their social network and thus gain more social support. However, little is known about the extent of China's policies on social connectedness. This paper aims to analyze the degree of social connectedness in Chinese policies and propose policy interventions that proved effective in past studies. The policies for people diagnosed with autism in China are extensive, but the proportion of social connectedness is not enough. Policymakers should consider increasing investment in rehabilitation centers and hospitals responsible for autism training. This can be achieved by conducting pilot projects and standardizing autism intervention by incorporating PEERS, pivotal response training, theater interventions, etc, which helps enhance social connectedness. Policies should also consider special job training for autistic young adults aside from only offering employment, thereby promoting their social skills with colleagues and working efficiency. Additionally, public knowledge of ASD should be included in the presentation in elementary schools and high schools to reduce bias and increase people's tolerance of autistic people.

Introduction

Autism Spectrum Disorder (ASD) is a neurological and developmental disorder that is characterized by impairments in communication and interactions, repetitive behaviors, restricted interest, and symptoms affecting daily activities (American Psychiatric Association). Families with autistic members can face several challenges, including financial pressure (Horlin et al.) and psychological burdens (Barker et al.). Among the autistic population, autistic children and autistic young adults are especially worth studying because of their special stages of development. For autistic children who are still in elementary school age, their developmental plasticity allows interventions to play an effective role that may affect their future lives. Children with ASD benefit their whole lives if the impediments to social interactions can be mitigated timely. On the other hand, autistic young adults may not be able to rely on guardians permanently, so developing the necessary skills to help autistic young adults sustain themselves.

Out of the approaches that aim to assist autistic people, this paper mainly focuses on enhancing social connectedness to autistic children and autistic young adults. Social connectedness represents the relationship bonding and the sense of belonging people perceive in their social groups (Haslam et al.). With increased social connectedness, autistic children and young adults can in turn be more likely to expand their structures of relationships and acquire social support (Haslam et al.).

The measurements of social network centrality and social network salience in a peer relationship are closely related to the social connectedness of an individual. Social network

centrality refers to the relative distance of social actors from the center of the social network (Zhang and Luo). The higher the social centrality of a person, the more convenience or power one can get from the social network (Zhang and Luo). Social network salience indicates the prominence of an individual in the structure of the group he or she belongs to (Locke et al.).

Public policy can effectively assist autistic people given its extensive affecting range and mandatory characteristics of implementation. This paper offers an overview and intervention proposal regarding social policy on autism in China. The literature review discusses the previous knowledge about effective interventions to enhance the social connectedness of autistic children and autistic young adults, thereby analyzing the focus of social connectedness in current Chinese policy and offering reasonable suggestions. This paper first arranges the current situations of social connectedness of children and young adults diagnosed with ASD. Then, the paper provides interventions useful for enhancing social connectedness. I clarify different policies for the autistic population in China. Incorporating the early analysis of effective interventions, I then analyze the extent of social connectedness in policies and suggest new policies.

Current Social Connectedness of Autistic Children and Autistic Young Adults

School is the setting most relevant with autistic children in elementary school age, in which social relationships are developed and most social connections occur, so focusing on their current situations of perceived social connectedness in the school setting helps identify the primary problems. Previous research dabbled in social networks, social involvement, peer relationships, and loneliness.

Previous research paid attention to the changes in social networks and several of their predictors. According to one research comparing autistic children with their non-autistic children, the social network salience for both groups increased over childhood and adolescent development, with upper-grade students having a more salient social network (Zhang & Luo, 2017). However, children diagnosed with ASD had significantly lower social network salience than non-autistic children (Locke et al.). For this reason, several potential predictors of social network connectivity of children with autism may aid in anchoring particular need-improving aspects, including gender, age, classroom size, and IQ (Anderson et al.). Initially, boys are more socially successful, but their connectivity gradually decreases as they grow older (Anderson et al.). In terms of gender and classroom size, girls typically perform better in larger classrooms, whereas the opposite is true for boys (Anderson et al.). Therefore, despite the increasing social network salience and relatively stable classroom connections over time, the social network connectivity of autistic children is disadvantaged, with predictors that help assist the direction of improvement.

The social involvement of children diagnosed with autism is also weaker than that of normal children. By analyzing the friend nomination reported by children with and without autism, the social network centrality and companionship of autistic pupils were less than typical developing pupils (Chamberlain et al.). What's more, the reciprocity and acceptance of friendships for autistic children were also fewer, while the level of loneliness was higher (Chamberlain et al.). Overall, the social involvement of autistic children gradually increases as their age gets older. However, they still

have significantly lower social involvement and more rejections from others compared to typically developing kids (Rotheram-Fuller et al.).

Peer relationship is a big component of social networks that have a huge impact on children's social development. Based on the studies investigating the experience of peer relationships of autistic children, similar interests and sharing activities are important for their friendships (Cresswell et al.). Because autistic people tend to have impairments in communication, some of them may experience peer rejection and victimization from others, hindering them from making new friends and developing a full understanding of friendships (Cresswell et al.). Children with ASD tend to have fewer friends, fewer contacts with friends, shorter-lasting friendships, and more friends with disabilities than typically developing children (Petrina et al.). The reciprocal friendships of children diagnosed with ASD are also fewer, becoming even fewer and fewer as they grow older (Petrina et al.).

When the endeavor of making friends fails, children with autism may suffer loneliness, sadness, and anxiety. In a previous study researching social competence and loneliness among autistic children, social difficulties among autistic children existed, such as difficulties in making friends (Zeedyk et al.). The loneliness of autistic children is associated with several factors, including social skills, behavioral problems reported by teachers, and conflicts between students and teachers (Zeedyk et al.). In addition, higher classmate closeness centrality is related to less level of loneliness (Tsou et al.), addressing the importance of joining interactions with individuals in the school community.

As for most young adults with ASD, their life settings are not only limited to school, but also community, workplace, or family. Thus, the components of their current social connectedness and difficulties are more complicated but still have generality.

Previous studies examined the current social network of young adults diagnosed with autism. Compared to typically developing individuals, autistic young adults tend to have more restrictive social networks and fewer reciprocal initiative contacts (Van Asselt-Goverts et al.). More often, young adults with ASD tend to be neutral or dissatisfied with their social network, and they desire to increase the range of their social networks (Van Asselt-Goverts et al.). Besides, some autistic young adults also suffer from social isolation, which is evinced by fewer social invitations and interactions (Orsmond et al.). Young adults with ASD also tend to have difficulties making friends, which is mainly because of their challenges in social relationships, lack of opportunities to socialize with others, uncomfortable social norms, and social differences (Sosnowy et al.). Beyond the friendship, autistic young adults have various ways to have meaningful social contacts. Some autistic adults perceive their workplaces as locations to have positive interactions or practice their social interactions, which can enhance their social connections (Chan et al.). Social activities with common interests, neighborhoods, support services, and online social networks can provide young people with ASD with a sense of belonging (Chan et al.).

The experiences of autistic young adults can be the exemplifications of the challenges they face. During real-world interaction, the interaction quality, desire to hang out, and warmth in the relationship are less important for some autistic adults compared to typically developed individuals

(Morrison et al.). Autistic participants tend to be seen as less attractive, more awkward, and less warm than typically developing adults (Morrison et al.). College students with ASD reported having difficulties navigating the social environment, which included difficulties in understanding social cues, decreased privacy, lack of familiarity, and isolation (Bailey et al.). What's more, college students with autism found it difficult to connect with others, leading to the trade-offs between efforts to establish relationships and academics (Bailey et al.).

For these reasons, factors supporting or hindering the social participation of young adults with autism are worth knowing. The understanding and adaptations of adolescents' security needs have a positive effect on social participation, such as engaging in home-based activities and exploring the environment at an early time (Krieger et al.). For example, in terms of physically controlled scenarios, the environment with proper control of noise, light, or spaces enhances the social participation of autistic people (Krieger et al.). In contrast, the nonreciprocal peer relationship, exclusion, over-inclusion, and school staff with little knowledge and flexibility can have hindered effects on social participation (Krieger et al.)

Overall, both autistic children and autistic young adults face similar challenges such as restrictive social networks and social isolation. The reasons for their difficulties are identical in some ways, such as their discomfort with social norms and differences between their peers and themselves. However, the life settings for autistic young adults are more than those of children with ASD. Aside from campus, adults also form connections with neighborhoods and colleagues in the workplace. On the other hand, different genders of autistic children have different extents of social interaction success, which is not obvious in young adults with autism. Furthermore, autistic young adults seem to develop more mature coping strategies compared to autistic kids, which is likely due to more life experiences.

Interventions to Enhance the Social Connectedness of Artistic Children and Young Adults

The challenges faced by people with ASD necessitate policy interventions centered on improving social skills. With reasonable instructions and training, their efficiency and efficacy in social engagement can be enhanced.

First of all, researchers applied various interventions to promote the social skills of children diagnosed with autism. In a study, researchers compared the efficacy of two interventions by conducting a randomized trial experiment (Kasari et al.). One group of children received the social skills group approach, emphasizing didactic and clinic-based social skills, and the other group received the engaging approach, which focused on shared activities (Kasari et al.). For kids in the skills group, the social network salience of children who had less closeness and more conflicts with their teachers enhanced significantly, with decreased time of isolation and increased peer engagement (Kasari et al.). For children in the engaged group, however, those with greater closeness to teachers gained more social network salience (Kasari et al.). A study that used music therapy found that the social interaction and communication skills of autistic children improved because music therapy provided social interaction and engagement (Eren).

With the same aim of improving the social skills of autistic children, some research restricts the scenarios in autistic students in a school with a recess setting, in which a regular period of physical activity is given to students in the school schedule (*Keep Recess in Schools*). In a study that used the peer network recess intervention (PNRI), which combined social skills instruction, priming, token systems, peer prompting, reinforcement, and group contingencies, children showed improvements in social communication, peer initiations, and reciprocal communications between peers (McFadden et al.). In another study assessing the effectiveness of an intervention including professional learning for teachers, peer training for participants, activities with key messages, feedback to parents in a recess setting, the time participants engage with peers, the frequency of initiations, the social reciprocity, and the number of friends nominated all increased, despite the huge variations in variables because of the small sample size (Radley et al.).

Previous research also dabbled in how to improve the social competence of autistic children, which is the ability to evaluate social settings and behave appropriately in a given social context (*APA Dictionary of Psychology*). In a study applying theater intervention, the children showed improvements in the memory of faces, social cognition, theories of minds, and communication skills (Corbett et al.). According to another systematic review, peer mediation intervention (PMI), pivotal response training, LEGO therapy, and The Program for the Education and Enrichment of Relational Skills (PEERS) are introduced. PMI, which consists of social skills training with typically developing peers, is responsible for enhancing social competence and social skills (Chang and Locke); pivotal response training, which covers motivation, self-management, initiation of social interactions, and response to multiple cues, can promote communication and social skills (Autism Speaks); LEGO therapy, a program that uses LEGO to support the development of social skills, impels social interactions between peers, problem-solving skills, and friendship-making skills (Seath); and PEERS, which contains instructions in small groups, demonstrations for role-play, social coaching, socialization assignments, and training for caregivers (University at Albany), improves social responsiveness, motivation, awareness, communication, and cognition (Özerk et al.).

In addition to implementing interventions, previous research conducted evaluations of different dimensions of interventions. In a study testing the fidelity, feasibility, and preliminary effects of an intervention combining professional learning for teachers, peer training for participants, activities with key messages, and feedback to parents, there are several groups of beneficiaries (Hodges, Cordier, Joosten, Bourke-Taylor, et al.). Autistic children enhanced their class participation, peer interactions, responsiveness, inattentive behaviors, self-awareness, confidence, and sense of belonging (Hodges, Cordier, Joosten, Bourke-Taylor, et al.). Teachers and typically developing students knew more about autism and became more receptive to people with autism (Hodges, Cordier, Joosten, Bourke-Taylor, et al.). Researchers also gathered the opinions of experts (Hodges, Cordier, Joosten, and Bourke-Taylor). Experts reached a consensus on the application of the family of participation-related constructs and the bidirectional relationship between the preferences of autistic children and their competence (Hodges, Cordier, Joosten, and

Bourke-Taylor). They also recommended offering professional training to teachers, who ought to be aware of the challenges in interventions (Hodges, Cordier, Joosten, and Bourke-Taylor).

Compared to autistic children, some autistic young adults need more comprehensive skills to enhance their social connectedness because of the more complex social settings. Autistic young adults not only maintain their relationships with peers but also develop their social networks with their colleagues, superiors, neighborhoods, college professors, etc. The interventions for different phases of young adults are also different since the support for those who transferring from youth to young adults, young adults who are attending colleges, and young adults who already have their careers differentiates.

Past studies examined improving social connectedness using different intervention programs or assessing the efficacy of different intervention models. A study has shown that the PEERS program can help autistic young adults improve their knowledge of social skills, social skills, social functioning, social responsiveness, and peer interactions and reduce ASD symptoms (Laugeson et al.). In the follow-up analysis, autistic young adults' knowledge of social skills, improvements in ASD symptoms regarding social responsiveness, and peer interactions were maintained, with improvements in social motivation, social cognition, social communication, responsibility, and empathy (Laugeson et al.). In another study that examined the effectiveness of employment-related interventions called Assistive Soft Skills and Employment Training (ASSET) that entails communications, teamwork, networking, enthusiasm and attitude, problem-solving and critical thinking, and professionalism, researchers concluded that participants felt highly satisfied with the usefulness, quality, appropriateness, and relevance of learning materials, with reasonable time and work, clear expectations for participants, easy content for participants to understand, and effective delivery in the intervention (Sung et al.). Improvements were visible in participants' social communication, social awareness, social motivation, social cognition, and skills in career networking (Sung et al.). Previous research also assessed the effect of virtual humans on the social connectedness of young autistic adults. For instance, in a study in which autistic and non-autistic young adults were asked to communicate with the virtual human Kuki, researchers found that autistic participants were more likely to communicate with Kuki and were ready to develop a relationship with it, but Kuki seemed not to fulfill the expectations of the participants, leading to frustration and barriers of affective exchange (Xygkou et al.).

Among the interventions aiming to develop social skills for autistic young adults, some researchers specify the range of autistic young adults in college settings. A previous study implemented a social planning intervention including motivational interest incorporation, social activity choice based on their interest, organizational skills regarding social activity, peer mentor support, and communication and interaction training (Ashbaugh et al.). researchers found that the number of social activities, peer interactions, extracurricular activities, and academic performance increased for all participants (Ashbaugh et al.). Furthermore, all participants enhanced their satisfaction with the social and college experience, the number of social activities, and the availability of campus social activities (Ashbaugh et al.). Another research focusing on a support group model including academic skills, teamwork, interpersonal communication, time and stress

management indicated that participants had higher self-esteem, lower loneliness, and reduced generalized anxiety after the intervention (Hillier et al.). Based on the feedback from participants, researchers identified five main themes, which were stress and anxiety, executive functioning, goal setting, academic resources, and social (Hillier et al.).

Other studies also offered suggestions on how to design and modify the model of interventions. In a systematic review of effective interventions helping autistic young adults get involved in communities, the author pointed out that postsecondary education, the formal education beyond high school such as earning academic credentials and learning occupations (*IES Topics - Postsecondary Education*), and learning opportunities offered by schools can make contributions to young autistic adults (Carter et al.). Effective interventions should pay attention to the cultivation of social and communication skills, set scenarios in which the confidence, skills, and opportunities for social interaction all exist, and be aware of the importance of adults' roles in creating socially supportive environments, thus enhancing the quality of interactions (Carter et al.). Another review study pointed out the limitations of the assessment of interventions in previous studies, revealing that most studies asked for feedback from caregivers instead of the autistic participants, and the samples from most studies were heavily skewed to white males as participants (Monahan et al.). The autistic input in the assessment of interventions should be emphasized (Monahan et al.).

In general, the interventions for autistic young adults have more comprehensive effects on participants than those for autistic children because young autistic adults deal with more complicated social problems in various scenarios. The dimensions of skills included in the interventions for autistic young adults are more extensive and premium, such as career-related skills and time management, whereas knowledge included in interventions for autistic children is more basic, such as participating in social activities and interacting with peers. Among interventions chosen for autistic children and autistic young adults, the PEERS program is suitable for both of them, which is beneficial for extensive use and convenient unification of standards and content.

Analysis of Policies in China

Most policies in China regarding autistic populations are applicable to all autistic people no matter their ages. In general, the support from policies include financial aid, assistance in autism care services, and support to parents in autism families and public knowledge. To be more specific, the screening of early autistic symptoms and enhancements in education are especially for autistic kids while the provision of employment is for autistic young adults.

The first and the most common one is financial assistance. In recent years, China has enhanced the intensity of fund input, which includes enlarging the assisting range from 0-14 years old to 0-17 years old and offering a basic standard of living and insurance (13th Provincial People's Congress). Furthermore, financial investment also increases input on enhancing rehabilitation services. In Fuzhou City, the government increases the standard of financial support for rehabilitation services by offering rent subsidies and discounts on electricity and water (Fuzhou Municipal People's Government).

The second area of policy improvement is mainly on rehabilitation services for autistic people. From a macro perspective, the State Council required impelling the medical service and enhancing its quality, developing the service of rehabilitation aids, strengthening the precaution of disability, increasing the public culture service, and reinforcing the standard of service and industrial management for the disabled. The hierarchical diagnosis and rehabilitation service system are to be established (Nanjing Municipal People's Government). Proposal No. 6622 of the National People's Congress also emphasized enhancing the facilities of autism care services (14th National Congress) and establishing a boarding system for autistic people (13th Provincial People's Congress).

Support for family members of autistic people and public advertisement is the third focus in Chinese policy. For instance, policies asked to construct parent self-help services, develop parent organizations (14th National Congress), organize parent lectures and parents-child campaigns (13th Provincial People's Congress), and advertise the knowledge and policies regarding autistic people on online platforms to increase public understanding (Nanjing Municipal People's Government).

Besides the policies available to all Chinese with autism and their family, a small portion of policies still have specific requirements of age, which can be briefly divided into those applicable to autistic children and those to autistic young adults.

To be more specific on the policies for autistic children, some pay attention to autistic children regarding early screening of autistic symptoms. The Central People's Government announced that for children from 0 to 6 years old, township health centers and community healthcare centers should provide health education for parents, screening of autism symptoms, diagnosis, and interventions (Nanjing Municipal People's Government). In Nanjing province, similar specific policies include early identification, diagnosis, and intervention services for autistic children from maternal and childcare health centers (Nanjing Municipal People's Government).

The second age-specific characteristic of Chinese policies is the focus on the education of autistic children. Some provinces offer training and research opportunities to improve the abilities of teachers of autistic children (Nanjing Municipal People's Government). Moreover, in Shenzhen, educational policies indicate that autistic children can match with a companion teacher at school, and local authorities will enhance the integrated education (Guangdong Provincial People's Government). In Fuzhou, the government increased the provision of degrees in rehabilitation centers by increasing public autism schools and subsidies for institutions helping autistic children (*"Fuzhou Municipal People's Government on Several Measures to Help Autistic Children Rehabilitation (Trial)" Policy Interpretation _ Municipal Policy Interpretation _ Fuzhou Municipal People's Government Portal*).

For autistic young adults, the specified policies for them more aim to help solve their employment issues. According to the policies published by the State Council, principal and municipal authorities and public institutions should increase the employment of disabled people (Guo). In addition, state-owned enterprises and private enterprises should recruit a certain number of disabled people every year (Guo). Policies also improve employment by enhancing laws and regulations, improving employment qualities and entrepreneurial abilities, improving employment service, and protecting employment rights (Ma).

Overall, the Chinese government emphasizes the issues of the autistic population and has put a lot of effort into improving the life qualities of people with autism in recent years. The range and aspects of assistance are extensive, which is helpful for the holistic mitigation of this social issue. However, examining the social connectedness of autistic children and autistic young adults is not enough. Most policies assist autistic people by providing physical support. More importantly, the goal of policy should incorporate the self-sufficiency of people with autism. This process needs help from training institutions and rehabilitation centers, so more effort should be given to the mental construction of autistic people with a moderately decreasing proportion of physical input. The decreased proportion of financial assistance may not necessarily contradict the well-being of autistic families since a great proportion of families may save money for interventions.

To be specific, policymakers can consider testing the effectiveness of the modes of interventions mentioned in the previous pages in several areas as a pilot project, such as PEERS, pivotal response training, theater interventions, etc. If increased social skills and social connectedness are observed, governments can spread and standardize effective interventions to other regions in case the unbalanced development of different provinces may hinder the applications of efficient interventions. It is worth noting that the efficacy should emphasize the feedback of autistic participants instead of only caregivers or experimenters.

In terms of assistance for autistic young adults, the main helping aspect is their employment. Additionally, governments should also offer special training to autistic employees. For example, training can incorporate PEERS and ASSET to not only promote working efficiency but also enhance their social skills and social connectedness with other colleagues in a working environment.

Besides, enhancing public knowledge of autism by organizing lectures or distributing folding papers may not be enough since the number of people receiving them and the efficiency of spreading public knowledge is limited. Elementary schools and middle schools of typically developing students should include presentations of some basic scientific knowledge about autistic people and people with other disorders in their syllables, hence increasing tolerance of school students and reducing prejudice toward autistic people, which decreases the difficulties of autistic people to form connections with typically developing people.

Discussion

Increasing the social connectedness of autistic children and young adults is urgent given the challenges they face. Since public policy can effectively spread compulsory obligations throughout the country, examining the extent of focus on social connectedness in Chinese policies and seeking improvements is the main focus of this literature review. By searching for autism-related policies issued in previous years, this paper explores a clear hierarchy of authorities and treatment systems and a lack of supporting social connectedness in policies, with primary skewness of physical help such as money and facilities. To better emphasize the importance of social connectedness, policymakers can standardize interventions in rehabilitation centers and hospitals in selected areas as a pilot project, incorporating PEERS, pivotal response training, theater interventions, etc. to enhance the social skills of autistic children. For young autistic adults, policies should not only provide them

with more employment but also special training, which may apply PEERS and ASSET to impel their social connectedness and performance in a work setting. Finally, policymakers should consider including knowledge related to autism in elementary schools and high school education to increase the tolerance of autistic populations in society.

There are also some limitations concerning the effectiveness of applied improvement methods for policies. First of all, though interventions mentioned in this literature review all proved to bring enhancements in social connectedness in autistic children and autistic young adults, some studies have sample sizes too small to generalize the conclusion to more autistic people because of the considerable variabilities between participants. Therefore, future studies can further explore the effectiveness of peer network recess intervention (PNRI), the Superheroes Social Skills program, and music interventions on autistic children and interventions including motivational interest incorporation, social activity choice based on their interest, organizational skills regarding social activity, peer mentor support, and communication and interaction training for college students with autism.

In addition, the differences between the current situations in China and Western countries, in which most studies regarding interventions are conducted, cast doubts on the true effectiveness of interventions if they are applied in China. Cultural differences (Triandis and Suh) and economic development, which is related to family income (Duncan et al.), can all affect the development of autistic children and young adults. For this reason, more studies in China should be conducted in order to verify whether the interventions effective in Western countries are still applicable in Chinese society.

Improving policies for children and young adults diagnosed with ASD does not completely equate to the enhancement of their social connectedness, considering the degree of policy implementation. If comprehensive policies are specious instead of pragmatic to the general public, the problem of autistic children and young adults cannot be improved. Future studies should gather data from the general public to assess whether the policies are truly followed and enhance the well-being of autistic people, filling the gap between policy planning and the tangible improvement in the corresponding autistic population.

Works Cited

- 13th Provincial People's Congress. People's Government of Jiangsu Province Public Response to Recommendation No. 6004 of the Third Session of the 13th Provincial People's Congress (Suggestions on Strengthening Assistance to Elderly Autistic Patients). 1 July 2020, https://www.js.gov.cn/art/2020/7/1/art_59167_9454853.html.
- 14th National Congress. China Disabled Persons' Federation - Proposal No. 6622 of the First Session of the 14th National People's Congress and Handling Reply. 14 Dec. 2023, <https://www.cdpcf.org.cn/ztzl/zyzt1/qglhjytafw/2023nlhjytablfwgk/qgrdjyblfw2023/ef6712083da24127b9c278abc9258a4d.htm>.
- American Psychiatric Association. *Anxiety Disorders: DSM-5® Selections*. American Psychiatric Pub, 2015.
- Anderson, Ariana, et al. "Social Network Analysis of Children with Autism Spectrum Disorder: Predictors of Fragmentation and Connectivity in Elementary School Classrooms." *Autism*, vol. 20, no. 6, Aug. 2016, pp. 700–09. DOI.org (Crossref), <https://doi.org/10.1177/1362361315603568>.
- APA Dictionary of Psychology. <https://dictionary.apa.org/>. Accessed 18 Aug. 2024.
- Ashbaugh, Kristen, et al. "Increasing Social Integration for College Students with Autism Spectrum Disorder." *Behavioral Development Bulletin*, vol. 22, no. 1, Apr. 2017, pp. 183–96. DOI.org (Crossref), <https://doi.org/10.1037/bdb0000057>.
- Autism Speaks. Pivotal Response Treatment (PRT) | Autism Speaks. <https://www.autismspeaks.org/pivotal-response-treatment-prt>. Accessed 18 Aug. 2024.
- Bailey, Kathryn M., et al. "The Relationship between Social Experience and Subjective Well-Being in Autistic College Students: A Mixed Methods Study." *Autism*, vol. 24, no. 5, July 2020, pp. 1081–92. DOI.org (Crossref), <https://doi.org/10.1177/1362361319892457>.
- Barker, Erin T., et al. "Trajectories of Emotional Well-Being in Mothers of Adolescents and Adults with Autism." *Developmental Psychology*, vol. 47, no. 2, 2011, pp. 551–61. APA PsycNet, <https://doi.org/10.1037/a0021268>.
- Carter, Erik W., et al. "CONNECTING YOUTH AND YOUNG ADULTS WITH AUTISM SPECTRUM DISORDERS TO COMMUNITY LIFE." *Psychology in the Schools*, vol. 50, no. 9, Nov. 2013, pp. 888–98. DOI.org (Crossref), <https://doi.org/10.1002/pits.21716>.
- Chamberlain, Brandt, et al. "Involvement or Isolation? The Social Networks of Children with Autism in Regular Classrooms." *Journal of Autism and Developmental Disorders*, vol. 37, no. 2, Feb. 2007, pp. 230–42. DOI.org (Crossref), <https://doi.org/10.1007/s10803-006-0164-4>.
- Chan, Dara V., et al. "Beyond Friendship: The Spectrum of Social Participation of Autistic Adults." *Journal of Autism and Developmental Disorders*, vol. 53, no. 1, Jan. 2023, pp. 424–37. DOI.org (Crossref), <https://doi.org/10.1007/s10803-022-05441-1>.
- Chang, Ya-Chih, and Jill Locke. "A Systematic Review of Peer-Mediated Interventions for Children with Autism Spectrum Disorder." *Research in Autism Spectrum Disorders*, vol. 27, July 2016, pp. 1–10. PubMed Central, <https://doi.org/10.1016/j.rasd.2016.03.010>.
- Corbett, Blythe A., et al. "Improvement in Social Competence Using a Randomized Trial of a Theatre Intervention for Children with Autism Spectrum Disorder." *Journal of Autism and*

- Developmental Disorders, vol. 46, no. 2, Feb. 2016, pp. 658–72. DOI.org (Crossref), <https://doi.org/10.1007/s10803-015-2600-9>.
- Cresswell, Lily, et al. “The Experiences of Peer Relationships amongst Autistic Adolescents: A Systematic Review of the Qualitative Evidence.” *Research in Autism Spectrum Disorders*, vol. 61, May 2019, pp. 45–60. DOI.org (Crossref), <https://doi.org/10.1016/j.rasd.2019.01.003>.
- Duncan, Greg J., et al. “Boosting Family Income to Promote Child Development.” *The Future of Children*, vol. 24, no. 1, 2014, pp. 99–120.
- Eren, Bilgehan. “The Use of Music Interventions to Improve Social Skills in Adolescents with Autism Spectrum Disorders in Integrated Group Music Therapy Sessions.” *Procedia - Social and Behavioral Sciences*, vol. 197, July 2015, pp. 207–13. DOI.org (Crossref), <https://doi.org/10.1016/j.sbspro.2015.07.125>.
- Fuzhou Municipal People’s Government. “Fuzhou Municipal People’s Government on Several Measures to Help Autistic Children Rehabilitation (Trial)” Policy Interpretation _ Municipal Policy Interpretation _ Fuzhou Municipal People’s Government Portal. 9 Aug. 2021, https://www.fuzhou.gov.cn/zcjd/bs/202108/t20210809_4158771.htm.
- “Fuzhou Municipal People’s Government on Several Measures to Help Autistic Children Rehabilitation (Trial)” Policy Interpretation _ Municipal Policy Interpretation _ Fuzhou Municipal People’s Government Portal. https://www.fuzhou.gov.cn/zcjd/bs/202108/t20210809_4158771.htm. Accessed 11 Aug. 2024.
- Guangdong Provincial People’s Government. Shenzhen Plans to Implement Lifelong Support for Autistic People Guangdong Provincial People’s Government Portal. 26 Oct. 2023, https://www.gd.gov.cn/zwgk/zdlyxxgkzl/mzxx/content/post_4272330.html.
- Guo, Xiaoting. Notice of The General Office of the State Council on Issuing a Three-Year Action Plan to Promote the Employment of Persons with Disabilities (2022-2024) _ Document of The State Council _ Gov.Cn. 8 Apr. 2022, https://www.gov.cn/zhengce/zhengceku/2022-04/08/content_5684090.htm.
- Haslam, Catherine, et al. *Social Connectedness and Health*. 2015, pp. 1–10. ResearchGate, https://doi.org/10.1007/978-981-287-080-3_46-2.
- Hillier, Ashleigh, et al. “Supporting University Students with Autism Spectrum Disorder.” *Autism*, vol. 22, no. 1, Jan. 2018, pp. 20–28. DOI.org (Crossref), <https://doi.org/10.1177/1362361317699584>.
- Hodges, Amy, Reinie Cordier, Annette Joosten, Helen Bourke-Taylor, et al. “Evaluating the Feasibility, Fidelity, and Preliminary Effectiveness of a School-Based Intervention to Improve the School Participation and Feelings of Connectedness of Elementary School Students on the Autism Spectrum.” *PLOS ONE*, edited by Amanda A. Webster, vol. 17, no. 6, June 2022, p. e0269098. DOI.org (Crossref), <https://doi.org/10.1371/journal.pone.0269098>.
- . “Expert Consensus on the Development of a School-Based Intervention to Improve the School Participation and Connectedness of Elementary Students on the Autism Spectrum: A Delphi Study.” *Focus on Autism and Other Developmental Disabilities*, vol. 37, no. 1, Mar. 2022, pp. 13–23. DOI.org (Crossref), <https://doi.org/10.1177/10883576211030483>.

- Horlin, Chiara, et al. “The Cost of Autism Spectrum Disorders.” PLOS ONE, vol. 9, no. 9, Sept. 2014, p. e106552. PLoS Journals, <https://doi.org/10.1371/journal.pone.0106552>.
- IES Topics - Postsecondary Education. <https://ies.ed.gov/topics/postsecondary.asp>. Accessed 18 Aug. 2024.
- Kasari, Connie, et al. “Children with Autism Spectrum Disorder and Social Skills Groups at School: A Randomized Trial Comparing Intervention Approach and Peer Composition.” *Journal of Child Psychology and Psychiatry*, vol. 57, no. 2, Feb. 2016, pp. 171–79. DOI.org (Crossref), <https://doi.org/10.1111/jcpp.12460>.
- Keep Recess in Schools.
https://www.cdc.gov/healthyschools/physicalactivity/pdf/Recess_Data_Brief_CDC_Logo_FINAL_191106.pdf. Accessed 18 Aug. 2024.
- Krieger, Beate, et al. “Supporting and Hindering Environments for Participation of Adolescents Diagnosed with Autism Spectrum Disorder: A Scoping Review.” PLOS ONE, edited by Amanda A. Webster, vol. 13, no. 8, Aug. 2018, p. e0202071. DOI.org (Crossref), <https://doi.org/10.1371/journal.pone.0202071>.
- Laugeson, Elizabeth A., et al. “A Randomized Controlled Trial to Improve Social Skills in Young Adults with Autism Spectrum Disorder: The UCLA PEERS® Program.” *Journal of Autism and Developmental Disorders*, vol. 45, no. 12, Dec. 2015, pp. 3978–89. DOI.org (Crossref), <https://doi.org/10.1007/s10803-015-2504-8>.
- Locke, Jill, et al. “Social Network Changes Over the School Year Among Elementary School-Aged Children with and Without an Autism Spectrum Disorder.” *School Mental Health*, vol. 5, no. 1, Mar. 2013, pp. 38–47. DOI.org (Crossref), <https://doi.org/10.1007/s12310-012-9092-y>.
- Ma, Juan. Notice of The State Council on the Issuance of the “14th Five-Year Plan” for the Protection and Development of the Disabled _ Documents of The State Council _ Chinese Government Website. 21 July 2021,
https://www.gov.cn/zhengce/zhengceku/2021-07/21/content_5626391.htm.
- McFadden, Brandon, et al. “Social Communication Effects of Peer-Mediated Recess Intervention for Children with Autism.” *Research in Autism Spectrum Disorders*, vol. 8, no. 12, Dec. 2014, pp. 1699–712. DOI.org (Crossref), <https://doi.org/10.1016/j.rasd.2014.08.015>.
- Monahan, Jessica, et al. “Autistic Input in Social Skills Interventions for Young Adults: A Systematic Review of the Literature.” *Review Journal of Autism and Developmental Disorders*, vol. 10, no. 1, Mar. 2023, pp. 1–21. DOI.org (Crossref), <https://doi.org/10.1007/s40489-021-00280-9>.
- Morrison, Kerriane E., et al. “Outcomes of Real-World Social Interaction for Autistic Adults Paired with Autistic Compared to Typically Developing Partners.” *Autism*, vol. 24, no. 5, July 2020, pp. 1067–80. DOI.org (Crossref), <https://doi.org/10.1177/1362361319892701>.
- Nanjing Municipal People’s Government. Reply to Proposal No. 0159 of the First Session of the 15th Municipal CPPCC _ Proposal of the Municipal CPPCC _ Nanjing Municipal People’s Government. 26 July 2023,
https://www.nanjing.gov.cn/xxgkn/jytabljggk/2023njytabl/shizxta/202311/t20231127_4107597.html.

- Orsmond, Gael I., et al. "Social Participation Among Young Adults with an Autism Spectrum Disorder." *Journal of Autism and Developmental Disorders*, vol. 43, no. 11, Nov. 2013, pp. 2710–19. DOI.org (Crossref), <https://doi.org/10.1007/s10803-013-1833-8>.
- Özerk, Guil, et al. "Developing Social Skills and Social Competence in Children with Autism." *International Electronic Journal of Elementary Education*, vol. 13, no. 3, Mar. 2021, pp. 341–63. DOI.org (Crossref), <https://doi.org/10.26822/iejee.2021.195>.
- Petrina, Neysa, et al. "The Nature of Friendship in Children with Autism Spectrum Disorders: A Systematic Review." *Research in Autism Spectrum Disorders*, vol. 8, no. 2, Feb. 2014, pp. 111–26. DOI.org (Crossref), <https://doi.org/10.1016/j.rasd.2013.10.016>.
- Radley, Keith C., et al. "The Effects of a Social Skills Training Package on Social Engagement of Children With Autism Spectrum Disorders in a Generalized Recess Setting." *Focus on Autism and Other Developmental Disabilities*, vol. 29, no. 4, Dec. 2014, pp. 216–29. DOI.org (Crossref), <https://doi.org/10.1177/1088357614525660>.
- Rotheram-Fuller, Erin, et al. "Social Involvement of Children with Autism Spectrum Disorders in Elementary School Classrooms." *Journal of Child Psychology and Psychiatry*, vol. 51, no. 11, Nov. 2010, pp. 1227–34. DOI.org (Crossref), <https://doi.org/10.1111/j.1469-7610.2010.02289.x>.
- Seath, Nicole. "What Is LEGO-Based Therapy? Expert Advice." *Therapy Focus*, 23 Apr. 2020, <https://therapyfocus.org.au/on-the-blog/what-is-lego-therapy/>.
- Sosnowy, Collette, et al. "Setbacks and Successes: How Young Adults on the Autism Spectrum Seek Friendship." *Autism in Adulthood*, vol. 1, no. 1, Mar. 2019, pp. 44–51. DOI.org (Crossref), <https://doi.org/10.1089/aut.2018.0009>.
- Sung, Connie, et al. "Development, Feasibility, and Preliminary Efficacy of an Employment-Related Social Skills Intervention for Young Adults with High-Functioning Autism." *Autism*, vol. 23, no. 6, Aug. 2019, pp. 1542–53. DOI.org (Crossref), <https://doi.org/10.1177/1362361318801345>.
- Triandis, Harry C., and Eunkook M. Suh. "Cultural Influences on Personality." *Annual Review of Psychology*, vol. 53, no. Volume 53, 2002, Feb. 2002, pp. 133–60. www.annualreviews.org, <https://doi.org/10.1146/annurev.psych.53.100901.135200>.
- Tsou, Yung-Ting, et al. "Social Connectedness and Loneliness in School for Autistic and Allistic Children." *Autism*, June 2024, p. 13623613241259932. DOI.org (Crossref), <https://doi.org/10.1177/13623613241259932>.
- University at Albany. The Program for the Education and Enrichment of Relational Skills (PEERS®) | University at Albany. <https://www.albany.edu/autism/programs/program-education-and-enrichment-relational-skills-peersr>. Accessed 18 Aug. 2024.
- Van Asselt-Goverts, A. E., et al. "Do Social Networks Differ? Comparison of the Social Networks of People with Intellectual Disabilities, People with Autism Spectrum Disorders and Other People Living in the Community." *Journal of Autism and Developmental Disorders*, vol. 45, no. 5, May 2015, pp. 1191–203. DOI.org (Crossref), <https://doi.org/10.1007/s10803-014-2279-3>.
- Xygykou, Anna, et al. "Can I Be More Social with a Chatbot?: Social Connectedness Through Interactions of Autistic Adults with a Conversational Virtual Human." *International Journal*

of Human–Computer Interaction, Jan. 2024, pp. 1–18. DOI.org (Crossref),
<https://doi.org/10.1080/10447318.2023.2292880>.

Zeedyk, Sasha M., et al. “Perceived Social Competence and Loneliness Among Young Children with ASD: Child, Parent and Teacher Reports.” *Journal of Autism and Developmental Disorders*, vol. 46, no. 2, Feb. 2016, pp. 436–49. DOI.org (Crossref),
<https://doi.org/10.1007/s10803-015-2575-6>.

Zhang, Junlong, and Yu Luo. Degree Centrality, Betweenness Centrality, and Closeness Centrality in Social Network. Atlantis Press, 2017, pp. 300–03. www.atlantis-press.com,
<https://doi.org/10.2991/msam-17.2017.68>.

MSCs-Derived Exosomes as a Biomaterial-Based Cell-Free Treatment for Spinal Cord Injury

By Davidoula Georgopoulou

Abstract

Spinal cord injury (SCI) arises from an abrupt and traumatic impact on the spine, resulting in fractures or displacement of the vertebrae. It is a profound medical condition with substantial social and economic consequences. SCI has witnessed significant advancements in its treatment through surgical procedures and rehabilitation. Unfortunately, these approaches often fail to achieve complete functional restoration and address secondary pathophysiological changes, which might be the key to treating SCI. Mesenchymal stem cells (MSCs) show promising new treatment options for SCI patients. According to recent studies, their therapeutic effects are likely related to paracrine action, where MSCs release exosomes, important mediators of cell-cell communication. More importantly, when biomaterials are combined with MSC-derived exosomes, the exosome retention is augmented at the designated site and controlled release of exosomes is ensured, enhancing the therapeutic impact on SCI. In this paper, relevant papers' significance and limitations are discussed with their objectivity taken into consideration. Since many studies are focused only on certain hallmarks of the secondary SCI being reversed by MSC-exosomes, a holistic approach would be more beneficial.

Keywords: spinal cord injury, exosomes, mesenchymal-stem cells (MSCs), cell-free therapy, biomaterials, regenerative medicine

Introduction

A spinal cord injury (SCI) is damage to the clusters of cells and nerves that are responsible for transmitting signals between the brain and the rest of the body (1). SCI often arises from an abrupt and traumatic impact on the spinal column, resulting in fractures or displacement of the vertebrae (2). It can lead to either a full or partial absence of voluntary motor control and sensory impairment, while also carrying a high risk for complications. As reported in the 2020 data from the National Center for Spinal Cord Injury Statistics Facts and Figures, there are an estimated 294,000 new cases of SCI annually across the globe, making SCIs an urgent health concern (3). Early surgical decompression, drug treatment for preventing secondary injury, and rehabilitation are the main treatments for SCI (4). The financial burden is also substantial, with lifetime costs exceeding 3 million dollars per patient (5). However, despite this significant expense, SCI patients often suffer from poor quality of life, as the available treatments have numerous limitations and only offer supportive relief for individuals (6). Regenerative medicine strategies based on stem-cell therapy for the treatment of SCI have attracted much attention recently. Stem cells are undifferentiated cells that can self-renew and differentiate into different types of cells once stimulated. This kind of therapy promotes the recovery of SCI mainly by replacing lost or damaged cells, providing nutritional support to neurons, protecting neurons, or improving the microenvironment in the spinal cord to allow the

regeneration of damaged axons (7). Up to this point, various cell types have been used in models to explore treatments for SCI, such as mesenchymal stem cells (MSCs), embryonic stem cells (ESCs), neural stem cells (NSCs), Schwann cells (SCs), and olfactory ensheathing cells (8). Specifically, MSCs are the most commonly used stem cells in animal research and human clinical trials. However, the use of MSCs has some drawbacks, like tumor formation, and their therapeutic effects are more likely to be related to paracrine secretion of exosomes by MSCs, instead of MSC substitution and multidirectional differentiation (9).

Exosomes are nanoparticles that originate both from common cells and stem cells and are responsible for communication between cells. Exosomes can be released into the extracellular environment (a non-living space outside of cells within tissues and organs) that consists of the extracellular matrix (ECM), extracellular vesicles, soluble molecules, and physical structures that support and interact with living cells (10). Exosomes' most common origin is MSCs, as MSCs secrete larger amounts of exosomes compared to other kinds of stem cells (11). The advancement of tissue engineering methods has introduced a new approach to SCI treatment by incorporating exosomes into biomaterial scaffolds, which serve as the physical framework for exosome proliferation and differentiation (12). Here an overview of existing studies using MSC-exosomes to cope with the secondary stage of hallmarks of SCI is presented. This paper also studies a biomaterial-cell-free therapy involving MSC-exosomes and sheds light on its potential challenges and future applications.

Understanding the Pathophysiology of Spinal Cord Injury (SCI)

The pathophysiology of SCI is characterized by primary and secondary injury. The primary is an irreversible process and stems from the application of physical forces like compression, contusion, tearing, or tension (5). Such mechanical insults eventually cause axonal disruption, vascular compromise, and cellular membrane damage, forming the foundation upon which subsequent secondary pathogenic events unfold (13). Secondary SCI is generally more complicated than primary injury (14). It can be divided into acute, intermediate, and chronic stages according to the time after injury and the pathological mechanism. It causes an inflammatory response and neuronal apoptosis at the site of the injury, followed by cavity formation and astrocytic scars, leading to the inhibition of axonal regeneration (15). SCI's variability in symptom severity highly depends on the injury location, extent and impact of injury, and individual-specific variables (14).

Exosomes

Exosomes are membrane-bound extracellular vesicles released by almost all types of cells and are paracrine factors (16). Exosomes harbor a diverse array of molecular constituents encompassing proteins, lipids, and nucleic acids (17), with their composition intricately linked to parent cells (18, 19). The proteins encapsulated within exosomes have diverse functions spanning metabolism, cellular energy pathways, growth regulation, intercellular communication, and cargo transportation (20). Moreover, exosomes derived from different cell types exhibit

distinctive lipid compositions, a divergence that also extends to lipid constituents within cellular compartments of the same source cell population (21). Beyond their protein and lipid cargoes, exosomes harbor nucleic acids encompassing messenger RNA (mRNA), microRNAs (miRNAs), and other noncoding RNAs (ncRNAs). miRNAs are particularly abundant within exosomes, a trait that has garnered escalating attention due to its functional implications (22). Heterogeneity in exosomes is beneficial, enabling their utilization across a range of medical applications.

Mesenchymal Stem Cells (MSCs) as a Source of Exosomes

Traditionally, stem cell-based therapeutic strategies require the direct transplantation of cells into the injured tissue. Among the various cell types, MSCs have emerged as a prominent focus due to their ease of acquisition (23), relative ease of isolation (24), diminished immunogenicity (25), and reduced ethical constraints compared to other stem cell types (26). However, this approach is not without its challenges. Direct transplantation of MSCs carries inherent risks, including the potential for chromosomal abnormalities even in early passages, which may give rise to the development of malignant tumors (27). Additionally, despite exhibiting neuron-like traits after transplantation, comprehensive examination underscores the difficulty in categorizing differentiated MSCs as genuine neurons (28, 29). Moreover, the therapeutic effects of MSCs have been associated with the paracrine activity of exosomes, and thus, recent studies focus on MSC-derived exosomes, rather than MSCs themselves. In contrast to direct cell transplantation, the concept of cell-free treatments has emerged, where therapeutic interventions use elements derived from cells, such as exosomes, without the actual transplantation of the cells (11).

The action of MSCs-Exosomes in Secondary Spinal Cord Injury

Since primary SCI injury is an irreversible process (5), focusing on reversing the negative consequences of secondary injury might be the key to treating SCI. There are several symptoms that fall under the category of secondary injury (14, 30). I will provide an overview of existing studies using MSC-exosomes to cope with the hallmarks of secondary SCI.

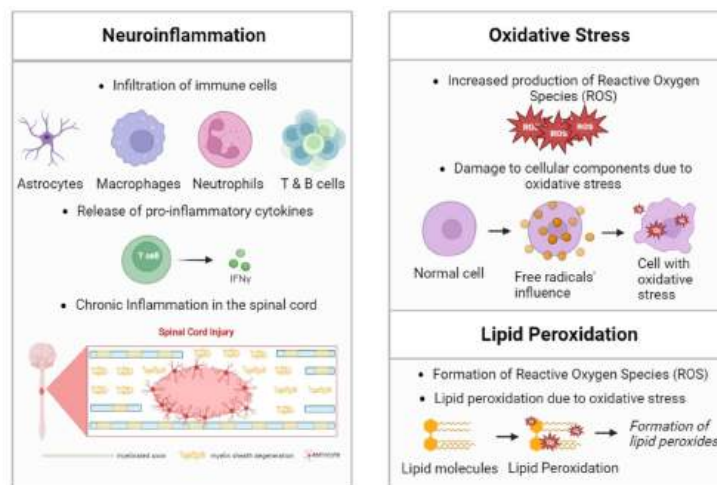


Figure 1: Neuroinflammation, Oxidative Stress, Lipid Peroxidation. Created with biorender.com.

Neuroinflammation

One of the main hallmarks is neuroinflammation (Figure 1). Neuroinflammation refers to the intricate immune response within the CNS, characterized by the activation of various immune cells, such as microglia and astrocytes. Pro-inflammatory cytokines, chemokines, and other immune mediators are released in response to injury. While it aims to counteract threats and facilitate tissue repair, excessive or prolonged neuroinflammation can lead to neuronal damage and cell death (31, 32).

A pivotal mediator in this process is the NLRP3 inflammasome, a critical protein complex, which notably contributes to secondary SCI. Suppression of NLRP3 inflammasome activation enhances functional recovery in rat SCI models (33, 34). Huang et al. found that exosomes derived from MSCs originating in epidural fat (EFMSCs) promote neural recovery and reduce the extent of injury. Systemic administration of EFMSC-exosomes dampens NLRP3 inflammasome activation and reduces inflammatory cytokine expression (35).

Moreover, the balance between pro-inflammatory cytokines like IL-6, interleukin (IL)-1 β , tumor necrosis factor (TNF)- α , and anti-inflammatory factors significantly influences the functional recovery of SCI patients (36). The pro-inflammatory and anti-inflammatory milieu's composition closely associates with SCI prognosis, and thus, restraining the establishment of a pro-inflammatory environment is a promising therapeutic approach. Romanelli et al. illustrated that exosomes originating from MSCs found in the human umbilical cord (hUCMSC-exosomes) engage directly with activated microglia in vitro, thereby suppressing the expression of pro-inflammatory cytokines in the context of secondary injury. In rat SCI models, intravenous hUCMSC-exosome administration suppressed IL-1 β and IL-6 expression and mitigated scar formation, thereby aiding motor function recovery (16). Additionally, Sun et al. revealed that hUCMSC-derived exosomes reduced levels of pro-inflammatory cytokines like interferon- γ , TNF- α , IL-6, and elevated anti-inflammatory cytokines IL-4 and IL-10 (37).

Lipid Peroxidation and Oxidative Stress

Activated microglia and macrophages are also called macrophage/microglia because they are not easily distinguished (38). These cells are heterogeneous and have been categorized into M1 and M2 types. M1 macrophages are responsible for generating pro-inflammatory cytokines, reactive oxygen species (ROS), and nitric oxide, which encourage inflammation and injury (39, 40). ROS are highly reactive molecules containing oxygen that can react with lipids (fats) in cell membranes and cause damage to the membrane structure. This process is called lipid peroxidation and it can also be a consequence of oxidative stress, which occurs when there is an imbalance between the production of ROS and the body's ability to counteract them with antioxidants (Figure 1) (30). The nitric oxide that M1 macrophages produce may induce oxidative stress and thus, further impact the survival of cells in the injured spinal cord tissue. Although M2 macrophages have an anti-inflammatory function, after SCI, macrophages switch from the M2 phenotype to the M1 phenotype (37). Thus, macrophage/microglia polarization is a promising therapeutic avenue. Exosomes released by MSCs promote macrophage polarization as

a treatment approach for SCI. Lankford and al. revealed that when MSC-derived exosomes are administered intravenously, they rapidly target the injured spinal cord, as opposed to the uninjured spinal cord, and selectively bind to M2 macrophages, underscoring the potential of M2 macrophages in alleviating SCI. Sun et al. found that human umbilical cord mesenchymal stem cells (hucMSC) also triggered polarization from M1 to an M2 phenotype (41).

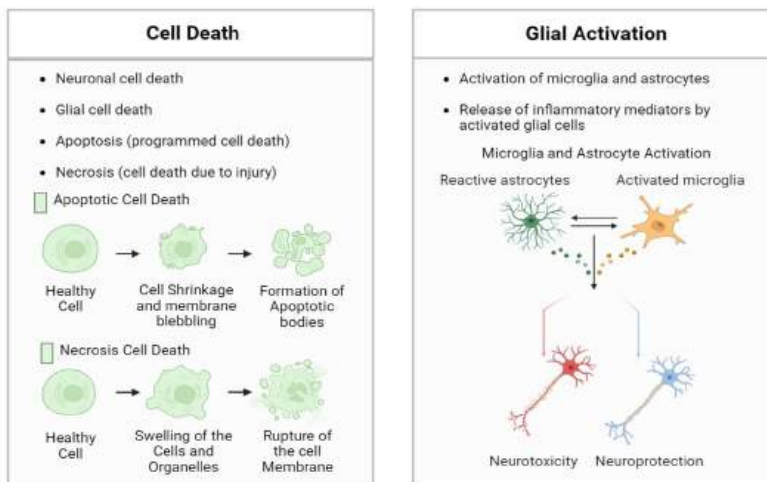


Figure 2: Cell Death, Cell Activation. Created with biorender.com.

Glial activation

Astrocytes either hinder or facilitate CNS recovery (Figure 2). In 2017, Liddelow et al. spotted the emergence of two distinct reactive astrocyte phenotypes, A1 and A2 astrocytes, in response to neuroinflammation and ischemia. The safeguarding role of A2 astrocytes manifests through the elevation of neurotrophic factor expression, contributing to protective mechanisms. Conversely, the formation of A1 astrocytes post-SCI lose the ability to promote neuronal survival, synapse formation, growth, and phagocytosis (an immune system response that destroys harmful substances) and can also lead to neuronal and oligodendrocyte death (42). Evidently, targeting A1 astrocytes emerges as a prospective avenue for SCI treatment. A study in 2018 reported that BMSC-derived exosome therapy reduces SCI-induced A1 astrocytes by effectively curbing the nuclear translocation of NF- κ B p65 (43). Moreover, another study reported the efficacy of exosomes derived from cells BMSCs in fostering functional recovery following SCI by restraining the activation of A1 neurotoxic reactive astrocytes (44).

Cell Death

Apoptosis, which is a form of programmed cell death, is a tightly regulated mechanism that serves to uphold tissue homeostasis and eliminate cells that are damaged or unnecessary within an organism. However, it contributes to the loss of neurons, oligodendrocytes, and other vital cells in the injured spinal cord (Figure 2). Therefore, it is necessary to suppress it (45). Li et al. demonstrated that BMSC-exosomes can promote motor function recovery via the Wnt/beta-catenin signaling pathway that can exhibit antiapoptotic effects. BMSC-exosomes also

significantly downregulate apoptosis-related markers, like caspase-3 and caspase-9 in rats after SCI and increase the level of the anti-apoptotic proteins, like Bcl-2 (46). In addition, BMSC-exosomes can reduce neuronal apoptosis by promoting neuronal autophagy. MSC-exosomes can activate autophagy and increase the number of autophagosomes (47). Another study revealed that exosomes obtained from cultured MSCs' supernatant possess the ability to diminish the expression of PTEN (tumor suppressor gene), programmed cell death factor 4 (PDCD4), and FasL (protein involved in apoptosis). This action contributes to the prevention of neuronal apoptosis, showcasing a protective influence (48). Additionally, targeting A1 astrocytes is another therapeutic option that copes with neuronal and oligodendrocyte death(42).

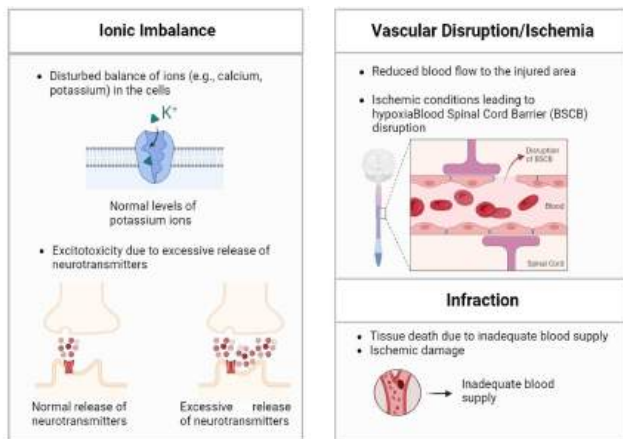


Figure 3: Ionic Imbalance, Vascular Disruption/Ischemia, Infraction. Created with biorender.com.

Ionic Imbalance

In the context of SCI, ionic imbalance refers to the disruption of the normal levels of ions (Figure 3), such as calcium and potassium within cells, which can lead to neural dysfunction and hinder recovery processes (30). Furthermore, excitotoxicity is a process where neurons become overstimulated and damaged due to excessive release of neurotransmitters, which are chemicals that transmit signals between nerve cells. Especially glutamate excitotoxicity results in an influx of calcium ions into neurons, triggering a series of events that result in cell damage and death (30). MSC-exosomes have been found to have neuroprotective properties, as they can contain factors that counteract excitotoxicity, such as anti-inflammatory molecules and growth factors. Zhuang et al. reported that MSCs-exosomes were found to alleviate glutamate-mediated excitotoxicity in the context of traumatic brain injury, another medical condition in the CNS. The study demonstrated that MSC-exosomes reduced neuronal apoptosis induced by glutamate, both in vitro and in vivo (49).

Vascular Infraction/Ischemia, Infraction

Angiogenesis, the formation of new blood vessels, plays a crucial role in repairing SCI. Injury-induced blood vessel loss and disruption of the blood-spinal cord barrier (BSCB) can result in inadequate blood supply (ischemia) and inflammation (Figure 3). Enhancing

angiogenesis helps restore blood flow, oxygen, and nutrients to the injured area and facilitates the healing process (50).

Recent investigations illustrate how MSC-exosomes amplify angiogenesis. For instance, vascular endothelial cells are essential constituents of blood vessel walls and MSC-exosomes have been shown to enhance the expression of crucial genes related to angiogenesis in human umbilical vein endothelial cells (HUVECs). Particularly, the involvement of vascular endothelial growth factor (VEGF) is demonstrated in the process of MSC-exosome-induced proliferation and migration of HUVECs (51). Following SCI, the spinal cord tissue frequently experiences reduced oxygen supply (hypoxia). Hypoxia-treated MSC-exosomes exhibit enhanced uptake by HUVECs compared to untreated MSC-exosomes. Additionally, these hypoxia-treated exosomes stimulate VEGF expression in HUVECs through the activation of the protein kinase A signaling pathway (52).

The BSCB is a specialized protective barrier in the CNS, composed of tightly interconnected endothelial cells, pericytes, and glial cells. BSCB tightly regulates the passage of substances between the blood circulation and the spinal cord tissue. Its main purpose is to maintain a stable microenvironment while preventing the entry of harmful molecules and immune cells that could cause inflammation. Disruption of the BSCB occurs in SCI, allowing inflammatory cells and molecules to infiltrate the spinal cord tissue, contributing to neuroinflammation and secondary injury processes (53).

Pericytes are integral components of the neurovascular unit and preserve the structural integrity and barrier functions of blood vessels. A research study showcased that the introduction of extracellular vesicles derived from bone mesenchymal stem cells (BMSC-EV) not only improved the survival and regeneration of neurons but also enhanced the integrity of the blood-spinal cord barrier (BSCB). They also found that exosomes reduced pericyte migration, with a consequent decrease in the permeability of the BSCB, meaning it became more stable (54). Matsushita et al. also suggest that intravenously delivered BMSC-exosomes have important effects on reducing BSCB leakage (55).

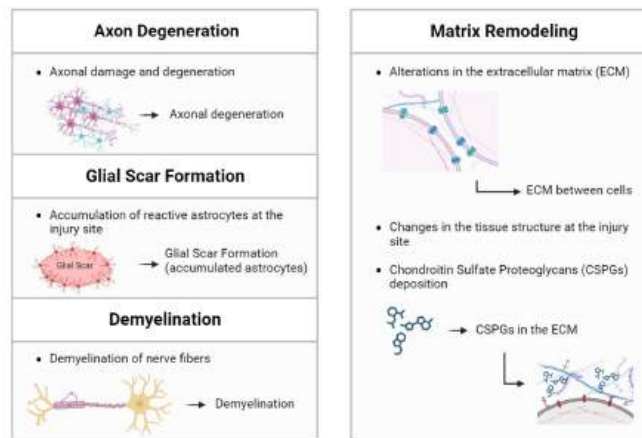


Figure 4: Axon Degeneration, Demyelination, Matrix Remodeling, Glial Scar Formation. Created with biorender.com.

Axon Degeneration, Demyelination, Matrix Remodeling, Glial Scar Formation

Promoting the regeneration of damaged axons or facilitating the growth of unaffected axons at the site of injury holds crucial importance in the restoration of spinal cord function following SCI (Figure 4). The impediment to axonal regeneration primarily hinges on the generation of chondroitin sulfate proteoglycans (CSPGs), complex molecules that are prominently present in the ECM of the injured spinal cord. These CSPGs create an inhibitory environment for axonal growth by forming a physical and biochemical barrier that restrains the extension of regenerating axons through the injury site. The accumulation of CSPGs contributes to the formation of a scar tissue matrix, known as the glial scar. The formation of the glial scar post-SCI involves the deposition of ECM by reactive astrocytes, which acts as a formidable obstacle to axonal regrowth (56). Liu et al. documented that on day 28 post-SCI, the group treated with BMSC-exosomes exhibited remarkable improvements compared to the control group. Specifically, the lesion size decreased by almost 60%, glial scar formation was reduced by nearly 75%, average blood vessel density increased by nearly 60%, and axonal regeneration saw an almost 80% increase (44). Furthermore, Lee et al. fabricated macrophage membrane-fused exosome-mimetic nanovesicles (MF-NVs) from macrophage membrane-fused umbilical cord blood-derived MSCs (MF-MSCs) to assess the therapeutic potential of modified MSC-exosomes. They confirmed that inside the glial scar, the ECM is notably abundant in CSPGs within the no-treatment group on the 28th day after injury. In contrast, the injection of MF-NV led to a noteworthy reduction in CSPGs expression at the lesion site 28 days post-injury, presenting a marked contrast to the no-treatment group (57).

Biomaterials in Secondary Spinal Cord Injury

Biomaterials hold substantial potential in advancing the field of spinal cord injury repair. Diverse categories of biomaterials have found application within tissue engineering, primarily originating from natural and synthetic polymers (58).

Natural biomaterials encompass a diverse array of substances, including chitosan, hyaluronic acid, fibrin, and collagen (59). These natural biomaterials stand out due to their remarkable biocompatibility, minimal immunogenicity, and non-toxic degradation pathways (59). Synthetic biomaterials are represented by polylactic acid (PLA), polyethylene glycol (PEG), polylactic acid-hydroxyacetic acid copolymer (PLGA), polyvinyl alcohol (PVA), polyacrylamide (PAM), and polymethyl methacrylate (PMMA). They exhibit compelling attributes, like robust mechanical properties, customizable structures, and limited immunogenicity and the porosity, stiffness, and degradation rates of synthetic biomaterials can be tailored to align precisely with the specific requisites of distinct tissue types (60, 58). Hydrogels are natural, synthetic, or hybrid biomaterials and present an ideal scaffold for addressing spinal cord contusions. The irregular contours of contused spinal cords can be suitably accommodated through hydrogel injection, while the capacity to emulate the mechanical traits of natural spinal cord tissue remains noteworthy (61).

Action of Biomaterial-Based Cell-Free Treatment using MSCs-Derived Exosomes

The process of integrating MSCs-derived exosomes with biomaterials initiates with the isolation of MSCs from suitable sources, followed by their cultivation and expansion in controlled laboratory conditions. Once an adequate population is obtained, they are prompted to release exosomes. Simultaneously, biomaterials are engineered from either synthetic and/or natural components and the isolated MSC-exosomes are loaded onto the biomaterials. The goal is to secure their attachment to the biomaterials, ensuring their controlled release. In a surgical procedure tailored to the specific injury's location and severity, the biomaterials combined with MSC-exosomes are strategically implanted at the SCI site. Over time, as the biomaterials degrade or dissolve, the exosomes are released. Their interaction within the local microenvironment initiates a cascade of regenerative processes and they counteract the symptoms of secondary SCI (58).

When biomaterials are combined with MSC-exosomes the therapeutic impact on SCI is enhanced due to certain characteristics of the chosen biomaterials. Firstly, biomaterials can retain exosomes at the injury site while conserving their intrinsic properties and structural attributes. Secondly, the scaffolds facilitate the gradual exosome release into the ECM over an extended duration. The scaffolds can also establish a binding affinity with the injured tissue, thereby promoting neighboring cell migration into the scaffold, a pivotal step in orchestrating a conducive microenvironment for tissue repair and regeneration (12, 62).

Although research is in the exploratory stage of this methodology, there are numerous studies that show promising results for this therapeutic paradigm. For instance, MSC-derived exosomes immobilized in a peptide-modified adhesive hydrogel (Exo-pGel) were found to be an effective treatment of CNS diseases based on exosome implantation. The localized application of Exo-pGel introduced an exosome-embedded ECM directly to the injured neural tissue. Thus, the exosomes interacted with the damaged area, demonstrating notable retention and controlled release within the host neural tissues, while also mitigating inflammation and oxidative stress (63). Another study created an injectable adhesive hydrogel called F127-polycitrate-polyethyleneimine hydrogel (FE). This hydrogel was designed to be anti-inflammatory and capable of releasing extracellular vesicles over an extended period of time. Ultimately, it also decreased fibrotic scar formation and inflammation, as well as enhanced the recovery of motor function after SCI (64). When a 3D gelatin methacrylate hydrogel (GelMA) was employed as a delivery platform for transplanted exosomes, it again confirmed that the GelMA hydrogel effectively facilitated the retention of exomes (65). To address the inefficient delivery of antioxidants, a novel hyaluronic acid (HA) hydrogel, augmented with the antioxidant compound 2,2,6,6-tetramethylpiperidinyloxy (TEMPO) was developed. The HA hydrogel was imbued with exosomes derived from human placental amniotic membrane-MSCs. At the 28-day mark post-rat-transplantation, motor function recovery was facilitated. Notably, according to magnetic resonance imaging, there was a substantial reduction in the fracture area within the spinal cord (66). Electroconductive hydrogels stand as compelling contenders for expediting the repair of SCI due to their alignment with the electrical and mechanical attributes

of neural tissue. Nevertheless, the introduction of electroconductive hydrogels may potentially exacerbate inflammation and impede their therapeutic efficacy. Since BMSC-exosomes harness the immunomodulatory and tissue regenerative potentials, neural tissue-like electroconductive hydrogels infused with BMSC-exosomes. Overall, they managed to exert control over microglial M2 polarization and suppress astrocyte differentiation while promoting axon outgrowth (67). By harnessing the inherent properties of MSCs and their exosomes, coupled with the supportive framework of biomaterials, this strategy holds the potential to address the challenges posed by secondary SCI.

Limitations and Future Directions

Utilizing MSC-derived exosomes as a biomaterial-based, cell-free treatment for SCI raises several considerations. Firstly, there are challenges associated with transitioning from preclinical to clinical trials, such as the timing of graft transplantation, purification standards, and safety measures. To ensure successful clinical translation, it's imperative to establish international standards (68, 69). Determining the ideal source of MSC-derived exosomes, such as BMSC or hucMSC-exosomes, is also important for the effectiveness and consistency of treatment outcomes (70). In addition, a current limitation is the development of innovative preservation methods, such as lyophilization, for preserving the functionality of exosomes. This advancement can enhance their storage capabilities and make them more readily available for therapeutic use (69).

Furthermore, it's essential to gain a comprehensive understanding of the specific components of exosomes that contribute to the treatment of SCI. This knowledge should also encompass the identification of potentially harmful constituents (70, 71). Another current limitation is determining the long-term effects of MSC-exosome treatment, including the sustainability of therapeutic benefits and any potential risks over time. This assessment is essential for evaluating treatment durability and safety, reducing the need for repeated interventions (71). Lastly, understanding the intricate interactions between MSC-exosomes and the host spinal tissue, including potential immune responses and integration into the neural environment, is pivotal for ensuring the safe delivery of treatment. Clinical heterogeneity, characterized by variations in patient profiles, types of injuries, and SCI stages, adds complexity. As a result, personalized treatment protocols should be developed to optimize clinical outcomes. To address these limitations and advance the field of MSC-derived exosome therapy for SCI, one promising avenue of research involves the development of standardized methods for MSC-exosome separation. This includes the optimization of techniques like ultracentrifugation to ensure consistent and reliable results (68). Furthermore, it is important to investigate and standardize the optimal dosing, frequency, and administration of MSC-exosomes to achieve long-term therapeutic effects. Understanding the relationship between these factors and treatment outcomes is essential for successful clinical translation (72). Another future direction is expanding preclinical studies to incorporate larger animals with longer lifespans and diverse injury models (71).

Introducing personalized medicine and tailoring MSC-exosome treatments based on individual patient characteristics can significantly improve treatment response and patient outcomes, promoting a more patient-centered approach to SCI therapy. Lastly, leveraging the immunomodulatory properties of MSC-exosomes to modulate immune responses and facilitate tissue integration represents a promising approach to enhance treatment efficacy while minimizing adverse reactions.

Conclusion

The integration of MSC-derived exosomes with biomaterials presents a promising approach for addressing the hallmarks of secondary SCI, which might be the key to treating SCI. These cell-free therapies offer opportunities to enhance treatment precision and outcomes, as well as minimize risks. While limitations like sourcing, administration, and long-term effects must be addressed, this field holds the potential for safer, more effective, and patient-centered treatments, leading to successful clinical translation.

Acknowledgments

I would like to thank my parents and my mentors, Drs. Birol Ay and Cigdem Sahin, for their constant support that allowed me to explore my passion for science and research. Additionally, I would like to acknowledge the Cambridge Center for International Research for their assistance and resources.

Works Cited:

1. Ahuja, C. S., Wilson, J. R., Nori, S., Kotter, M. R. N., Druschel, C., Curt, A., & Fehlings, M. G. (2017). Traumatic spinal cord injury. *Nature Reviews Disease Primers*, 3(1). <https://doi.org/10.1038/nrdp.2017.18>
2. O'Shea, T. M., Burda, J. E., & Sofroniew, M. V. (2017). Cell biology of spinal cord injury and repair. *Journal of Clinical Investigation*, 127(9), 3259–3270. <https://doi.org/10.1172/jci90608>
3. Shen, H., Fan, C., You, Z., Xiao, Z., Zhao, Y., & Dai, J. (2021). Advances in Biomaterial-Based Spinal Cord Injury Repair. *32(13)*, 2110628–2110628. <https://doi.org/10.1002/adfm.202110628>
4. Wang, T. Y., Park, C., Zhang, H., Rahimpour, S., Murphy, K. R., Goodwin, C. R., Karikari, I. O., Than, K. D., Shaffrey, C. I., Foster, N., & Abd-El-Barr, M. M. (2021). Management of Acute Traumatic Spinal Cord Injury: A Review of the Literature. *Frontiers in Surgery*, 8. <https://doi.org/10.3389/fsurg.2021.698736>
5. Katoh, H., Yokota, K., & Fehlings, M. G. (2019). Regeneration of Spinal Cord Connectivity Through Stem Cell Transplantation and Biomaterial Scaffolds. *Frontiers in Cellular Neuroscience*, 13. <https://doi.org/10.3389/fncel.2019.00248>
6. Tian, T., Zhang, S., & Yang, M. (2023). Recent progress and challenges in the treatment of spinal cord injury. *Protein & Cell*. <https://doi.org/10.1093/procel/pwad003>
7. Sahni, V., & Kessler, J. A. (2010). Stem cell therapies for spinal cord injury. *Nature Reviews Neurology*, 6(7), 363–372. <https://doi.org/10.1038/nrneurol.2010.73>
8. Huang, L., Fu, C., Xiong, F., He, C., & Wei, Q. (2021). Stem Cell Therapy for Spinal Cord Injury. *Cell Transplantation*, 30. <https://doi.org/10.1177/0963689721989266>
9. Janockova, J., Slovinska, L., Harvanova, D., Spakova, T., & Rosocha, J. (2021). New therapeutic approaches of mesenchymal stem cells-derived exosomes. *Journal of Biomedical Science*, 28(1). <https://doi.org/10.1186/s12929-021-00736-4>
10. Kalluri, R., & LeBleu, V. S. (2020). The biology, function, and Biomedical Applications of Exosomes. *Science*, 367(6478). <https://doi.org/10.1126/science.aau6977>
11. Rao, D., Huang, D., Sang, C., Zhong, T., Zhang, Z., & Tang, Z. (2022). Advances in Mesenchymal Stem Cell-Derived Exosomes as Drug Delivery Vehicles. *Frontiers in Bioengineering and Biotechnology*, 9. <https://doi.org/10.3389/fbioe.2021.797359>
12. Wang, B., Liu, S., & Xie, Y.-Y. (2019). Role and prospects of regenerative biomaterials in the repair of spinal cord injury. *Neural Regeneration Research*, 14(8), 1352. <https://doi.org/10.4103/1673-5374.253512>
13. Rowland, J. W., Hawryluk, G. W. J., Kwon, B., & Fehlings, M. G. (2008). Current status of acute spinal cord injury pathophysiology and emerging therapies: promise on the horizon. *Neurosurgical Focus*, 25(5), E2. <https://doi.org/10.3171/foc.2008.25.11.e2>
14. Alizadeh, A., Dyck, S. M., & Karimi-Abdolrezaee, S. (2019). Traumatic Spinal Cord Injury: An Overview of Pathophysiology, Models and Acute Injury Mechanisms. *Frontiers in Neurology*, 10(282). <https://doi.org/10.3389/fneur.2019.00282>

15. Wertheim, L., Edri, R., Goldshmit, Y., Kagan, T., Noor, N., Ruban, A., Shapira, A., Gat-Viks, I., Assaf, Y., & Dvir, T. (2022). Regenerating the Injured Spinal Cord at the Chronic Phase by Engineered iPSCs-Derived 3D Neuronal Networks. *Advanced Science*, 9(11), 2105694. <https://doi.org/10.1002/advs.202105694>
16. Romanelli, P., Bieler, L., Scharler, C., Pachler, K., Kreutzer, C., Zaunmair, P., Jakubecova, D., Mrowetz, H., Benedetti, B., Rivera, F. J., Aigner, L., Rohde, E., Gimona, M., Strunk, D., & Couillard-Despres, S. (2019). Extracellular Vesicles Can Deliver Anti-inflammatory and Anti-scarring Activities of Mesenchymal Stromal Cells After Spinal Cord Injury. *Frontiers in Neurology*, 10. <https://doi.org/10.3389/fneur.2019.01225>
17. Colombo, M., Raposo, G., & Théry, C. (2014). Biogenesis, Secretion, and Intercellular Interactions of Exosomes and Other Extracellular Vesicles. *Annual Review of Cell and Developmental Biology*, 30(1), 255–289. <https://doi.org/10.1146/annurev-cellbio-101512-122326>
18. Muthu, S., Bapat, A., Jain, R., Jeyaraman, N., & Jeyaraman, M. (2021). Exosomal therapy—a new frontier in regenerative medicine. *Stem Cell Investigation*, 8, 7–7. <https://doi.org/10.21037/sci-2020-037>
19. Mathivanan, S., Ji, H., & Simpson, R. J. (2010). Exosomes: Extracellular organelles important in intercellular communication. *Journal of Proteomics*, 73(10), 1907–1920. <https://doi.org/10.1016/j.jprot.2010.06.006>
20. Yang, C., Guo, W. ., Zhang, W. ., Bian, J., Yang, J. ., Zhou, Q. ., Chen, M. ., Peng, W., Qi, T., Wang, C. ., & Liu, C. . (2017). Comprehensive proteomics analysis of exosomes derived from human seminal plasma. *Andrology*, 5(5), 1007–1015. <https://doi.org/10.1111/andr.12412>
21. Viet Anh Dang, Kishore Kumar Jella, Ragheb, R., Denslow, N. D., & Alli, A. A. (2017). Lipidomic and proteomic analysis of exosomes from mouse cortical collecting duct cells. *The FASEB Journal*, 31(12), 5399–5408. <https://doi.org/10.1096/fj.201700417r>
22. Haraszti, R. A., Didiot, M.-C., Sapp, E., Leszyk, J., Shaffer, S. A., Rockwell, H. E., Gao, F., Narain, N. R., DiFiglia, M., Kiebish, M. A., Aronin, N., & Khvorova, A. (2016). High-resolution proteomic and lipidomic analysis of exosomes and microvesicles from different cell sources. *Journal of Extracellular Vesicles*, 5(1), 32570. <https://doi.org/10.3402/jev.v5.32570>
23. Stewart, M. C., & Stewart, A. A. (2011). Mesenchymal Stem Cells: Characteristics, Sources, and Mechanisms of Action. *Veterinary Clinics of North America: Equine Practice*, 27(2), 243–261. <https://doi.org/10.1016/j.cveq.2011.06.004>
24. Zhang, W., Zhang, F., Shi, H., Tan, R., Han, S., Ye, G., Pan, S., Sun, F., & Liu, X. (2014). Comparisons of Rabbit Bone Marrow Mesenchymal Stem Cell Isolation and Culture Methods In Vitro. *PLoS ONE*, 9(2), e88794. <https://doi.org/10.1371/journal.pone.0088794>
25. GROH, M., MAITRA, B., SZEKELY, E., & KOC, O. (2005). Human mesenchymal stem cells require monocyte-mediated activation to suppress alloreactive T cells. *Experimental Hematology*, 33(8), 928–934. <https://doi.org/10.1016/j.exphem.2005.05.002>
26. Volarevic, V., Markovic, B. S., Gazdic, M., Volarevic, A., Jovicic, N., Arsenijevic, N., Armstrong, L., Djonov, V., Lako, M., & Stojkovic, M. (2018). Ethical and Safety Issues

- of Stem Cell-Based Therapy. *International Journal of Medical Sciences*, 15(1), 36–45.
<https://doi.org/10.7150/ijms.21666>
27. Jeong, J.-O., Han, J. W., Kim, J.-M., Cho, H.-J., Park, C., Lee, N., Kim, D.-W., & Yoon, Y. (2011). Malignant Tumor Formation after Transplantation of Short-Term Cultured Bone Marrow Mesenchymal Stem Cells in Experimental Myocardial Infarction and Diabetic Neuropathy. *Circulation Research*, 108(11), 1340–1347. <https://doi.org/10.1161/CIRCRESAHA.110.239848>
 28. Vismara, I., Papa, S., Rossi, F., Forloni, G., & Veglianesi, P. (2017). Current Options for Cell Therapy in Spinal Cord Injury. *Trends in Molecular Medicine*, 23(9), 831–849.
<https://doi.org/10.1016/j.molmed.2017.07.005>
 29. Corcione, A. (2006). Human mesenchymal stem cells modulate B-cell functions. *Blood*, 107(1), 367–372. <https://doi.org/10.1182/blood-2005-07-2657>
 30. Anjum, A., Yazid, M. D., Fauzi Daud, M., Idris, J., Ng, A. M. H., Selvi Naicker, A., Ismail, O. H. R., Athi Kumar, R. K., & Lokanathan, Y. (2020). Spinal Cord Injury: Pathophysiology, Multimolecular Interactions, and Underlying Recovery Mechanisms. *International Journal of Molecular Sciences*, 21(20), 7533. <https://doi.org/10.3390/ijms21207533>
 31. Pang, Q.-M., Chen, S.-Y., Xu, Q.-J., Fu, S.-P., Yang, Y.-C., Zou, W.-H., Zhang, M., Liu, J., Wan, W.-H., Peng, J.-C., & Zhang, T. (2021). Neuroinflammation and Scarring After Spinal Cord Injury: Therapeutic Roles of MSCs on Inflammation and Glial Scar. *Frontiers in Immunology*, 12. <https://doi.org/10.3389/fimmu.2021.751021>
 32. Kempuraj, D., Thangavel, R., Natteru, P. A., Selvakumar, G. P., Saeed, D., Zahoor, H., Zaheer, S., Iyer, S. S., & Zaheer, A. (2016). Neuroinflammation Induces Neurodegeneration. *Journal of Neurology, Neurosurgery and Spine*, 1(1), 1003.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5260818/#:~:text=This%20chronic%20inflammation%20contributes%20to>
 33. Jiang, W., Li, M., He, F., Zhou, S., & Zhu, L. (2017). Targeting the NLRP3 inflammasome to attenuate spinal cord injury in mice. *Journal of Neuroinflammation*, 14(1).
<https://doi.org/10.1186/s12974-017-0980-9>
 34. Jiang, W., Huang, Y., He, F., Jia, L., Li, M., Sun, T., Ren, W., Hou, J., & Zhu, L. (2016). Dopamine D1 Receptor Agonist A-68930 Inhibits NLRP3 Inflammasome Activation, Controls Inflammation, and Alleviates Histopathology in a Rat Model of Spinal Cord Injury. *Spine*, 41(6), E330–E334. <https://doi.org/10.1097/brs.0000000000001287>
 35. Huang, J.-H., Fu, C., Xu, Y., Yin, X.-M., Cao, Y., & Lin, F.-Y. (2020). Extracellular Vesicles Derived from Epidural Fat-Mesenchymal Stem Cells Attenuate NLRP3 Inflammasome Activation and Improve Functional Recovery After Spinal Cord Injury. *Neurochemical Research*, 45(4), 760–771. <https://doi.org/10.1007/s11064-019-02950-x>
 36. Zhang, J.-M., & An, J. (2007). Cytokines, Inflammation, and Pain. *International Anesthesiology Clinics*, 45(2), 27–37. <https://doi.org/10.1097/aia.0b013e318034194e>
 37. Sun, G., Li, G., Li, D., Huang, W., Zhang, R., Zhang, H., Duan, Y., & Wang, B. (2018). hucMSC derived exosomes promote functional recovery in spinal cord injury mice via attenuating inflammation. *Materials Science and Engineering: C*, 89, 194–204.
<https://doi.org/10.1016/j.msec.2018.04.006>

38. Kigerl, K. A., Gensel, J. C., Ankeny, D. P., Alexander, J. K., Donnelly, D. J., & Popovich, P. G. (2009). Identification of Two Distinct Macrophage Subsets with Divergent Effects Causing either Neurotoxicity or Regeneration in the Injured Mouse Spinal Cord. *Journal of Neuroscience*, 29(43), 13435–13444. <https://doi.org/10.1523/jneurosci.3257-09.2009>
39. Mosser, D. M., & Edwards, J. P. (2008). Exploring the full spectrum of macrophage activation. *Nature Reviews Immunology*, 8(12), 958–969. <https://doi.org/10.1038/nri2448>
40. Wolfs, I., Donners, M., & de Winther, M. (2011). Differentiation factors and cytokines in the atherosclerotic plaque micro-environment as a trigger for macrophage polarisation. *Thrombosis and Haemostasis*, 106(11), 763–771. <https://doi.org/10.1160/th11-05-0320>
41. Lankford, K. L., Arroyo, E. J., Nazimek, K., Bryniarski, K., Askenase, P. W., & Kocsis, J. D. (2018). Intravenously delivered mesenchymal stem cell-derived exosomes target M2-type macrophages in the injured spinal cord. *PLOS ONE*, 13(1), e0190358. <https://doi.org/10.1371/journal.pone.0190358>
42. Liddelow, S. A., Guttenplan, K. A., Clarke, L. E., Bennett, F. C., Bohlen, C. J., Schirmer, L., Bennett, M. L., Münch, A. E., Chung, W.-S., Peterson, T. C., Wilton, D. K., Frouin, A., Napier, B. A., Panicker, N., Kumar, M., Buckwalter, M. S., Rowitch, D. H., Dawson, V. L., Dawson, T. M., & Stevens, B. (2017). Neurotoxic reactive astrocytes are induced by activated microglia. *Nature*, 541(7638), 481–487. <https://doi.org/10.1038/nature21029>
43. Wang, L., Pei, S., Han, L., Guo, B., Li, Y., Duan, R., Yao, Y., Xue, B., Chen, X., & Jia, Y. (2018). Mesenchymal Stem Cell-Derived Exosomes Reduce A1 Astrocytes via Downregulation of Phosphorylated NFκB P65 Subunit in Spinal Cord Injury. *Cellular Physiology and Biochemistry*, 50(4), 1535–1559. <https://doi.org/10.1159/000494652>
44. Liu, W., Wang, Y., Gong, F., Rong, Y., Luo, Y., Tang, P., Zhou, Z., Zhou, Z., Xu, T., Jiang, T., Yang, S., Yin, G., Chen, J., Fan, J., & Cai, W. (2019). Exosomes Derived from Bone Mesenchymal Stem Cells Repair Traumatic Spinal Cord Injury by Suppressing the Activation of A1 Neurotoxic Reactive Astrocytes. *Journal of Neurotrauma*, 36(3), 469–484. <https://doi.org/10.1089/neu.2018.5835>
45. Shi, Z., Yuan, S., Shi, L., Li, J., Ning, G., Kong, X., & Feng, S. (2021). Programmed cell death in spinal cord injury pathogenesis and therapy. *Cell Proliferation*, 54(3). <https://doi.org/10.1111/cpr.12992>
46. Li, C., Jiao, G., Wu, W., Wang, H., Ren, S., Zhang, L., Zhou, H., Liu, H., & Chen, Y. (2019). Exosomes from Bone Marrow Mesenchymal Stem Cells Inhibit Neuronal Apoptosis and Promote Motor Function Recovery via the Wnt/β-catenin Signaling Pathway. *Cell Transplantation*, 28(11), 1373–1383. <https://doi.org/10.1177/0963689719870999>
47. Gu, J., Jin, Z. S., Wang, C. M., Yan, X. F., Mao, Y. Q., & Chen, S. (2020). Bone Marrow Mesenchymal Stem Cell-Derived Exosomes Improves Spinal Cord Function After Injury in Rats by Activating Autophagy. *Drug Design, Development and Therapy*, Volume 14, 1621–1631. <https://doi.org/10.2147/dddt.s237502>

48. Kang, J., Li, Z., Zhi, Z., Wang, S., & Xu, G. (2019). MiR-21 derived from the exosomes of MSCs regulates the death and differentiation of neurons in patients with spinal cord injury. *Gene Therapy*, 26(12), 491–503. <https://doi.org/10.1038/s41434-019-0101-8>
49. Zhuang, Z., Liu, M., Luo, J., Zhang, X., Dai, Z., Zhang, B., Chen, H., Xue, J., He, M., Xu, H., & Liu, A. (2022). Exosomes derived from bone marrow mesenchymal stem cells attenuate neurological damage in traumatic brain injury by alleviating glutamate-mediated excitotoxicity. *Experimental Neurology*, 357, 114182–114182. <https://doi.org/10.1016/j.expneurol.2022.114182>
50. Oudega, M. (2012). Molecular and cellular mechanisms underlying the role of blood vessels in spinal cord injury and repair. *Cell and Tissue Research*, 349(1), 269–288. <https://doi.org/10.1007/s00441-012-1440-6>
51. Qu, Q., Pang, Y., Zhang, C., Liu, L., & Bi, Y. (2020). Exosomes derived from human umbilical cord mesenchymal stem cells inhibit vein graft intimal hyperplasia and accelerate reendothelialization by enhancing endothelial function. *Stem Cell Research & Therapy*, 11(1). <https://doi.org/10.1186/s13287-020-01639-1>
52. Xue, C., Shen, Y., Li, X., Li, B., Zhao, S., Gu, J., Chen, Y., Ma, B., Wei, J., Han, Q., & Zhao, R. C. (2018). Exosomes Derived from Hypoxia-Treated Human Adipose Mesenchymal Stem Cells Enhance Angiogenesis Through the PKA Signaling Pathway. *Stem Cells and Development*, 27(7), 456–465. <https://doi.org/10.1089/scd.2017.0296>
53. Jin, L.-Y., Li, J., Wang, K.-F., Xia, W.-W., Zhu, Z.-Q., Wang, C.-R., Li, X.-F., & Liu, H.-Y. (2021). Blood–Spinal Cord Barrier in Spinal Cord Injury: A Review. *Journal of Neurotrauma*, 38(9), 1203–1224. <https://doi.org/10.1089/neu.2020.7413>
54. Lu, Y., Zhou, Y., Zhang, R., Wen, L., Wu, K., Li, Y., Yao, Y., Duan, R., & Jia, Y. (2019). Bone Mesenchymal Stem Cell-Derived Extracellular Vesicles Promote Recovery Following Spinal Cord Injury via Improvement of the Integrity of the Blood-Spinal Cord Barrier. *Frontiers in Neuroscience*, 13. <https://doi.org/10.3389/fnins.2019.00209>
55. Matsushita, T., Lankford, K. L., Arroyo, E. J., Sasaki, M., Neyazi, M., Radtke, C., & Kocsis, J. D. (2015). Diffuse and persistent blood–spinal cord barrier disruption after contusive spinal cord injury rapidly recovers following intravenous infusion of bone marrow mesenchymal stem cells. *Experimental Neurology*, 267, 152–164. <https://doi.org/10.1016/j.expneurol.2015.03.001>
56. Anderson, M. A., Burda, J. E., Ren, Y., Ao, Y., O’Shea, T. M., Kawaguchi, R., Coppola, G., Khakh, B. S., Deming, T. J., & Sofroniew, M. V. (2016). Astrocyte scar formation aids central nervous system axon regeneration. *Nature*, 532(7598), 195–200. <https://doi.org/10.1038/nature17623>
57. Lee, J.-R., Jae Won Kyung, Kumar, H., Sung Pil Kwon, Seuk Young Song, Han, I., & Byung Soo Kim. (2020). Targeted Delivery of Mesenchymal Stem Cell-Derived Nanovesicles for Spinal Cord Injury Treatment. *International Journal of Molecular Sciences*, 21(11), 4185–4185. <https://doi.org/10.3390/ijms21114185>

58. Feng, C., Deng, L., Yong, Y.-Y., Wu, J.-M., Qin, D.-L., Yu, L., Zhou, X.-G., & Wu, A.-G. (2023). The Application of Biomaterials in Spinal Cord Injury. *International Journal of Molecular Sciences*, 24(1), 816. <https://doi.org/10.3390/ijms24010816>
59. Wang, M., Zhai, P., Chen, X. R., Schreyer, D. J., Sun, X., & Cui, F. (2011). Bioengineered Scaffolds for Spinal Cord Repair. *Tissue Engineering Part B-Reviews*, 17(3), 177–194. <https://doi.org/10.1089/ten.teb.2010.0648>
60. Doblado, L. R., Martínez-Ramos, C., & Pradas, M. M. (2021). Biomaterials for Neural Tissue Engineering. *Frontiers in Nanotechnology*, 3. <https://doi.org/10.3389/fnano.2021.643507>
61. McKay, C. A., Pomrenke, R. D., McLane, J. S., Schaub, N. J., DeSimone, E. K., Ligon, L. A., & Gilbert, R. J. (2014). An Injectable, Calcium Responsive Composite Hydrogel for the Treatment of Acute Spinal Cord Injury. *ACS Applied Materials & Interfaces*, 6(3), 1424–1438. <https://doi.org/10.1021/am4027423>
62. Poongodi, R., Chen, Y.-L., Yang, T.-H., Huang, Y.-H., Yang, K. D., Lin, H.-C., & Cheng, J.-K. (2021). Bio-Scaffolds as Cell or Exosome Carriers for Nerve Injury Repair. *International Journal of Molecular Sciences*, 22(24), 13347. <https://doi.org/10.3390/ijms222413347>
63. Li, L., Zhang, Y., Mu, J., Chen, J., Zhang, C., Cao, H., & Gao, J. (2020). Transplantation of Human Mesenchymal Stem-Cell-Derived Exosomes Immobilized in an Adhesive Hydrogel for Effective Treatment of Spinal Cord Injury. *Nano Letters*, 20(6), 4298–4305. <https://doi.org/10.1021/acs.nanolett.0c00929>
64. Wang, C., Wang, M., Xia, K., Wang, J., Cheng, F., Shi, K., Ying, L., Yu, C., Xu, H., Xiao, S., Liang, C., Li, F., Lei, B., & Chen, Q. (2021). A bioactive injectable self-healing anti-inflammatory hydrogel with ultralong extracellular vesicles release synergistically enhances motor functional recovery of spinal cord injury. *Bioactive Materials*, 6(8), 2523–2534. <https://doi.org/10.1016/j.bioactmat.2021.01.029>
65. Cheng, J., Chen, Z., Liu, C., Zhong, M., Wang, S., Sun, Y., Wen, H., & Shu, T. (2021). Bone mesenchymal stem cell-derived exosome-loaded injectable hydrogel for minimally invasive treatment of spinal cord injury. *Nanomedicine*, 16(18), 1567–1579. <https://doi.org/10.2217/nnm-2021-0025>
66. Zhang, Y., Li, L., Mu, J., Chen, J., Feng, S., & Gao, J. (2020). Implantation of a functional TEMPO-hydrogel induces recovery from rat spinal cord transection through promoting nerve regeneration and protecting bladder tissue. *Biomaterials Science*, 8(6), 1695–1701. <https://doi.org/10.1039/C9BM01530B>
67. Fan, L., Liu, C., Chen, X., Zheng, L., Zou, Y., Wen, H., Guan, P., Lu, F., Luo, Y., Tan, G., Yu, P., Chen, D., Deng, C., Sun, Y., Zhou, L., & Ning, C. (2022). Exosomes-Loaded Electroconductive Hydrogel Synergistically Promotes Tissue Repair after Spinal Cord Injury via Immunoregulation and Enhancement of Myelinated Axon Growth. *Advanced Science*, 9(13), 2105586. <https://doi.org/10.1002/advs.202105586>

68. Li, P., Kaslan, M., Lee, S. H., Yao, J., & Gao, Z. (2017). Progress in Exosome Isolation Techniques. *Theranostics*, 7(3), 789–804. <https://doi.org/10.7150/thno.18133>
69. Bari, E., Perteghella, S., Di Silvestre, D., Sorlini, M., Catenacci, L., Sorrenti, M., Marrubini, G., Rossi, R., Tripodo, G., Mauri, P., Marazzi, M., & Torre, M. (2018). Pilot Production of Mesenchymal Stem/Stromal Freeze-Dried Secretome for Cell-Free Regenerative Nanomedicine: A Validated GMP-Compliant Process. *Cells*, 7(11), 190. <https://doi.org/10.3390/cells7110190>
70. Lopatina, T., Bruno, S., Tetta, C., Kalinina, N., Porta, M., & Camussi, G. (2014). Platelet-derived growth factor regulates the secretion of extracellular vesicles by adipose mesenchymal stem cells and enhances their angiogenic potential. *Cell Communication and Signaling*, 12(1), 26. <https://doi.org/10.1186/1478-811x-12-26>
71. Liu, W., Ma, Z., Li, J., & Kang, X. (2021). Mesenchymal stem cell-derived exosomes: therapeutic opportunities and challenges for spinal cord injury. *Stem Cell Research & Therapy*, 12(1). <https://doi.org/10.1186/s13287-021-02153-8>
72. Gupta, D., Zickler, A. M., & El Andaloussi, S. (2021). Dosing extracellular vesicles. *Advanced Drug Delivery Reviews*, 178, 113961. <https://doi.org/10.1016/j.addr.2021.113961>

The Effect of Student Loan Debt Forgiveness on Social Mobility

By Pranav Pradeep

Abstract:

This paper explores the impact of student loan debt forgiveness on social mobility in the United States. It argues that student loan debt forgiveness could create economic opportunities by promoting job mobility among borrowers, facilitating homeownership, and enabling entrepreneurship. By analyzing empirical data and existing literature, the paper highlights how debt forgiveness can eliminate barriers to economic participation, particularly for the average American and low-income individuals.

Despite concerns about the economic implications of forgiveness and moral hazards, the potential benefits of increased social mobility through debt forgiveness are substantial. They have the power to reshape the American economy and promote long-term financial stability.

Introduction:

The student debt crisis in the United States is not just a personal burden but a significant national economic issue. With a staggering 44 million individuals trapped in student debt, the outstanding student loan debt has reached \$1.4 trillion. This financial burden is a personal struggle and a drag on the United States economy, hindering various aspects of life, from small business formation to new home buying, marriage, and reproduction (Khattar et al.).

The Saving on a Valuable Education (SAVE) Plan:

In a pivotal moment in July 2023, the United States Supreme Court delivered a significant ruling on the Biden Administration's student loan forgiveness program. In *Biden v. Nebraska*, a 6-3 majority of the court invalidated the administration's plan, supporting the six challenging states in their right to sue. The court's ruling clarified that the Higher Education Relief Opportunities for Students Act, or HEROES Act, does not authorize the administration's student loan forgiveness plan. The HEROES Act allows the Secretary of Education to waive or modify provisions for loan forgiveness under the Higher Education Act in case of a war, military action, or national emergency. The court viewed student debt forgiveness as too sweeping of a change compared to simply waiving or modifying it. The court's decision effectively halted a plan that would have erased almost \$430 billion in student debt and reduced the median amount of non-forgiven loan repayments from \$29,400 to \$13,600 per individual (Brookings). In a blow to millions, the decision left America's most vulnerable seeking financial freedom from student loan debt behind.

In response to the escalating issue, the Biden Administration has proposed the Saving on a Valuable Education (SAVE) Plan, which could benefit over 20 million borrowers. The plan, designed to halve payments or reduce them to 5% of income, could offer significant relief to many borrowers, providing hope in the student debt crisis. Like other Income-Driven Repayment (IDR) plans, the SAVE Plan calculates the borrower's monthly payment amount based on the

family's income and size. However, the SAVE Plan offers unique benefits to low-income borrowers to reduce their monthly payment amounts, such as an interest benefit that covers the borrower's monthly interest if the monthly payment is fully paid. This benefit prevents the borrower's balance from ballooning due to unpaid interest payments (Federal Student Aid).

It's crucial to note that while the SAVE Plan could provide some relief, it is not a comprehensive solution. The plan does not forgive any debt. Instead, it extends the debt over time, compelling America's most vulnerable to make constant monthly payments. Those who borrowed \$12,000 or less can expect complete forgiveness only after a decade of payments, and those with higher levels of debt will be forced into a situation where monthly payments and interest will burden them for decades (Federal Student Aid 1).

Furthermore, Biden's SAVE policy shifts the burden of student debt onto the backs of taxpayers, mitigating the benefits of forgiveness (Cochling). For these reasons, many income-driven repayment plans feel like a lifetime sentence. While IDR plans can reduce monthly payments, the admin process is cumbersome and confusing, leading to more outstanding debt (Ananda et al.).

Thus, the United States federal government forgiving all federal student loan debt would enhance social mobility by promoting entrepreneurship, homeownership, and job mobility. While short-term deficit spending may result, the economic advantages of increased social mobility outweigh the drawbacks of deficit spending.

Job Mobility:

Student loan debt forgiveness gives Americans the financial flexibility to seek higher-paying jobs and incomes while repaying existing expenses. Debtholders are worried that collectors will "garnish their wages," so they do not seek higher-paying jobs. Additionally, debtholders are deterred from pursuing higher-paying jobs as they believe that any income made will be relinquished to debt collectors. Empirically, when student loans were forgiven, borrowers "were 30% more likely to change jobs" and also observed a "4,000 dollar boost to their income" (Harvard Business School). Some argue that the decline in job mobility is due to a lack of motivation and incentives. However, because employers perform credit checks, borrowers in default are less likely to be employed. Therefore, the decline in job mobility is most closely correlated with the enormous debt impeding borrowers from seeking employment.

There are concerns that student debt forgiveness is a moral hazard, disincentivizing Americans from becoming employed. However, student debt forgiveness motivates Americans to seek high-paying jobs, as debtholders understand that debt collectors will not garnish their high wages. However, due to this fear, debtholders "may not be keen on looking for higher-paying work just to pay collectors more." (Harvard Business School 1).

Smart Spending:

Borrowers who observed increased income and upward job mobility followed smart spending habits, which involved making deliberate financial choices aligned with long-term

goals. When loan debt is forgiven, borrowers can repay personal debts, such as credit card, auto, and mortgage payments. They even generate wealth with any leftover income. For example, a poll asked what adults with student debt would do with the extra money if the debt was forgiven. The adults answered that “they would contribute to long-term savings (43%), pay off other debts (40%), invest (38%), or purchase a home (24%) or car (22%)” (National Association of Realtors). When loan debts are paid off, former debtholders “say they use the additional funds to pay off other debts (34%), contribute to long-term savings (31%), or invest (26%)” (National Association of Realtors 1). Further, “borrowers whose loans were forgiven reduced total debts by 26%” (National Association of Realtors 2). Thus, debt forgiveness must be used as an economic tool to alleviate immediate financial burdens, paving the way for individuals to foster a stable financial future.

However, as debt accumulates, individuals forgo smart spending habits and frivolously purchase, known as the “what-the-hell effect,” referring “to the loss of restraint following the initial violation of a goal” (Zhang et al.). When an individual violates their financial goals by assuming more debt, this “goal abandonment leads to overspending” (Zhang et al. 1).

Some argue that welfare programs that increase welfare benefits decrease employee outcomes and workforce participation; however, empirical studies show that well-designed forgiveness programs support workforce participation by providing a safety net for individuals pursuing higher-skilled and better-paying jobs, eliminating the fear of poverty (Zhang et al. 2).

Some argue that student loan debt forgiveness will incentivize unemployment. However, individuals with aspirations of bettering their economic mobility will never leave the workforce or become unemployed due to credit incentives, “Since students will still need to be responsible to secure good credit, it is highly unlikely that students stop putting in efforts” (Zhao et al.).

It is essential to distinguish between forgiving student loans, which alleviates existing financial burdens and fosters economic participation, and giving money to individuals without conditions. Welfare programs are designed to support individuals during times of need while promoting long-term economic stability. Forgiven student loans relieve individuals of crushing debt and enable them to invest in their futures, pursue higher education, and contribute more effectively to the economy.

Homeownership:

High levels of student debt and low incomes present significant barriers to homeownership: student debt burdens incomes, preventing savings for a down payment and precluding borrowers from accessing a mortgage. For example, loan levels are growing, and “student debt as a share of income is highest” in the lowest-income areas (Perry et al.). This debt is preventing low-income Americans from owning a home in the first place. Student debt inhibits Americans from making mortgage payments and disincentivizes American homeownership due to the fear of lower credit scores. Debtholders, especially those with low incomes, do not have enough liquid cash for a downpayment, as high debt precludes individuals' ability to spend their disposable income. Instead, they pay off their debt (Perry et al. 1).

High levels of student debt among millennials increase their reliance on credit cards and hinder their ability to qualify for mortgages, perpetuating a cycle of debt and limiting social mobility. Consequently, “because millennial borrowers have less liquid cash, they also end up taking on more credit card debt” (Safier); “low credit scores and high Debt-to-Income ratio preclude a mortgage”(Huffman), as banks are turned away due to the fear of default and financial risk. Even for those who qualify for mortgages, there is a never-ending cycle of debt as those with higher levels of debt must also pay higher interest rates on their mortgage; this is because “as rates go up, so does a borrower’s debt-to-income (DTI) ratio” (Walsh et al.). The cycle continues as individuals must make consistent payments, or "monthly loan payments that are not made on time cause substantial effects on credit score” (Huffman 1). Americans with high debt and low credit may not qualify for loans or down payments from banks, hindering all facets of social mobility.

There is a misconception that student debtholders only make up a small portion of home buyers. However, “nearly one-quarter of all home buyers, and 37% of first-time buyers, have student loan debt” (National Association of Realtors). Even worse, because student debt cripples individuals' financial flexibility, many student debtholders have the intent to own a home; however, they lack the financial capability to do so as “61% of non-homeowning millennials said that student loan debt is delaying their ability to buy a home” (National Association of Realtors 1).

Owning a home is of the utmost financial importance; historically, it has been the best way to accumulate wealth. This is because a home is a generational asset that can bestow families with decades of wealth, as homes only appreciate over time. For this reason, giving Americans the wealth-building assets they deserve is critical, as “Housing inequality may be the biggest contributor to our economic divides” (Florida et al.). Additionally, homeownership is a vital component of the American dream, symbolizing financial success and independence. In fact, “surveys have suggested that many place owning a home above their career, family and college as a sign of prosperity” (Snyder).

Student debtholders “with high debt have a lower willingness to spend - especially for new equity,” as most debtholders' savings are spent repaying their debts instead of consumption (Huffman). The economic theory of marginal returns summarizes this financial peril. For low-income debtholders, even marginal amounts of extra income at their disposal increase the likelihood of individuals spending it, boosting the economy writ large. Encouraging spending among lower-income groups has a distributive effect, boosting consumption and supporting economic growth at various levels of society (Hayes et al.).

Some argue that Biden’s SAVE plan already bestows debtholders with additional disposable income. However, anything short of universal forgiveness is ineffective. Biden’s SAVE plan does not forgive any debt; instead, it “leaves millions with the same monthly payment they had before, negating cancellation” (Robinson). Instead of debtholders spending their disposable income on a mortgage or the economy, they must make consistent decade-long monthly payments to pay off their student loans. Some argue that trickle-down economics will

eventually benefit those lower on the economic ladder; however, if this were true, debtholders would already be experiencing economic relief, and the United States debt crisis should have shown signs of improvement. Instead, "the federal student loan debt balance in the United States has almost tripled in the past 15 years", contradicting the theory of trickle-down economics (Rose).

Entrepreneurship:

High levels of student debt inhibit aspiring entrepreneurs' access to capital and risk-taking, which is crucial for starting businesses reliant on personal debt financing. Small businesses depend on funding from personal debt; student debt rose as a share of total personal debt. When entrepreneurs cannot purchase capital goods and invest in their businesses, their business output flounders, which is why "a 3% increase in relative student debt is associated with a 14% decline in small business" (Fullwiler). Starting a business also comes with an associated risk that many debtholders do not have the privilege of taking, as "the ability to take risks is largely predicated on your capacity to finance your lifestyle while you aren't making consistent money" (Raines). Because debtholders cannot finance themselves during the initial start-up of a business, they are often deterred from starting a business in the first place. This deterrence incentivizes debtholders to forget their entrepreneurial dreams and pursue the workforce. When most graduates leave college with debt, doing anything but getting a well-paying job to reduce this burden might seem irresponsible. For this reason, "the rate of new businesses has been dropping by 29%" (Revzin et al.).

Some may argue that debt is not the root cause of declined business growth. However, debt relief accelerates existing business investment, and because a business's output is only as strong as its input and investment, "47% of small business owners with student loans say that the collection pause enabled them to invest more in business" (Buttle). This is also historically true as "student debt led to two million fewer new businesses being formed between 2006 and 2015", highlighting the significant impact of student debt on entrepreneurial activity and the potential benefits of debt forgiveness in fostering business innovation and growth (Kaufman et al.).

The United States economy is only as strong as its small businesses. They curb the inevitable recessionary dips created by the boom-and-bust business cycle. Small businesses are the heart of employment during recessions. They are thirteen times more innovative and are more open to providing jobs to people who face barriers to employment (Pollak). When so many businesses are dying, encouraging the birth of new ones becomes more crucial than ever.

Conclusion:

Addressing the burden of student loan debt through comprehensive forgiveness measures can significantly enhance social mobility in America by fostering job mobility, promoting greater homeownership, and alleviating financial constraints on entrepreneurship. Even better, student loan debt forgiveness pays for itself by incentivizing economic activity writ large and empowering contributions to the economy through more business and investment. In times of

inevitable recessionary dips and economic slowdowns, student loan debt forgiveness is the policy prescription that America needs.

Despite concerns about economic implications and moral hazards, the benefits of increased social mobility and economic participation outweigh the costs, making student debt forgiveness a crucial step towards a more equitable and prosperous society.

Works Cited

- Ananda, Kitana, et al. "Cancelling Student Debt Is Necessary for Racial Justice - Non Profit News." Nonprofit Quarterly, 1 December 2022, <https://nonprofitquarterly.org/cancelling-student-debt-necessary-for-racial-justice/>. Accessed 5 July 2024.
- Brookings. "student-loans-the-racial-wealth-divide-and-why-we-need-fullstudent-debt-cancellation." Brookings, Brookings, <https://www.brookings.edu/articles/student-loans-the-racial-wealth-divide-and-why-we-need-fullstudent-debt-cancellation>.
- Buttle, Rhett. "Student Debt Relief Could Spur Innovation And Encourage Small Business Growth." Forbes, Forbes, <https://www.forbes.com/sites/rhettbuttle/2022/09/30/student-debt-relief-could-spur-innovation-and-encourage-small-business-growth/?sh=196239053880>.
- Cochling, Eric. "Why student loan forgiveness plan is bad for the poor and working class." Georgia Center For Opportunity, <https://foropportunity.org/why-student-loan-forgiveness-plan-is-bad-for-the-poor-and-working-class/>. Accessed 5 July 2024.
- Florida, Richard, and Feargus O'Sullivan. "Economic Inequality Largely Boils Down to Housing Inequality." Bloomberg, 13 April 2018, <https://www.bloomberg.com/news/articles/2018-04-13/economic-inequality-largely-boils-down-to-housing-inequality>. Accessed 5 July 2024.
- Fullwiler, Scott. "THE MACROECONOMIC EFFECTS OF STUDENT DEBT CANCELLATION." Levy Economics Institute of Bard College, https://www.levyinstitute.org/pubs/rpr_2_6.pdf. Accessed 5 July 2024.
- Harvard Business School. "Forgiving Student Loan Debt Leads To Better Jobs, Stronger Consumers." Harvard Business School, <https://hbswk.hbs.edu/item/forgiving-student-loan-debt-leads-to-better-jobs-stronger-consumers>.
- Hayes, Adam. "Law of Diminishing Marginal Returns: Definition, Example, Use in Economics." Investopedia, <https://www.investopedia.com/terms/l/lawofdiminishingmarginalreturn.asp>. Accessed 5 July 2024.
- Huffman, Anna. "Forgive and Forget? An Analysis of Student Loan Forgiveness Plans." Carolina Law Scholarship Repository, 1 March 2020, <https://scholarship.law.unc.edu/cgi/viewcontent.cgi?article=1512&context=ncbi>. Accessed 5 July 2024.
- "The Impact of Student Loan Debt." National Association of REALTORS®, <https://www.nar.realtor/research-and-statistics/research-reports/the-impact-of-student-loan-debt>. Accessed 5 July 2024.
- Kaufman, Ben, and Mark Herbert. "Research Roundup: The Student Debt Crisis is a Crisis for Small Businesses and Entrepreneurship." Student Borrower Protection Center, https://protectborrowers.org/smallbiz_studendebt/. Accessed 5 July 2024.
- Khattar, Rose, and Zahir Rasheed. "Canceling at Least \$10,000 of Student Loan Debt Will Help Lower the Cost of Living." American Progress, American Progress, <https://www.americanprogress.org/article/canceling-student-debt-of-at-least-10000-will-help-lower-the-cost-of-living/>.
- Perry, Andre. "Student loans, the racial wealth divide, and why we need full student debt cancellation | Brookings." Brookings Institution, 23 June 2021, <https://www.brookings.edu/articles/student-loans-the-racial-wealth-divide-and-why-we-need-full-student-debt-cancellation/>. Accessed 5 July 2024.

- Pollak, Julia. "Small Business is Key to the COVID Recovery." CWI Labs, CWI Labs, <https://www.cwilabs.org/small-business-is-key-to-the-covid-recovery/>.
- Raines, Jack. "The Privilege to Take Risks." Young Money, Jack Raines, <https://www.youngmoney.co/p/privilege-take-risks>.
- Revzin, Vadim. "Student Debt Is Stopping U.S. Millennials from Becoming Entrepreneurs." Harvard Business Review, Harvard, <https://hbr.org/2019/04/student-debt-is-stopping-u-s-millennials-from-becoming-entrepreneurs>.
- Robinson, Nathan J. "Destroying All the Arguments Against Cancelling Student Debt." Current Affairs, 21 May 2022, <https://www.currentaffairs.org/news/2022/05/destroying-all-the-arguments-against-cancelling-student-debt>. Accessed 5 July 2024.
- Rose, Marley. "Average Student Loan Debt: 2024 Statistics." BestColleges.com, <https://www.bestcolleges.com/research/average-student-loan-debt/>. Accessed 5 July 2024.
- Safier, Rebecca. "Study Shows Student Debt Can Kill 75% of Millennials' Average Net Worth." Lending Tree, <https://www.lendingtree.com/student/student-debt-can-kill-millennials-net-worth-study/>. Accessed 5 July 2024.
- "The Saving on a Valuable Education (SAVE) Plan Offers Lower Monthly Loan Payments." Federal Student Aid, <https://studentaid.gov/announcements-events/save-plan>. Accessed 5 July 2024.
- Snyder, Shawn. "The American Home Dream in Today's Economy." J.P. Morgan, 7 December 2023, <https://www.jpmorgan.com/insights/real-estate/housing-market/the-american-homeowners-dream-vs-recent-reality>. Accessed 5 July 2024.
- "Student Loan Debt." National Association of REALTORS®, <https://www.nar.realtor/student-loan-debt>. Accessed 5 July 2024.
- Walsh. "Rising Interest Rates Put the Brakes on the Mortgage Market for Borrowers with Low Incomes." Urban Institute, 24 August 2023, <https://www.urban.org/urban-wire/rising-interest-rates-put-brakes-mortgage-market-borrowers-low-incomes>. Accessed 5 July 2024.
- Zhang. "The Effect of Student Loan Debt on Spending." Faculty Darden Virginia, https://faculty.darden.virginia.edu/wilcoxr/pdf_docs/StudentLoanDebt-WorkingPaper.pdf.
- Zhao, Jiayu, and Joseph Zhang. "Policy Recommendation for Managing College Debt Crisis and Retirement Savings." iOme Challenge 2016-2017, Columbia University, https://iomechallenge.org/wp-content/uploads/2017%20iOme%20Essay_wAuthors_Managing%20College%20Debt%20Crisis%20and%20Retirement.pdf.

A Review of Brain Stimulation as an Alternative Method of Treating OCD

By Aiden Chung

Abstract

The exact causes of many psychiatric conditions are not fully understood by the scientific community today, such as obsessive-compulsive disorder (OCD). Obsessions and compulsions, the symptoms associated with OCD, hinder people from even carrying out simple tasks and living their daily lives without trouble. To alleviate this issue, serotonin reuptake inhibitors (SRI) and cognitive behavioral therapy (CBT) have become used as the most common treatments for any patient with OCD. However, since every patient responds differently to therapeutics, some find themselves unresponsive to both SRIs and CBT. This means that if their OCD symptoms are disrupting their daily lives, they have to find an equally, if not more, effective alternative option for treating their disorder, which is where brain stimulation comes in. Brain stimulation, the techniques and therapies aimed at modulating brain activity, has been used for various psychiatric-related disorders for a while now; however, it has just been recently explored, roughly over the past two decades, for its potential of being therapeutic for OCD. This paper will explore some of the significant brain stimulation techniques that have been experimented with for treating OCD to ultimately address the question of their effectiveness as an alternative therapeutic for patients who are unresponsive to SRIs and CBT. It will also discuss some of the pitfalls associated with this form of treatment, what scientists can do to advance understanding in this relatively new field, and how that understanding can potentially lead to the creation of more effective brain stimulation techniques and an improved treatment experience for both patients and administrators.

Introduction

Obsessive-compulsive disorder (OCD) is a neurological disorder where people find themselves repeatedly having unwanted, intrusive thoughts (obsessions), uncontrollable, repetitive behaviors (compulsions), or both (*Obsessive-Compulsive Disorder - National Institute of Mental Health (NIMH)*). Examples of obsessions include constant doubts about whether the door is locked or not and intense stress when a group of objects is not oriented in a specific way. Compulsions involve behaviors such as constantly washing your hands and repeatedly checking the stove to see if it is turned off. Someone is diagnosed with OCD if these time-consuming obsessions and compulsions cause significant distress or interfere with daily life (*Obsessive-Compulsive Disorder - National Institute of Mental Health (NIMH)*). An example of this interference is not being able to get ready in the morning without having to repeat some steps multiple times. It is not as uncommon as people think, with 2.5 million adults in the United States being affected by it, translating to roughly 1.2% of the United States population (*Facts & Statistics | Anxiety and Depression Association of America, ADAA*).

As a concept, symptoms of OCD have been recognized for centuries, even as early as the 18th century, when they were thought of as being linked to someone's religious melancholy and

guilt (“History”). Since then, scientific advancements over the years have helped us develop our understanding of this condition from moral weakness or intellectual disability to one that complexly affects someone’s mental health and behavior. Despite this, it is still not precisely known what causes someone to develop OCD; however, it is generally thought that genetics, brain abnormalities, and the environment all play a role (*Obsessive-Compulsive Disorder (OCD) | Johns Hopkins Medicine*). The genetics component of this current theory comes from the fact that having family members with OCD can increase the chances of one getting this disorder (“Obsessive-Compulsive Disorder (OCD) - Symptoms and Causes”). Brain abnormalities, another aspect of this theory, comes from how mental disorders (e.g., PTSD) can alter someone’s brain functions or body chemistry, allowing the potential of them getting OCD (Bremner). The environmental aspect is attributed to how obsessive and compulsive behaviors can be learned if a person constantly watches the others closest to them exhibit these symptoms (“Obsessive-Compulsive Disorder (OCD) - Symptoms and Causes”).

Based on current knowledge of how OCD emerges and exhibits itself in people, one common type of therapeutic scientists formed to lower the symptoms associated with the disorder is psychotherapy (Fornaro et al.). Psychotherapy, also known as talk therapy, refers to a type of collaborative treatment (between the psychiatrist and individual) that uses psychological techniques to help a patient identify and change distressing behaviors, emotions, or thoughts (*Psychotherapies - National Institute of Mental Health (NIMH)*). One example of this that is commonly used is cognitive behavior therapy (CBT), which aims to help you identify negative behaviors or ways of thinking so that you can work on fixing them in a clear and effective manner (*Cognitive Behavioral Therapy - Mayo Clinic*). Another common type of therapeutic is medications, such as serotonin reuptake inhibitors (SRI), which is a type of antidepressant that works by increasing serotonin levels in the brain (Fornaro et al.; “Overview - SSRI Antidepressants”). There are patients, however, for whom neither SRIs nor CBT prove effective, so scientists have been experimenting with other methods to treat OCD, such as various brain stimulation therapies. Brain stimulation essentially aims to treat various psychiatric and neurological disorders by modulating brain activity via activating or inhibiting specific brain parts with electricity (*Brain Stimulation Therapies - National Institute of Mental Health (NIMH)*). All the different forms of brain stimulation therapy were formed to serve as the primary third treatment option for patients who are drug-resistant or are unresponsive to psychotherapies (*Brain Stimulation Therapies - National Institute of Mental Health (NIMH)*). The merit of brain stimulation as another main method of treating OCD comes from how OCD, as a mental and behavioral disorder, arises due to abnormalities in specific regions of the brain. Therefore, a therapeutic that directly targets the parts of the brain responsible for OCD, like brain stimulation, has a great potential of being an effective treatment.

The following sections will delve into the various types of brain stimulation therapies used for OCD. This paper will specifically look at how these specific methods work and their effectiveness when put into practice. This is to ultimately answer the question of whether or not brain stimulation is an effective alternative method for treating OCD. It will also touch upon

what scientists can do to further our understanding of the mechanisms behind effective brain stimulation and how they can use that to develop the most effective and accessible form of treatment possible.

Discussion

Transcranial Magnetic Stimulation

Transcranial Magnetic Stimulation (TMS) is a noninvasive form of brain stimulation that modifies neural activity in particular brain regions using magnetic fields (Carmi, Tendler, et al.). It also adjusts neuronal excitability, the readiness of neurons to generate electrical impulses/signals in response to stimuli, and plasticity, the ability of a brain to change its structure and function of the whole brain (Carmi, Tendler, et al.).

Patients with OCD have been found to have hyperactivity in the cortical-striatal-thalamic-cortical loop circuit (CSTC) (Carmi, Tendler, et al.). The CSTC is a broad network of brain regions that controls habit formation, reward, and the execution of movement (Rădulescu et al.). It begins in the cortex, which sends signals to the striatum (Carmi, Tendler, et al.). From the striatum, the circuit progresses to the thalamus, where it then finally loops back to the cortex (Figure 1) (Carmi, Tendler, et al.). The CSTC is divided into three loops, which are interconnected brain pathways responsible for carrying out the functions of CSTC: sensorimotor, cognitive, and limbic (Carmi, Tendler, et al.). The sensorimotor loop, which starts from the brain's sensorimotor cortices, is responsible for sensory perception and motor control (Riemann and Lephart). The cognitive loop is responsible for higher-order cognitive functions of the brain; it starts from the dorsolateral prefrontal cortex (DLPFC), which takes part in executing executive functions (the set of cognitive processes and mental skills that enable individuals to plan, organize, regulate, and execute complex goal-directed behaviors) such as decision making and planning (Hertrich et al.). The limbic loop controls learning, memory, motivation, and emotional regulation (Torricco and Abdijadid). This loop starts from the orbitofrontal cortex (OFC), which plays a key role in evaluating the emotional significance of stimuli and making decisions based on their reward value, or the anterior cingulate cortex (ACC), which is involved in emotional processing, error detection, and monitoring conflict (Rolls et al.; Apps et al.).

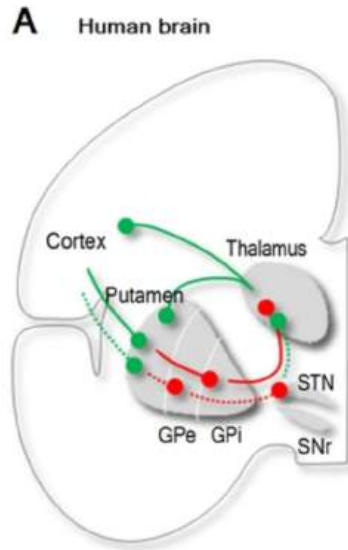
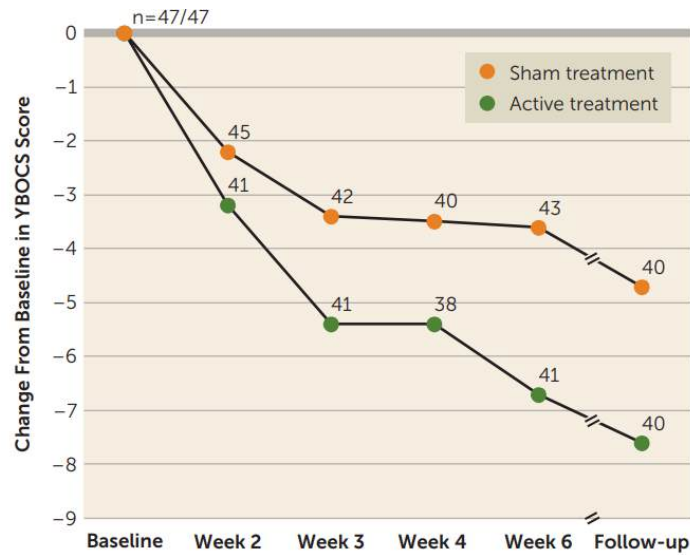


Fig 1: The synaptic connections of the CSTC pathway in the human brain. The dots represent the location of each region within the CSTC, while the lines represent the synapses that connect these regions together (Figure taken from Rădulescu et al., 2017).

Studies, in particular, have heavily suggested that abnormalities and hyperactivity of the CSTC's limbic loop, more specifically the ACC and OFC, lead to people getting symptoms of OCD (Carmi, Tendler, et al.). The ACC has been found to be involved in many of the cognitive processes that are impaired in OCD, such as the coordination of movement with emotions, motivations, and thoughts, leading to unintentional compulsions (Carmi, Tendler, et al.). It also has been reported that, after successful treatment, the activity of the limbic loop in patients decreased (Carmi, Tendler, et al.). So, the mechanisms behind TMS are designed to target this loop as well as, potentially, other regions of the CSTC, due to the other parts of CSTC also being found responsible for the dysfunctional or hyperactive brain functions that manifest as commonly seen symptoms (e.g., a hyperactive DLPFC leading to overactive motoring and checking, which are seen as obsessions, due its responsibility for planning and decision making) (Hertrich et al.).

During TMS, an electromagnetic coil is placed against the scalp of a patient's head, and this coil delivers magnetic pulses that stimulate nerve cells in specific regions of your brain responsible for the onset of OCD (*Brain Stimulation Therapies - National Institute of Mental Health (NIMH)*; Carmi, Tendler, et al.). More specifically, TMS uses an H-coil, the H7 coil (Carmi, Tendler, et al.). The H7 coil generates and delivers magnetic pulses of various frequencies to the medial prefrontal cortex (mPFC), a brain region containing ACC, OFC, and the ACC itself (Carmi, Tendler, et al.). The ultimate goal of targeting these regions and altering them with magnetic pulses is to lower the activity of the limbic loop, which, based on what studies suggest, should theoretically lower the symptoms associated with OCD (Carmi, Tendler, et al.).

So, to see the merit of TMS's theoretical usage, a study by Carmi et al. was conducted that tested the safety and efficacy of TMS, specifically deep transcranial magnetic stimulation (dTMS), in treating OCD. dTMS is a type of TMS that uses a specific coil that affects deeper and wider brain regions compared to other forms of TMS (*Transcranial Magnetic Stimulation - Mayo Clinic*). For this experiment, 100 OCD patients were randomly chosen around the world to be administered TMS with a high frequency (the active TMS) or a sham dTMS treatment for a period of six weeks. Patients chosen for this study were between 22 and 68 years old and did not respond adequately to past treatments of SRI or CBT (meaning that the patients either did not respond to SRIs after 2 months or one trial of SRI while undergoing CBT). For the patients that received the active TMS, the H7 coil of the device that administered TMS was placed four centimeters anterior to the foot motor cortex (which is responsible for controlling voluntary movements of the foot) and used with the maximum amount of magnitude needed to evoke a response in the leg muscles (100% resting motor threshold, RMT). Those particular settings were chosen because of how they were found to hit both the mPFC and ACC bilaterally; the RMT (for this and other studies in general), which is determined by single pulse trials (individual pulses delivered to the brain for testing purposes), helped calibrate the stimulation intensity needed to evoke a safe and efficacious response in patients (Carmi, Tendler, et al.). Patients under the active TMS were given 20 Hz dTMS at 100% RMT. The pulses were delivered in 2-second trains (sets of pulses) and 20-second intertrain intervals (periods of time between each train) for a total of 50 sets and 2000 pulses per session. Patients under the sham TMS were under similar conditions, meaning they could also feel the same sensations that a patient under active TMS would feel (e.g., scalp sensations), but the magnetic pulses did not directly penetrate the inside of the brain. There were many scales used to evaluate the efficacy of each of the treatments; however, one of the more major of these scales is the Yale-Brown Obsessive Compulsive Scale (Y-BOCS). Y-BOCS is one of the most widely used rating scales for OCD, and it is used to assess the severity of someone's OCD symptoms; it is on a scale from 0-40, with 0 being considered subclinical and 40 being considered extreme (*Diagnosis | Obsessive-Compulsive and Related Disorders | Stanford Medicine*). Overall, this study found that dTMS is an effective treatment for OCD as the Y-BOCS significantly decreased for patients who received the active treatment in comparison to those who received the sham ones (Figure 2) (Carmi, Tendler, et al.). They also found that the scores for the Clinical Global Impressions Severity Scale and the Clinical Global Impressions Improvement Scale (scales that assess the severity of someone's illness and how much they improve over time after using a treatment directed towards said illness) improved for patients receiving dTMS versus those not receiving the actual treatment (Carmi, Tendler, et al.; *The Clinical Global Impressions Scale - PMC*).



^a dTMS=deep repetitive transcranial magnetic stimulation; YBOCS=Yale-Brown Obsessive Compulsive Scale. Each data point includes the patients with recorded YBOCS scores at that time point.

Fig 2: The periodic change in Y-BOCS scores in patients when receiving sham or active dTMS. n=47/47 represents that there were 47 patients that were participating in the study before the treatment process (Figure taken from Carmi et al., 2019).

When interviewing patients, they found that the H7 coil did not have any major effects on them: the most prevalent side effect was mild headaches that occurred either during the dTMS or immediately after it, showing that dTMS is also a safe treatment on top of being an effective one (Carmi, Tendler, et al.). The general consensus of scientific studies published on this topic thus far is that active TMS is more effective than sham TMS; however, there is no universal agreement on how it should be administered (e.g., the specific parameters/settings TMS should be delivered with) (Carmi, Tendler, et al.). Specifically, targeting the OFC showed some positive results, while the dorsolateral prefrontal cortex showed mixed results (Carmi, Tendler, et al.). It was also found that targeting the supplementary motor area (SMA) and pre-supplementary motor area (pre-SMA), the areas of the brain involved in the planning and coordination of motor movements, led to positive results (Carmi, Tendler, et al.). There has not been any study prior that tested dTMS on other parts of the CSTC, specifically the mPFC and ACC; however, this study showed that stimulating the ACC also leads to a decrease in symptoms of OCD for patients (Carmi, Tendler, et al.).

As it is not fully agreed on how TMS should be administered to patients, the most effective frequency of TMS stimulation is also unknown (Carmi, Alyagon, et al.). The debate arises from how the CSTC is overactive in patients with OCD, so a low-frequency (LF) TMS may be effective due to how LF TMS is thought to produce inhibitory effects on the brain (Carmi, Alyagon, et al.). However, high-frequency (HF) TMS produces excitatory effects and can disrupt activity in the CSTC region (Carmi, Alyagon, et al.). Plus, HF TMS can have more long-term effects, as shown by how, when induced, it has had lasting effects on mice models'

neuroplasticity, which refers to the brain’s ability to reorganize its structure and function in response to experiences and environmental stimuli (Carmi, Alyagon, et al.). This enhanced neuroplasticity leads to an increase in learning and memory consolidation, improving cognitive restructuring, behavioral modification, and emotional regulation, which will significantly ease the process of recovering from OCD-related habits and symptoms for patients (*Neuroplasticity - StatPearls - NCBI Bookshelf*).

So, to see which range of frequencies is more ideal for treating OCD, another study by Carmi et al. aimed to test both high and low frequencies of TMS, specifically dTMS, using the H7 coil on the mPFC and ACC. 41 patients, who were unresponsive to 2 SRI trials and CBT, were chosen at random to be administered TMS at either 1 Hz (LF), 20 Hz (HF), or they received a sham treatment; every group was given treatment 5 times a week for a total of 5 weeks. Patients chosen for this study were also between 18 and 65 years old and had a Y-BOCS score of at least 20 and above. The device that administered the TMS was placed 4 centimeters anterior to the leg motor cortex, the part of the brain responsible for voluntary control of the legs. The HF treatment was administered at 100% of the leg RMT and consisted of 50 trains of magnetic pulses; each train had an interval of 2 seconds, with an intertrain interval of 20 seconds, giving a total of 2000 pulses per day. The LF treatment was administered at 110% of the leg RMT and consisted of 900 consecutive pulses, and the sham treatment was given in a manner that randomly mimicked the HF or the LF treatment. After the experiment, it was found that both the HF and LF treatments did not cause any adverse side effects in any of the patients, only mild headaches either during or after the treatment was administered (Carmi, Alyagon, et al.). However, it was found that the HF treatment led to a significantly higher response rate in patients compared to those given the sham or LF treatment, meaning that the patients who received the HF treatment generally had a greater reduction in OCD-related symptoms than the other patients, as proven through reductions in their Y-BOCS score (Figure 3) (Carmi, Alyagon, et al.). This study also reaffirmed the notion among previous scientific studies that the sham treatment is not as effective as the active treatments (HF and LF) (Carmi, Alyagon, et al.).

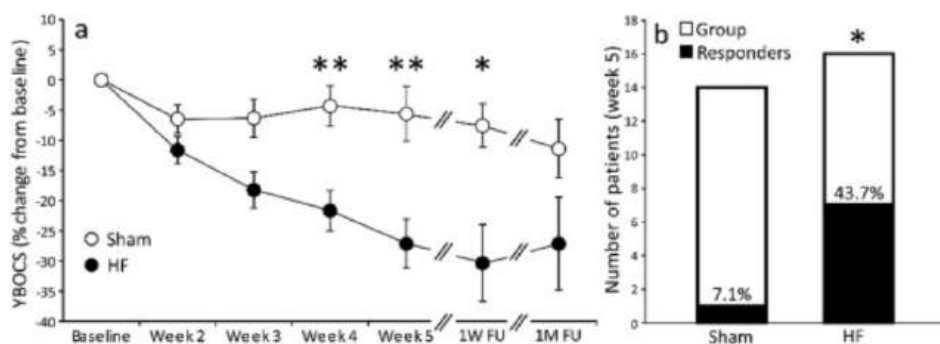


Fig 3: (A) The periodic change in Y-BOCS scores in patients when receiving HF or sham treatment. (B) The amount of patients responding to HF treatment versus sham treatment (Figure taken from Carmi et al., 2017).

Altogether, TMS is an effective third option for OCD treatment for patients who are drug-resistant or unresponsive to cognitive-behavioral therapies. This is particularly shown by how the active treatments are effective overall and are more efficacious than the sham treatment, proving that the effects patients experience from TMS are not a placebo and that giving TMS will lead to more desirable results for the patient (Carmi, Tendler, et al.; Carmi, Alyagon, et al.). Plus, it is safe for usage as none of the patients in both experiments had any major side effects when receiving the treatment (e.g., epilepsy); the most that the patients experienced were temporary minor headaches (Carmi, Tendler, et al.; Carmi, Alyagon, et al.).

Theta Burst Stimulation

Theta Burst Stimulation (TBS) is a type of repeated TMS (rTMS), a type of TMS that sends brief recurring magnetic pulses to a specific region of the brain (Mann and Malhi; Oberman et al.). More specifically, it revolves around delivering magnetic pulses in bursts of three at a frequency of 50 Hz (meaning there were 50 pulses per second); these bursts are separated by a short interval lasting 200 milliseconds (Oberman et al.). The merit behind this is that these specific protocols for administering TBS, in comparison to rTMS, are more aligned with the natural rhythm of brain activity, making it have the potential to enhance the process of long-term potentiation (LTP) (Jiang et al.). LTP essentially strengthens the connections between neurons, improving communication in a neuronal network (*Long-Term Potentiation - an Overview | ScienceDirect Topics*). An increase in LTP can be beneficial in treating OCD since patients with OCD are found to have defects in memory, attention, and executive functions (due to impaired synaptic plasticity, affecting the brain's ability to form and maintain new memories, leading to the persistence of obsessive thoughts and compulsive behaviors) (Jiang et al.). One type of executive function that is found to be defective in patients with OCD is response inhibition (RI), which is directly responsible for the brain's ability to prevent undesirable and inappropriate thoughts and behaviors (Guo et al.). So, using TBS to increase the strength and effectiveness of LTP can fix the aforementioned damaged executive functions via normalizing synaptic communication in the brain, making it a promising treatment for OCD (Jiang et al.; *Long-Term Potentiation - an Overview | ScienceDirect Topics*). Not to mention that, since TBS is a type of TMS, its treatment targets what is most associated with the onset of OCD: the hyperactivity of the CSTC. Therefore, TBS can be possibly seen as a more effective version of TMS, as it achieves similar outcomes while producing additional effects that could lower OCD symptoms (e.g., enhancing LTP in the brain).

In a study conducted by Guo et al., they tested the effects of using TBS, specifically continuous beta burst stimulation (cTBS), on patients with OCD. cTBS is a low-frequency version of TBS that produces inhibitory effects at the brain regions it targets (Guo et al.). During this particular experiment, they aimed at the SMA, a region of the brain found promising for targeting during TMS, since they found that targeting the mPFC, a brain region typically targeted for TMS, is ineffective for low-frequency stimulations like cTBS (Guo et al.). Plus, they also found that OFC, another ideal target for TMS, may be more difficult to aim at as it is close to the

brow bones, meaning that any stimulation must travel through overlying tissues (such as skin, muscle, and bone) to reach it (Guo et al.). The merit for targeting the SMA comes from how the hyperexcitability of the SMA, which leads to the loss of motor control, is shown to be a common defect in patients with OCD (Guo et al.). This loss of motor control is due to how the SMA has connections with the frontal lobe, the part of the brain responsible for executive functions (Guo et al.). So, researchers in this experiment hypothesized that the SMA may be the optimal area to direct magnetic pulses toward during cTBS (Guo et al.). In this study, there were 54 participants; they were all between 18-52 years old and scored between 16 and 31 on the Y-BOCS scale, making them have moderate to severe symptoms of OCD. They either received the active or sham treatment. For the active treatment, patients wore a device that placed a coil that was positioned anterior to the apex skull and located 15% of the way between the nasion and the occipital eminence. This treatment was administered on one side of the SMA at a time rather than having the cTBS target the SMA bilaterally. During the cTBS treatment, the intensity of the stimulus was at 110% RMT delivered in trains of three pulses for a total of 1200 pulses, which lasted for 48 seconds. In this instance, RMT is defined as the minimum intensity of stimulus needed for at least half of the single pulse trials to elicit a threshold electromyogram response (the activity produced by muscle fibers in response to nerve stimulation) of 50 mV in the contralateral abductor pollicis brevis when delivered to the contralateral primary cortex muscle. Each train occurred in sets of 3 pulses at 50 Hz and were repeated at 5 Hz (meaning that this burst cycle repeated 5 times per second). This led to every treatment containing 1200 pulses in total and lasting 48 seconds. Conditions were the same for the sham treatment; however, the coil was flipped 180° so that the device couldn't stimulate the brain. Both patients under the active and sham treatments received treatment 5 times a week for 4 weeks, meaning that patients received a total of 20 treatment sessions. They concluded that treatment with cTBS over the SMA could have the potential to relieve anxiety and depression in patients with OCD (Guo et al.). However, there was no significant evidence that the active treatment was more effective at reducing general symptoms than the sham treatment (Figure 4) (Guo et al.). This is primarily shown by the fact that the RI in patients did not improve significantly after cTBS (Guo et al.). The importance of RI as an indicator comes from how the hyperactivity of the SMA is negatively correlated with the RI, meaning that an increase in RI implies the lower activity of the SMA (which would subsequently lower symptoms of OCD) (Guo et al.). So, the RI not improving suggests that the cTBS did not effectively address the excess excitability of the SMA, leaving patients who received active cTBS without a significant increase in improvement of OCD symptoms in comparison to those who received the sham cTBS (Guo et al.) However, there were no recorded incidents of severe adverse events when cTBS was used on patients; plus, most patients recorded their treatment as somewhat satisfactory (Guo et al.). Therefore, cTBS is still a safe and acceptable form of treatment for OCD, but there is no significant difference in the efficacy between sham and active cTBS (Guo et al.).

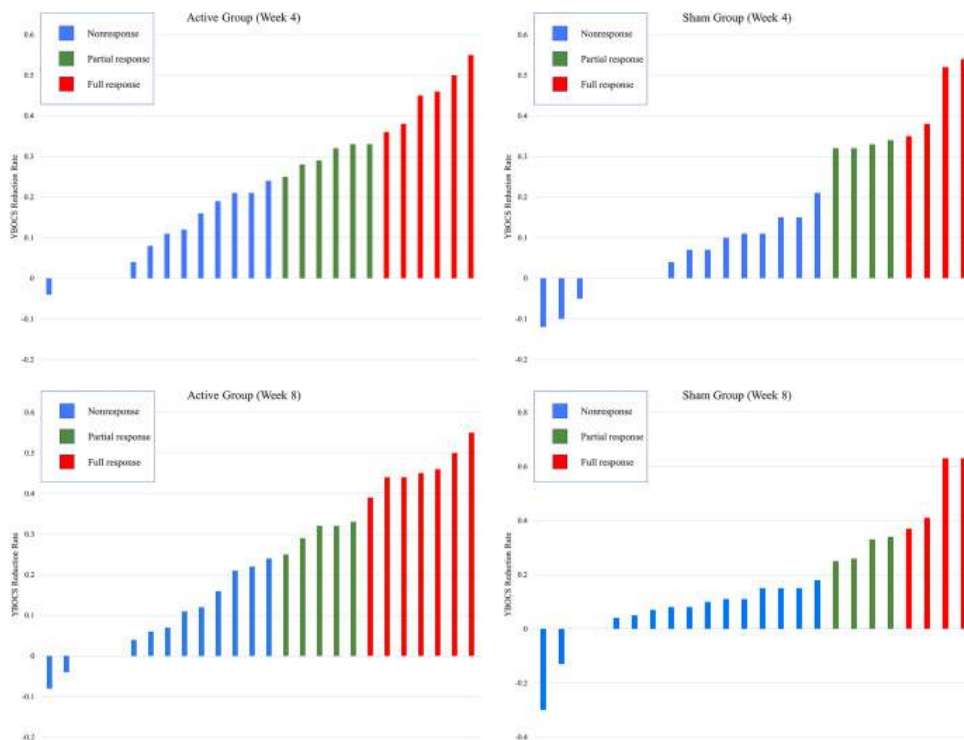


Fig 4: Individual distribution of patient response to sham and active treatments after 4 and 8 weeks (Figure taken from Guo et al., 2022).

A problem with traditional TMS is that treatment, specifically using rTMS, typically lasts about 4-6 weeks (Jiang et al.). This may pose a potential problem for many patients who cannot commit to or endure lengthy treatment durations such as this. So, in a study conducted by Jiang et al., they tested a newly developed form of TBS (accelerated high-dose TBS, ahTBS), whose treatment will take less time and result in longer-lasting effects for the patients it is used on (Jiang et al.). In this study, there were 45 patients who were between the ages of 18 and 45 and had a Y-BOCS score of 16 or higher (meaning that the patients had moderate to severe symptoms of OCD). Patients were either given ahTBS or 1 Hz rTMS for 5 consecutive days. Patients who received ahTBS were administered it to their right pre-SMA. In each session, they were given 1800 pulses in a continuous train of 600 theta bursts. Each theta burst consisted of 3 pulses with a frequency of 50 Hz, repeated at 5 Hz. Ten sessions were given a day, with 50-minute intervals between each session (equating to patients getting a total of 18,000 pulses a day). The ahTBS was given at 80% of the RMT, which is defined as the minimum intensity of stimulus needed for at least half of the single pulse trials to elicit a motor response of at least 50 μ V at the relaxed abductor pollicis brevis muscle when delivered to the contralateral primary cortex. The patients who received the 1 Hz rTMS were given 1800 TMS pulses to the right SMA every day of the treatment plan at 80% RMT. This study found that the Y-BOCS scores were significantly reduced in patients who received the ahTBS treatment and the traditional rTMS treatment (Figure 5) (Jiang et al.). There were, however, more individuals who had a significant decrease in

OCD symptoms when given ahTBS rather than rTMS; the difference is not enough to be considered significant, though (Figure 5) (Jiang et al.). Plus, the ahTBS treatment was well-tolerated by the patients who were administered it, so it is a safe and effective method for OCD (Jiang et al.). Since the results of those using ahTBS were similar to those using rTMS, ahTBS may be the next step in optimizing and improving the overall delivery of TMS (Jiang et al.).

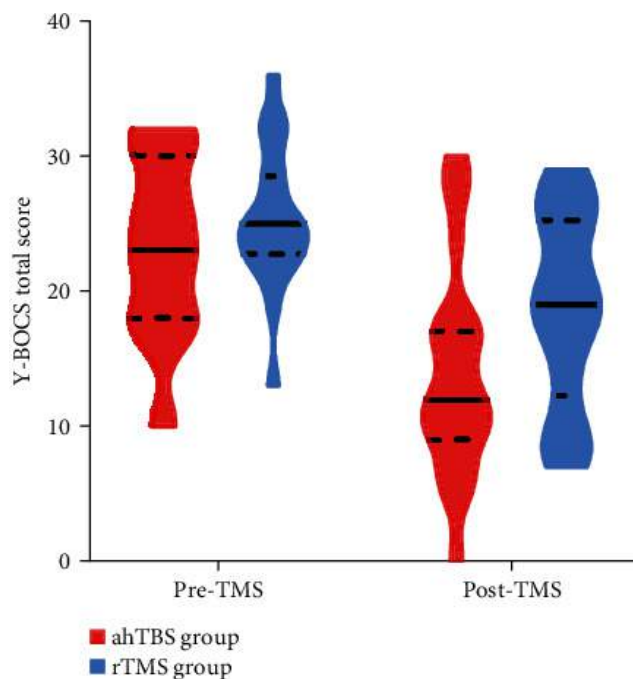


Fig 5: The change in Y-BOCS scores in patients after receiving ahTBS or rTMS (Figure taken from Jiang et al., 2023).

Overall, there were some mixed results regarding the efficacy of TBS. When specifically talking of cTBS, it is found to be somewhat effective in alleviating OCD symptoms (mainly by improving anxiety and depression in OCD patients) (Guo et al.). However, the outcomes did not differ significantly between using active cTBS and placebo (Guo et al.). In the case of ahTBS, even though it is not more effective than regular rTMS, it has been found that using ahTBS produces results similar to those of using rTMS (Jiang et al.). So, all in all, cTBS can be helpful as a supplementary therapeutic for OCD patients with more significant symptoms of anxiety and depression, but it is not as effective as a standalone treatment (Guo et al.). Moreover, ahTBS seems to be a promising alternative to rTMS for patients who can not physically, mentally, or financially afford to go through the weeks of treatment that rTMS requires of patients (Jiang et al.).

Deep Brain Stimulation

Deep Brain Stimulation (DBS) is another form of treatment, given via brain stimulation, designed for patients with refractory OCD (Graat et al.). Refractory OCD is defined as patients who do not see an improvement in OCD symptoms after taking at least three antidepressants for

three to six months, as well as attempting two atypical antidepressants in addition to their original treatment, essentially meaning that they are drug-resistant (Swierkosz-Lenart et al.). DBS, compared to many other neuromodulation techniques, is invasive and was initially introduced as an alternative to traditional surgery because it can be adjusted and reversed, based on patient response, if needed (Tyagi et al.). It works towards treating refractory OCD by surgically implanting two electrodes, conductive devices that deliver controlled electric stimulation to specific brain regions, to a subcutaneously implanted pulse generator (IPG) (Graat et al.). The electrodes used during DBS are thin, insulated wires inserted through a small opening in a patient's skull, which are then implanted into a specific brain region ("Deep Brain Stimulation"). These electrodes are connected to an IPG, which generates and delivers the electrical pulses to the electrodes via an insulated extension wire ("Deep Brain Stimulation"). The electrical pulses administered to the patient aim to control abnormal brain activity that is associated with OCD by directly targeting the brain regions primarily responsible for the onset of OCD symptoms ("Deep Brain Stimulation"). These pulses can also adjust the chemical imbalances in the brain that are also causing the occurrence of OCD symptoms (e.g., influencing the release of neurotransmitters, such as serotonin and dopamine, that are found imbalanced in patients with OCD) ("Deep Brain Stimulation").

DBS was initially developed and used for neurological disorders, primarily movement disorders; however, when it started being used for psychiatry-related purposes (such as OCD), it had, and still has, between a 2-30% risk for adverse side effects (adverse events, AE) like insomnia and mood swings (Graat et al.). These side effects depend on the voltage the pulses are given and the location the DBS treatment aims towards (Graat et al.). Yet, adjusting any of these stimulation settings can mitigate the effectiveness of DBS for the patient (Graat et al.). Plus, DBS, when used for treating OCD, requires a relatively high voltage, which lowers the typical lifespan of a non-rechargeable IPG battery (14 months) (Graat et al.). This leads to the administration of DBS being costly and burdensome as it requires the IPG to be frequently replaced (Graat et al.). There are rechargeable stimulators that can be used for DBS, but the time spent on recharging during treatment can lead to patients feeling this process is unnecessarily time-consuming (Graat et al.). So, researchers have been aiming to find a type of DBS that minimizes side effects and optimizes energy efficiency (for the IPG battery) to improve the experience of DBS for patients (Graat et al.).

Generally, DBS is given in a continuous stream of stimulation (known as continuous DBS); however, the IPG can be programmed to deliver stimulation in cycles, turning on and off at intervals (Graat et al.). This is known as cyclic DBS, and it has been found that the clinical effects of DBS can last longer and have more efficacious results for patients when administered as cyclic DBS rather than continuous DBS. (Graat et al.) Plus, cyclic DBS could have more efficient IPG battery usage and reduce the adverse side effects of regular DBS (Graat et al.). So, a research study was conducted by Graat et al. to test if cyclic DBS has the aforementioned benefits of reduced side effects and maximizing IPG battery efficiency. They also investigated whether or not cyclic DBS is as effective as continuous DBS in reducing symptoms of OCD. To

be included in this study, the patients had to have already been diagnosed with OCD and gone through the process of DBS, with a minimal 25% reduction in their Y-BOCS score after treatment in comparison to their initial score. There were a total of 16 patients in this study with a relatively equal gender ratio; the mean of their ages was around 50 years old. In this particular experiment, these patients were randomly given, in two-week blocks, either cyclic DBS first, then continuous DBS, or vice-versa (for a total of 4 weeks of active treatment). The electrical pulses of the DBS treatment were aimed at the ventral anterior limb of the internal capsule (vALIC). vALIC is a region of white matter in the brain that serves as a neural circuit involved in various motor, cognitive, and emotional functions as it connects the different parts of the CSTC, whose abnormalities have been found responsible for the onset of OCD symptoms (Safadi et al.). Specifically, an abnormal vALIC will lead to effects such as dysregulated communication within the CSTC circuit (causing symptoms like obsessive thoughts), disrupted inhibitory control (resulting in the inability to suppress intrusive thoughts or resist compulsive behaviors), and an imbalance in reward processing (leading to an overvaluation of compulsive behaviors as a means of reducing anxiety or distress associated with obsessions) (Safadi et al.). So, because the vALIC is a part of the CSTC, meaning that its disruption is a significant potential cause for OCD, it is a common spot for DBS during experiments such as this. After the patients were treated, the severity of their symptoms was measured using Y-BOCS to quantify the efficacy of cyclic DBS. Cyclic DBS, in this study, was found to increase the battery life of the IPG in comparison to continuous DBS (Graat et al.). However, it was also found to cause a relapse in OCD as most patients experienced a significant increase in OCD symptoms of 133% on average, with half of the patients experiencing a Y-BOCS score increase of at least 4 after receiving cyclic DBS (Figure 6) (Graat et al.). Not all patients experienced this increase, though, as half of the patients did not experience a significant increase in OCD symptoms after cyclic DBS, with even some patients experiencing a slight increase in their Y-BOCS score (although the length of these effects varied from patient to patient) (Graat et al.). Cyclic DBS was found to cause more AEs than continuous DBS (Graat et al.). Still, this difference was not statistically significant as some stimulation-related AEs were found to decrease in some patients with the usage of DBS (e.g., headaches, concentration problems, and loss of inhibition) (Graat et al.). Nonetheless, DBS remains an effective treatment for OCD, as patients quickly recovered from their temporary relapse and saw an improvement in their symptoms after switching from cyclic DBS to continuous DBS (Graat et al.).

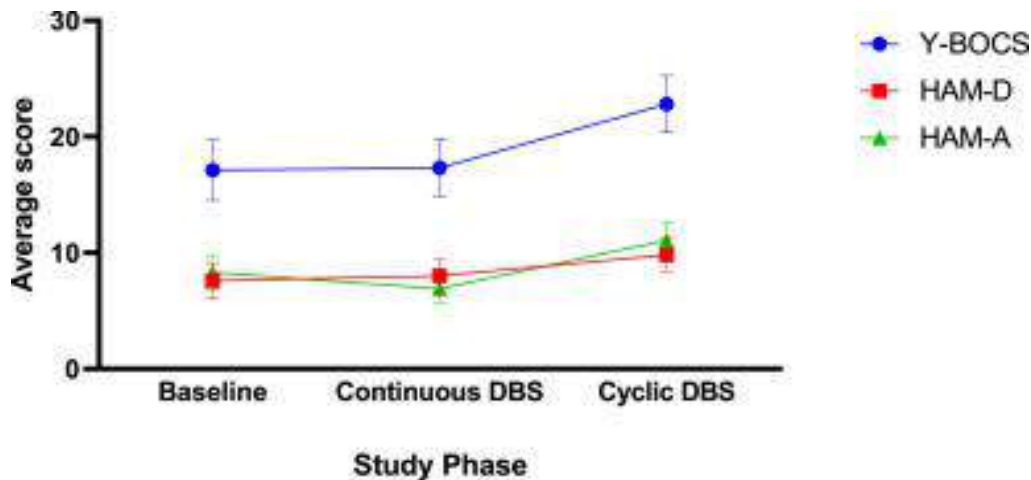


Fig 6: The change in Y-BOCS, HAM-D, and HAM-A scores in patients when receiving continuous or cyclic DBS (Figure taken from Graat et al., 2023).

The ideal location for targeting during DBS is still being researched. In addition to the vALIC, the ventral capsule/ventral striatum (VC/VS) and the anteromedial subthalamic nucleus (amSTN) have also been popular targets for TBS treatments (Tyagi et al.). This is from how both the VC/VS and amSTN are a part of the brain's basal ganglia, which is responsible for many functions such as motor control, habit formation, and reward processing and cognition, as it is a part of the CSTC (Senova et al.). Both the amSTN and VC/VS are responsible for motor control, but what differentiates them is that the amSTN is specifically more involved with the cognitive aspect of the basal ganglia; meanwhile, the VC/VS plays more of a role in reward processing and mood regulation (Senova et al.). Disruption of the amSTN can lead to impaired cognitive functions such as attention, decision-making, and cognitive flexibility (Senova et al.). This cognitive inflexibility is characterized by rigid thinking patterns and difficulty adapting to changing circumstances, contributing to obsessive thoughts and the deficit in RI found in OCD patients (Senova et al.). Disruption of the VC/VS can lead to reward circuitry dysfunctions and emotion regulation abnormalities, resulting in irregularity in incentive salience and heightened emotion reactivity, respectively, which both cause compulsive behaviors and obsessive thoughts and rituals commonly associated with OCD (Senova et al.). Essentially, since both the VC/VS and the amSTN are part of the CSTC, their dysregulation has been found to correlate with the rise of OCD, making them common targets in DBS. Studies have also found that stimulating the VC/VS influenced the mood of the patients while stimulating the amSTN modulated affected the patients' compulsive behaviors, suggesting that distinct neural circuits are involved in the symptomatology of OCD and the decrease of OCD symptoms (Tyagi et al.).

In a study conducted by Tyagi et al., they gave patients DBS to both the VC/VS and amSTN to investigate whether one location is more effective than the other in treating OCD. They also determined the precise, optimal location for DBS by calculating volumes of tissue activation (since it allows for direct comparison of which particular brain region target induces the most neuronal effects) and the validity of the aforementioned hypothesis of how targeting the

VC/VS versus the amSTN leads to the stimulation of completely different neural circuits. Another thing the researchers looked for was whether or not it is beneficial to go through CBT while taking DBS. There were 6 patients for this study, and they were all above 20 years old, had OCD for over 10 years (with continual severe symptoms for 2 years), had a Y-BOCS score above 32 (meaning that their symptoms were on the extreme end), and were drug-resistant (which was defined as having no benefit from at least 2 SRIs taken at optimal doses for at least 12 weeks, SRIs with antipsychotic medications, SRIs take above the optimal dose, or 2 trials of CBT for a minimum of 10 hours). During the study, patients first received DBS stimulation to either the amSTN or VC/VS for a period that lasted 12 weeks. They then received stimulation to the other brain region for the same amount of time. There were then two 12-week phases where DBS was given using optimized settings gathered from the data of the previous phases, followed by another period where patients received CBT, specifically exposure and response prevention CBT, while receiving the optimized DBS. Clinical and cognitive assessments, measured primarily with Y-BOCS, were given before and after each phase to test and quantify their specific effectiveness in treating OCD. Using Y-BOCS, this study found that stimulating the amSTN and VC/VS for DBS led to effective results for each brain region of equal magnitude (Figure 7) (Tyagi et al.). It also confirmed that targeting the different brain circuits led to different effects in patients, with the VC/VS DBS improving the mood of patients and the amSTN DBS improving cognitive flexibility in patients (Tyagi et al.). This, along with tractography (a 3D modeling technique used to represent nerve tracts using data collected by diffusion MRI visually) from the activated electrode contacts at the sites of DBS, suggest that the stimulation of VC/VS and amSTN led to the modulation of completely distinct brain networks (Tyagi et al.). Using the volume of tissue activation, this study also found that targeting the VC is what mainly leads to the clinical effect of increasing mood and decreasing OCD symptoms overall rather than the VS, as the most effective electrode contacts were the ones that were more dorsal (higher up) in the brain (the VC is more dorsal than the VS) (Tyagi et al.). Plus, they observed that utilizing CBT during DBS does not boost the beneficial effects of DBS mainly because most patients were already at mild Y-BOCS levels after DBS, making it difficult for further significant improvements to be made via DBS (Tyagi et al.). The adverse effects of DBS on both the VC/VS and amSTN were mainly hypomania induced by the stimulation; however, these effects did not occur more frequently when either site was stimulated (Tyagi et al.).

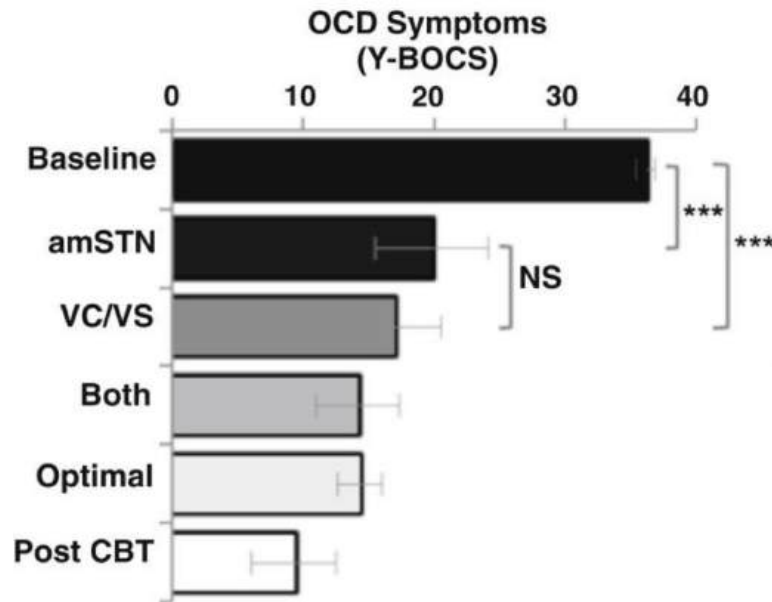


Fig 7: The change in Y-BOCS score in patients after receiving amSTN, VC/VS, combined amSTN and VC/VS treatment, optimal combined settings, and CBT (Figure taken from Tyagi et al., 2019).

In general, DBS is an effective treatment for OCD, as the majority of the patients in both of these studies experienced a decrease in OCD symptoms from some type of DBS treatment (Graat et al.; Tyagi et al.). However, it is not suggested to specifically use cyclic DBS as it results in patients experiencing relapses (Graat et al.). Even though it conserves the battery of the IPG and can reduce some common AEs, the risk of relapse outweighs the potential benefits cyclic DBS may bring compared to continuous DBS (Graat et al.). Regarding locations for DBS, the vALIC, VC/VS (specifically VC), and amSTN were all found to be effective targets (Graat et al.; Tyagi et al.). It is important to note that DBS, although it has continuously been found to be a successful therapeutic for OCD, can lead to patients experiencing adverse effects (the magnitude and probability of which can vary greatly between patients) (Graat et al.; Tyagi et al.).

Transcranial Direct Current Stimulation

Transcranial Direct Current Stimulation (tDCS) is another main non-invasive neuromodulation technique (a process of altering the activity and connectivity within certain brain networks by delivering stimuli directly to those regions) used to treat refractory OCD (Harika-Germaneau et al.). In the case of tDCS, it modulates brain activity and excitability via sending a weak direct electrical current to a specific part of the brain thought to be responsible for OCD symptoms (e.g., SMA) (Alizadehgoradel et al.). This electric current is delivered through two flat, large electrodes placed on a patient's scalp (Harika-Germaneau et al.). Out of these electrodes, there is an active and reference one (DaSilva et al.). The active electrode, placed directly on top of a cortical region (a specific part of the cerebral cortex crucial for many higher brain functions such as thought, memory, and decision-making), is the one responsible for

being the medium where the electrical current enters the brain tissue in the cerebral cortex via the skull (Harika-Germaneau et al.; DaSilva et al.). Meanwhile, the reference electrode, placed in an extracephalic location (a location not on the head), acts as a stable point of reference, completing the circuit between the active electrode and the tDCS machine (DaSilva et al.). This ensures the electrical current is stable and consistent, protecting the patient undergoing tDCS from harm by preventing adverse events such as concentrated currents, which can lead to damaged tissue (DaSilva et al.).

The active electrodes are classified as anodal or cathodal depending on their polarity (Harika-Germaneau et al.). More specifically, anodal stimulation involves applying a positive electrical current to a target area of the brain, depolarizing the neuronal membranes, which causes the inside of the neurons to be less negative than the outside (Harika-Germaneau et al.; Alizadehgoradel et al.). This causes a shift in the membrane that brings the neurons in it closer to the threshold needed to fire an active potential (a brief and rapid electrical signal that travels along the membrane of a neuron), meaning that anodal stimulation increases neuronal excitability (Harika-Germaneau et al.; Alizadehgoradel et al.). It is also associated with an increased regional cerebral blood flow in the stimulated cortical area (Harika-Germaneau et al.). This is because anodal stimulation causes enhanced neuronal activity, which increases the already high metabolic demands of the neurons (Pulgar). So, to communicate their demands, the neurons release signaling molecules that cause the local blood vessels in the area to dilate (vasodilation)(Pulgar). This vasodilation is what causes the blood flow to active neurons to increase, fulfilling the neurons' metabolic needs (Pulgar). Conversely, cathodal stimulation involves applying a negative electrical current to a target area of the brain, which hyperpolarizes the neuronal membranes (Harika-Germaneau et al.; Alizadehgoradel et al.). This causes the membrane to shift its neurons further from the threshold needed to fire an active potential, which means that cathodal stimulation decreases neuronal excitability. Harika-Germaneau et al.; Alizadehgoradel et al.). Also, it is associated with a decreased regional cerebral blood flow in the stimulated cortical area because of how it decreases neuronal excitability, subsequently lowering the metabolic needs of the neurons in that region (Harika-Germaneau et al.; Pulgar).

tDCS's ability to modulate cortical excitability and promote neuroplasticity (through said modulation in cortical excitability) is what sparked researchers to study it more in-depth due to its implied potential to enhance the brain's ability to recover its function (Harika-Germaneau et al.). It is thought that tDCS, through adaptive changes in neural circuits and the formation of new connections (made increasingly possible by increased plasticity) due to the electrical currents, may help restore normal brain function in individuals with psychiatric disorders like OCD (Harika-Germaneau et al.). Plus, compared to other neuromodulation techniques commonly used, tDCS is relatively simple and low-cost (Harika-Germaneau et al.).

However, it is still not exactly known what the most optimal targets are for tDCS, so a study by Harika-Germaneau *et al.* aimed to further understanding on this front as well as also to test whether the effects of tDCS can be long-lasting in patients after completion of the initial treatment (1-3 months)(Harika-Germaneau et al.). This study specifically tested to see whether

using cathodal stimulation over the SMA and anodal stimulation over the right supraorbital area would be efficient in lowering OCD symptoms. The hyperexcitability of the SMA has been found to lead to OCD symptoms, as it is a part of the CSTC and, therefore, responsible for functions such as the execution of motor movements and cognitive control (Harika-Germaneau et al.). So, it is thought that using cathodal stimulation will lower the SMA's excitability, normalizing it from its previous state of hyperactivity (Harika-Germaneau et al.). The right supraorbital area is a part of the DLPFC, a part of the CSTC, and it is responsible for cognitive flexibility, decision-making, and response inhibition, which is often found to be abnormal in patients with OCD (Harika-Germaneau et al.; Li et al.). So, modulation of the right supraorbital area via anodal stimulation is thought to enhance the DLPFC, which can help improve cognitive control and regulate emotional processing (as the DLPFC is also involved in regulating emotional responses and modulating amygdala, a brain region used in emotional processing, activity) (Harika-Germaneau et al.; Li et al.). Improving cognitive control and regulating emotional processing can lead to OCD patients being able to lower symptoms related to compulsive behaviors and intrusive thoughts, as well as reduce the intensity of their obsessive thoughts and OCD-induced anxiety (Harika-Germaneau et al.; Li et al.).

There were 21 patients in this study, and they received a tDCS session once a day, 5 days a week, for a span of 2 weeks, meaning that they received a total of 10 tDCS sessions. These patients were between the ages of 18 and 70 years old, and they had to have a Y-BOCS score of at least 20 (making them have moderate symptoms) for a span of 2 years; they also had to be treatment-resistant, which this study classified as when patients do not respond to at least 2 12-week treatments of SRI and CBT. A session of the tDCS involved the patients receiving a direct current of 2 mA for 30 minutes via electrodes. The cathodal electrode was placed in a location that was intended to hit the SMA bilaterally (the sagittal midline, approximately 15% of the distance from theinion to the nasion, in a direction toward the front of the head, anterior to the Cz location [a specific location on the scalp typically used as a reference point for neuromodulation techniques]). The anodal electrode was placed over the right orbitofrontal area. The outcomes of the tDCS treatments were measured using differences in the patients' total Y-BOCS scores. This study found that tDCS, specifically 10 sessions of cathodal tDCS over the bilateral SMA (since the anodal electrode over the right supraorbital area was used as an "inert" reference electrode), significantly reduced OCD symptoms in treatment-resistant patients with little adverse side effects, as shown from how patients had at least a 35% reduction in their Y-BOCS scores (Figure 8) (Harika-Germaneau et al.). However, the study said the effects could also have been from stimulating the right supraorbital area (Harika-Germaneau et al.). So, cathodal stimulation over the SMA and anodal stimulation over the right supraorbital area with bipolar tDCS, a type of tDCS where both electrodes are active and placed on the scalp (meaning that there is no polarity distinction between the two electrodes), is a safe and effective way of treating OCD; this makes bipolar tDCS a valid form of tDCS for widespread clinical usage (Harika-Germaneau et al.).

Since bipolar tDCS was found to be an effective form of tDCS without any polarity distinction between its electrodes, it suggests that its clinical effect of tDCS is not polarity-dependent (Harika-Germaneau et al.). This brings up the possibility of the effects of tDCS not being limited to the anodal and cathodal sites, further proven by research finding that tDCS can generate electric fields deep inside the brain, which shows that its effects may not be limited to superficial cortical regions (Harika-Germaneau et al.). On top of successfully modifying the anodal and cathodal sites and potentially other regions deeper in the brain, tDCS was also found to maintain its therapeutic effects for 1-3 months in multiple patients (Harika-Germaneau et al.). Patients also saw themselves having improvements in their depression symptoms (Harika-Germaneau et al.). Still, researchers of the study found that a reduction in someone's Y-BOCS score did not correlate with their reduction in depression, suggesting that tDCS has a secondary effect on depression when treating OCD (Harika-Germaneau et al.).

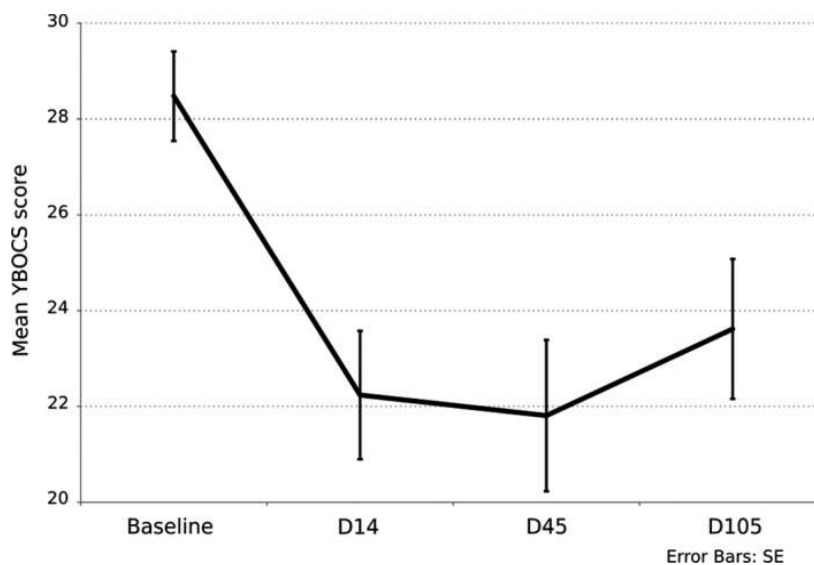


Fig 8: The periodic change in Y-BOCS scores in patients after 10 tDCS sessions. D stands for day. (Figure taken from Harika-Germaneau et al., 2020).

Along with the fact that there is no known ideal location for tDCS, it is also not known which parameters are best for administering this type of brain stimulation (e.g., the intensity of the electrical currents) (Alizadehgoradel et al.). Plus, there has not been much research on the effects of tDCS on core cognitive and electrophysiological correlates of the psychopathology of OCD, such as response inhibition and functional connectivity (Alizadehgoradel et al.). So, a study conducted by Alizadehgoradel et al. aimed to discover these three tDCS properties. In this experiment, there were 39 patients (with a relatively higher number of women than men), who were between the ages of 18 and 50 years old and were diagnosed with OCD without previous history of other neurological diseases, brain damage, or psychiatric disorders. They were all randomly assigned to active and sham stimulation groups. Each patient, both in the active (which

further consisted of a 1-mA and 2-mA group to test which intensity is best for tDCS) and sham groups, received 2 sessions of stimulation a day for 5 days, with 20-minute intervals between sessions, for a total of 10 sessions. To optimize the effects of tDCS, these stimulation sessions took place between 11:00 A.M. and 2:00 P.M. to avoid external effects caused by circadian rhythm and sleep pressure, which can negatively affect the factors of the brain that tDCS is targeting to improve, such as neuroplasticity and cognitive performance (Alizadehgoradel et al.). To generate the same sensation during the sham treatment as the active treatment, the electrical current was ramped up and down for 30 seconds and then turned off. For both the active and sham tDCS treatments, anodal and cathodal electrodes were placed on regions of the scalp over the left DLPFC and right pre-SMA, respectively. The left DLPFC was chosen as one of the targets because functional abnormalities of the DLPFC, such as response inhibition and decision-making, are often found in OCD patients, as the DLPFC is responsible for cognitive control (Alizadehgoradel et al.; Li et al.). So, putting the DLPFC under anodal stimulation would increase the region's activity, helping patients restore cognitive abilities that were hindered by the abnormality of their DLPFC (Alizadehgoradel et al.; Li et al.). The pre-SMA was chosen since it, which is responsible for motor control and cognitive processes, is often found to be hyperactive in OCD patients, especially when they have to do tasks that require attention and inhibitory control (Guo et al.; Alizadehgoradel et al.). So, using cathodal stimulation for the pre-SMA is supposed to counter the hyperactivity and allow the patient to regain control and function of some of their core motor control and cognitive processes (Guo et al.; Alizadehgoradel et al.). Plus, they were chosen to keep a minimum 6-centimeter distance between the edges of the electrodes, ensuring the patient's safety while preventing interference with the electrodes or disruption of the electrical current (Alizadehgoradel et al.; DaSilva et al.). To measure the outcome of the experiment, the Y-BOCS scale was used, and to measure the effect tDCS has on the core cognitive and electrophysiological correlates of the psychopathology of OCD, performance tasks, and data from monitoring patients' resting EEG were used, respectively. The study found that when stimulating the pre-SMA and the left DLPFC, both the 1-mA and 2-mA treatments were effective in reducing symptoms of depression (Figure 9)(Alizadehgoradel et al.). Still, it was ultimately the 2-mA treatment that was found to reduce OCD symptoms and anxiety more significantly in comparison to the effects of the 1-mA treatment (Figure 9) (Alizadehgoradel et al.). Both of the active treatments partially improved response inhibition (Alizadehgoradel et al.). However, patients who received 2-mA tDCS also saw improved working memory performance (the efficiency and effectiveness with which an individual can temporarily store and manipulate information for cognitive tasks) and reduced attentional bias to threat-related stimuli (Alizadehgoradel et al.). Overall, from observations of performance and behavioral tasks primarily related to cognitive control and flexibility often found impaired in OCD patients (e.g., response inhibition and working memory), improvements were seen in how the patients performed on these tasks after any type of active intervention (Alizadehgoradel et al.). However, results were more obvious in patients undergoing the 2-mA treatment (Alizadehgoradel et al.). Still, the study noted that these improvements, especially in response

inhibition, were more minor than expected, potentially because the right prefrontal region (which was not directly aimed at in the experiment) may have higher relevance in cognitive inhibition than expected (Alizadehgoradel et al.).

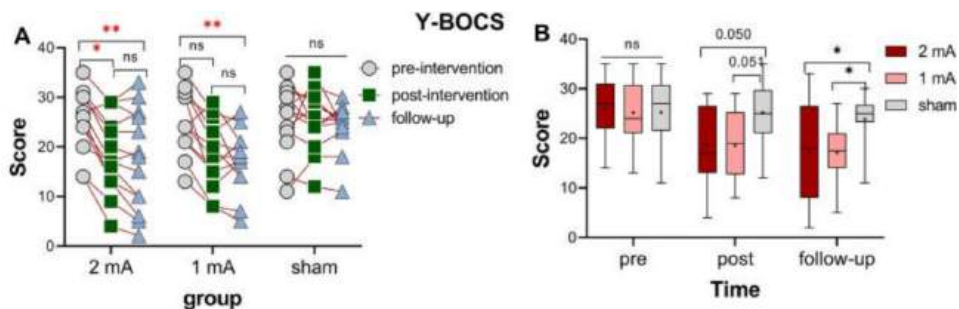


Fig 9: The periodic change in Y-BOCS scores when patients received 2 mA, 1 mA, or sham treatments. (Figure taken from Alizadehgoradel et al., 2024)(Alizadehgoradel et al.)

Both 1-mA and 2-mA tDCS also increased alpha oscillatory power in the brain, and the 2-mA decreased delta oscillatory power (Alizadehgoradel et al.). These oscillations represent different patterns of brain activity (Alizadehgoradel et al.). Specifically, alpha oscillations are associated with inhibitory processes and attentional control; therefore, abnormalities in alpha oscillations may reflect deficits in inhibitory control, a common problem in many OCD patients (Alizadehgoradel et al.; Ippolito et al.). So, the 1-mA and 2-mA treatments increase alpha oscillatory power, leading to strengthened inhibitory control mechanisms and improved attentional processes, helping people with OCD better regulate their thoughts and behaviors (Alizadehgoradel et al.; Ippolito et al.). Meanwhile, delta oscillations are associated with deep sleep and unconsciousness, so an excess of this during wakefulness may lead to sleep disturbances and abnormalities in arousal (the state of physiological or psychological activation or readiness of an organism) regulation, which is also commonly found in OCD patients as a dysfunctional arousal regulation can lead to symptoms such as heightened levels of anxiety and arousal (Alizadehgoradel et al.; *Delta Wave - an Overview | ScienceDirect Topics*; De Lecea et al.). These heightened levels of anxiety and arousal are what can lead to obsessive thoughts, leading to the need to perform compulsive behaviors (Alizadehgoradel et al.; *Delta Wave - an Overview | ScienceDirect Topics*; De Lecea et al.). So, the 2-mA treatment decreasing delta oscillatory power helps regulate arousal levels and promotes wakefulness (Alizadehgoradel et al.; *Delta Wave - an Overview | ScienceDirect Topics*; De Lecea et al.). This alleviates symptoms of excessive sleepiness or fatigue often reported with OCD, as well as combats common OCD symptoms via effects associated with higher levels of awakesness, such as cognitive engagement, which may help patients better resist OCD-related urges and instead implement coping strategies more effectively (Alizadehgoradel et al.; *Delta Wave - an Overview | ScienceDirect Topics*; De Lecea et al.). Furthermore, increasing alpha oscillations and decreasing delta oscillations lead to increased connectivity in higher-frequency bands at frontal-central areas in the brain (Alizadehgoradel et al.). Connectivity refers to the degree to which different brain regions

communicate or synchronize their activity, so higher connectivity indicates strong communication or coordination between various brain regions (Alizadehgoradel et al.; Lang et al.). In this instance, the brain regions are specifically the front-central areas of the brain, which are involved in higher-order functions, including cognitive control, decision-making, and response inhibition (Alizadehgoradel et al.; *Frontal Lobe: What It Is, Function, Location & Damage*). The higher frequency bands refer to how the oscillations connect at higher frequencies, such as beta and gamma oscillations, which are associated with cognitive processes such as attention, working memory, and executive function (Alizadehgoradel et al.; Uhlhaas and Singer). So, higher connectivity between front-central brain regions at higher frequency bands causes improved coordination and communication between brain regions involved in cognitive control and executive function, enhancing the brain's cognitive control mechanisms, which leads to individuals better regulating their OCD-related symptoms (Alizadehgoradel et al.; Lang et al.; *Frontal Lobe: What It Is, Function, Location & Damage*; Uhlhaas and Singer). Plus, because of strengthened connectivity in the front-central regions, individuals could also process and integrate information more efficiently and effectively, allowing them to manage their thoughts and behaviors better (Alizadehgoradel et al.; *Frontal Lobe: What It Is, Function, Location & Damage*).

This study found the 2-mA dTCS, specifically 2-mA intensified dTCS (meaning that the patients received the treatment twice per day with 20-minute intervals), to be more effective than both the sham and 1-mA stimulation for overall improvement of OCD symptoms as well as improvements in the core cognitive and electrophysiological correlates of the psychopathology of OCD (Alizadehgoradel et al.). The treatment's intensified protocol and higher frequency aspect are thought to be more effective due to how it leads to a greater increase in cortical excitability and longer-lasting aftereffects (Alizadehgoradel et al.). Also, the fact that the active dTCSs had a greater effect than the sham treatment further demonstrates that dTCS has a noticeable positive effect in patients (Alizadehgoradel et al.). Overall, dTCS, specifically 2 mA intensified dTCS, is an effective and safe treatment for OCD, as patients in the study tolerated the stimulation well with no reports of adverse effects from undergoing the treatment (Alizadehgoradel et al.).

Altogether, dTCS is a cost-efficient, simple, and effective brain stimulation method for the lowering of OCD symptoms (Harika-Germaneau et al.; Alizadehgoradel et al.). More specifically, higher frequency dTCS (2-mA) treatments were found to produce the best results for patients, not only in terms of the patient's OCD symptoms but also in the core cognitive and electrophysiological aspects of the patients' mind that are often found negatively affected by the onset of OCD (Alizadehgoradel et al.). It was also found that bipolar dTCS was found to be just as effective and safe in treating treatment-resistant patients as other "polar-dependent" types of dTCS and that the effects can last in patients from 1 month up to 3 months (Harika-Germaneau et al.). The most common targets for dTCS that led to successful results were the DLPFC and SMA, or at least regions that have some relation to those two, as they are both major parts of the CSTC (whose abnormality is thought to be the main cause of OCD arising in patients), making

them common targets for other neuromodulation techniques (Harika-Germaneau et al.; Alizadehgoradel et al.).

Conclusion

As a whole, brain stimulation is a valid, safe, and effective method of treating symptoms associated with OCD. This implies that most of the different types of neurostimulation methods brought up in this paper could be used as the automatic “third option” when patients are unresponsive to SRIs and CBT.

Brain stimulation generally aims to modulate brain activity and some properties of the brain (e.g., neuronal excitability and plasticity) found abnormal due to OCD by stimulating, via the use of currents (e.g., magnetic and electric), the brain regions most associated with the onset of OCD (*Brain Stimulation Therapies - National Institute of Mental Health (NIMH)*). This modulation is theoretically supposed to lower OCD symptoms because it directly affects the regions causing the issues for the patient (*Brain Stimulation Therapies - National Institute of Mental Health (NIMH)*). In the case of TMS, TBS, DBS, and tDCS, the studies that experimented with them all targeted some parts of the CSTC (the brain network responsible for functions such as reward, movement, and habit formation) due to its hyperactivity leading to patients developing OCD (Maraone et al.). This onset of OCD from the CSTC’s hyperactivity is due to how its dysfunction primarily leads to the interference of the balance between the goal-directed system and the habit-learning system of the brain, with overreliance on the latter (Maraone et al.). The goal-directed system allows for flexible and deliberate behavior changes in light of new situations a person finds themselves in, while the habit-learning system promotes repetitive behaviors triggered by specific cues of contexts (Maraone et al.). So, when the balance between these two systems is disrupted, specifically with more reliance on the habit-learning system, a person increasingly responds to everyday situations in life more rigidly and repetitively (Maraone et al.). This leads to a positive feedback loop that causes this person to perform their compulsions more as their responses to certain stimuli get more ingrained as a habit due to their habit-learning system (which is supposed to help them evaluate the irrationality of their compulsions and subsequently adapt their behaviors) being less effective (Maraone et al.). As a result, now that their compulsions are a habit, it is harder for them to modify their behaviors based on immediate feedback or changing circumstances, which means that they are also more susceptible to their obsessions “forcing” them to carry out said compulsions (Maraone et al.). So, targeting the CSTC for these brain stimulation methods was supposed to restore this balance and the functions of the network were found impaired by OCD (Maraone et al.). However, it is not exactly agreed upon which part of the CSTC is most effective for the treatment of OCD, leading to studies testing the effectiveness of targeting all sorts of brain regions (e.g., SMA, DLPFC, ACC, etc.). For the most part, all of the different regions that were used for these neuromodulation techniques were found to have positive results when it came to lowering OCD symptoms. It is important to note, though, that the results for targeting the SMA for cTBS, although effective for lowering anxiety and depression-related symptoms in OCD patients, were

similar to those for using a placebo.(Guo et al.) However, targeting the SMA for dTCS was found to be effective in improving patients with OCD, so the reason behind these findings is most likely not because of the region targeted (Harika-Germaneau et al.). Another thing to take into consideration is how the use of cyclic DBS on the vALIC led to some patients experiencing relapses of OCD symptoms, but this is most likely due to the method again, as the vALIC was found to be a valid target during other DBS experiments (Graat et al.).

These different therapeutic methods all arose due to how there is also not a “universal” agreement on which technique is the most effective in treating OCD. All of them essentially perform the same basic function of sending stimuli to a particular part of the brain to modulate its activity found abnormal due to OCD (*Brain Stimulation Therapies - National Institute of Mental Health (NIMH)*). However, they are differentiated in how they execute this function, with each technique either improving on an aspect of another technique or leading to additional therapeutic effects that further benefit the patient. An example of the former is how cyclic DBS was created as an alternative method to continuous DBS that would more efficiently use the IPG’s battery, which, in turn, would save money and improve the treatment experience for both the patient and the administrator (Graat et al.). An example of the latter is how TBS (a type of TMS) was used as a more “effective” treatment for OCD than TMS because of its potential to enhance LTP (due to its specific protocols for administering magnetic pulses to the brain) (Oberman et al.; Jiang et al.). This theoretically allows it to repair impaired executive functions and normalize brain activity of a specific brain region.(Jiang et al.) One drawback of brain stimulation as a treatment method is the adverse side effects that come with it. For most of the techniques, the side effects on patients are minimal to none (the most common being a mild headache), which is why brain stimulation is generally considered to be safe. However, every patient differs, so their magnitude and probability of experiencing a more-than-mild effect varies greatly. The side effects that a patient can undergo also heavily depend on the technique they use for their treatment. For example, the most prevalent side effect of using TMS was temporary minor headaches, while the most prevalent effect of DBS was hypomania (Carmi, Alyagon, et al.; Tyagi et al.). Another drawback is the large time commitment that comes with going through a brain stimulation treatment plan. Most require patients to come into the clinic and receive treatment many days a week for multiple weeks or even multiple months (which is why treatment methods have been created to shorten this treatment period, such as ahTBS) (Jiang et al.). Like the side effects that can come with brain stimulation, the amount of time a patient has to commit to getting treated varies a lot between each method, though. For instance, at the very least, the aforementioned ahTBS treatment lasts around 5 days, while, at the very most, a DBS treatment can last up to 24 weeks (which translates to around 6 months) (Jiang et al.; Tyagi et al.).

While brain stimulation techniques have been researched and used for a while to treat various psychiatric and neurological disorders, such as Parkinson’s disease and dystonia, their application to treating OCD has only been thoroughly explored more recently (Schermer). This implies that a lot of the methods that have been previously discussed, especially the newer ones

like TBS and tDCS, are still mainly in their experimental/trial phase. So, overall, all the neurostimulation methods, including the sub-methods (like how bipolar tDCS is a type of tDCS), need to be researched more to confirm their efficacy and safety further for patients with refractory OCD. Plus, since it is not still fully known which brain regions are the most effective targets for neurostimulation or even which technique is the most effective for treating OCD in general, there needs to be more comparative studies. More specifically, there should be 2 types of comparative studies: a direct comparison of two or more brain stimulation techniques (by testing their effectiveness when they target a specific brain region) and a direct comparison of a brain stimulation technique (ideally the one found to be the most effective from the previous type of study) on multiple commonly targeted brain regions. The results from these studies will, in turn, give the scientific community more clarity on what makes brain stimulation of OCD effective as a whole since they would be able to narrow down what exactly does and does not lead to the best results. From knowing the precise mechanisms of the “ideal” OCD brain stimulation treatment, scientists would then be able to further improve the treatment in a way that will still keep what made it effective in the first place but improve on the experience of administering it for both to administer and patient (e.g., lowering costs and the time commitment needed to receive the treatment). There should also be studies that follow the patients who underwent any sort of brain stimulation treatment for the next couple of years, and even decades, in order to detect any common short-term and long-term side effects of neuromodulation. These studies, in particular, will indicate if it is ultimately worth further researching and developing brain stimulation techniques as a therapeutic for OCD since following patients long-term could truly indicate how lasting the beneficial effects of the treatment are and if any major health defect could occur from it as they age (e.g., chronic seizures and strokes). Furthermore, it is important to further explore the treatments found to be not as effective as others, such as cTBS and cyclic DBS, to confirm if it is a method not to be used on patients and also to learn the exact aspects of them that led it to not producing the positive results that it was previously theorized to do.

Brain stimulation has already been found to be a promising alternative form of OCD treatment, but further research, development, and action in the aforementioned areas will help ensure that it can be as effective and accessible as it can be for all patients with refractory OCD. It has the potential to give everyone negatively affected by OCD hope for recovery and, ultimately, hope for a life without the burden that comes with having this particular disorder. Plus, with further development and improved accessibility, it even has the possibility of being the “first option” of treatment in the future, instead of SRIs or CBT, for refractory OCD.

Works Cited

- Alizadehgoradel, Jaber, et al. "Targeting the Prefrontal-Supplementary Motor Network in Obsessive-Compulsive Disorder with Intensified Electrical Stimulation in Two Dosages: A Randomized, Controlled Trial." *Translational Psychiatry*, vol. 14, no. 1, Feb. 2024, p. 78. DOI.org (Crossref), <https://doi.org/10.1038/s41398-024-02736-y>.
- Apps, Matthew A. J., et al. "The Anterior Cingulate Gyrus and Social Cognition: Tracking the Motivation of Others." *Neuron*, vol. 90, no. 4, May 2016, pp. 692–707. DOI.org (Crossref), <https://doi.org/10.1016/j.neuron.2016.04.018>.
- Brain Stimulation Therapies - National Institute of Mental Health (NIMH).
<https://www.nimh.nih.gov/health/topics/brain-stimulation-therapies/brain-stimulation-therapies>. Accessed 12 Aug. 2024.
- Bremner, J. Douglas. "Traumatic Stress: Effects on the Brain." *Dialogues in Clinical Neuroscience*, vol. 8, no. 4, Dec. 2006, pp. 445–61. DOI.org (Crossref), <https://doi.org/10.31887/DCNS.2006.8.4/jbremner>.
- Carmi, Lior, Uri Alyagon, et al. "Clinical and Electrophysiological Outcomes of Deep TMS over the Medial Prefrontal and Anterior Cingulate Cortices in OCD Patients." *Brain Stimulation*, vol. 11, no. 1, Jan. 2018, pp. 158–65. DOI.org (Crossref), <https://doi.org/10.1016/j.brs.2017.09.004>.
- Carmi, Lior, Aron Tendler, et al. "Efficacy and Safety of Deep Transcranial Magnetic Stimulation for Obsessive-Compulsive Disorder: A Prospective Multicenter Randomized Double-Blind Placebo-Controlled Trial." *American Journal of Psychiatry*, vol. 176, no. 11, Nov. 2019, pp. 931–38. DOI.org (Crossref), <https://doi.org/10.1176/appi.ajp.2019.18101180>.
- Cognitive Behavioral Therapy - Mayo Clinic.
<https://www.mayoclinic.org/tests-procedures/cognitive-behavioral-therapy/about/pac-20384610>. Accessed 12 Aug. 2024.
- DaSilva, Alexandre F., et al. "Electrode Positioning and Montage in Transcranial Direct Current Stimulation." *Journal of Visualized Experiments*, no. 51, May 2011, p. 2744. DOI.org (Crossref), <https://doi.org/10.3791/2744>.
- De Lecea, Luis, et al. "Shining Light on Wakefulness and Arousal." *Biological Psychiatry*, vol. 71, no. 12, June 2012, pp. 1046–52. DOI.org (Crossref), <https://doi.org/10.1016/j.biopsych.2012.01.032>.
- "Deep Brain Stimulation." AANS,
<http://www.aans.org/patients/conditions-treatments/deep-brain-stimulation/>. Accessed 12 Aug. 2024.
- Delta Wave - an Overview | ScienceDirect Topics.
<https://www.sciencedirect.com/topics/neuroscience/delta-wave>. Accessed 12 Aug. 2024.
- Diagnosis | Obsessive-Compulsive and Related Disorders | Stanford Medicine.
<https://med.stanford.edu/ocd/about/diagnosis.html>. Accessed 12 Aug. 2024.
- Facts & Statistics | Anxiety and Depression Association of America, ADAA.
<https://adaa.org/understanding-anxiety/facts-statistics>. Accessed 12 Aug. 2024.
- Fornaro, Michele, et al. "Obsessive-Compulsive Disorder and Related Disorders: A Comprehensive Survey." *Annals of General Psychiatry*, vol. 8, no. 1, 2009, p. 13. DOI.org (Crossref), <https://doi.org/10.1186/1744-859X-8-13>.
- Frontal Lobe: What It Is, Function, Location & Damage.
<https://my.clevelandclinic.org/health/body/24501-frontal-lobe>. Accessed 12 Aug. 2024.

- Graat, Ilse, et al. "Cyclic versus Continuous Deep Brain Stimulation in Patients with Obsessive Compulsive Disorder: A Randomized Controlled Trial." *Brain Stimulation*, vol. 16, no. 1, Jan. 2023, pp. 82–87. DOI.org (Crossref), <https://doi.org/10.1016/j.brs.2023.01.1664>.
- Guo, Qihui, et al. "Continuous Theta Burst Stimulation over the Bilateral Supplementary Motor Area in Obsessive-Compulsive Disorder Treatment: A Clinical Randomized Single-Blind Sham-Controlled Trial." *European Psychiatry*, vol. 65, no. 1, 2022, p. e64. DOI.org (Crossref), <https://doi.org/10.1192/j.eurpsy.2022.2323>.
- Harika-Germaneau, Ghina, et al. "Treating Refractory Obsessive–Compulsive Disorder with Transcranial Direct Current Stimulation: An Open Label Study." *Brain and Behavior*, vol. 10, no. 7, July 2020, p. e01648. DOI.org (Crossref), <https://doi.org/10.1002/brb3.1648>.
- Hertrich, Ingo, et al. "The Role of the Dorsolateral Prefrontal Cortex for Speech and Language Processing." *Frontiers in Human Neuroscience*, vol. 15, May 2021, p. 645209. DOI.org (Crossref), <https://doi.org/10.3389/fnhum.2021.645209>.
- "History." *Obsessive-Compulsive and Related Disorders*, <https://med.stanford.edu/ocd/treatment/history.html>. Accessed 12 Aug. 2024.
- Ippolito, Giuseppe, et al. "The Role of Alpha Oscillations among the Main Neuropsychiatric Disorders in the Adult and Developing Human Brain: Evidence from the Last 10 Years of Research." *Biomedicines*, vol. 10, no. 12, Dec. 2022, p. 3189. DOI.org (Crossref), <https://doi.org/10.3390/biomedicines10123189>.
- Jiang, Jin, et al. "A Controlled Clinical Study of Accelerated High-Dose Theta Burst Stimulation in Patients with Obsessive–Compulsive Disorder." *Neural Plasticity*, edited by Sergio Bagnato, vol. 2023, Dec. 2023, pp. 1–8. DOI.org (Crossref), <https://doi.org/10.1155/2023/2741287>.
- Lang, E. W., et al. "Brain Connectivity Analysis: A Short Survey." *Computational Intelligence and Neuroscience*, vol. 2012, 2012, pp. 1–21. DOI.org (Crossref), <https://doi.org/10.1155/2012/412512>.
- Li, Hailong, et al. "Neural Primacy of the Dorsolateral Prefrontal Cortex in Patients with Obsessive-Compulsive Disorder." *NeuroImage: Clinical*, vol. 28, 2020, p. 102432. DOI.org (Crossref), <https://doi.org/10.1016/j.nicl.2020.102432>.
- Long-Term Potentiation - an Overview | ScienceDirect Topics*. <https://www.sciencedirect.com/topics/neuroscience/long-term-potentiation>. Accessed 12 Aug. 2024.
- Mann, Sukhmanjeet Kaur, and Narpinder K. Malhi. "Repetitive Transcranial Magnetic Stimulation." *StatPearls*, StatPearls Publishing, 2024. *PubMed*, <http://www.ncbi.nlm.nih.gov/books/NBK568715/>.
- Maraone, Annalisa, et al. "Antiglutamatergic Agents for Obsessive-Compulsive Disorder: Where Are We Now and What Are Possible Future Prospects?" *World Journal of Psychiatry*, vol. 11, no. 9, Sept. 2021, pp. 568–80. DOI.org (Crossref), <https://doi.org/10.5498/wjp.v11.i9.568>.
- Neuroplasticity - StatPearls - NCBI Bookshelf. <https://www.ncbi.nlm.nih.gov/books/NBK557811/>. Accessed 12 Aug. 2024.
- Oberman, Lindsay, et al. "Safety of Theta Burst Transcranial Magnetic Stimulation: A Systematic Review of the Literature." *Journal of Clinical Neurophysiology*, vol. 28, no. 1, Feb. 2011, pp. 67–74. DOI.org (Crossref),

<https://doi.org/10.1097/WNP.0b013e318205135f>.

Obsessive-Compulsive Disorder - National Institute of Mental Health (NIMH).
<https://www.nimh.nih.gov/health/topics/obsessive-compulsive-disorder-ocd>. Accessed 12 Aug. 2024.

Obsessive-Compulsive Disorder (OCD) | Johns Hopkins Medicine.
<https://www.hopkinsmedicine.org/health/conditions-and-diseases/obsessivecompulsive-disorder-ocd>. Accessed 12 Aug. 2024.

“Obsessive-Compulsive Disorder (OCD) - Symptoms and Causes.” Mayo Clinic,
<https://www.mayoclinic.org/diseases-conditions/obsessive-compulsive-disorder/symptoms-causes/syc-20354432>. Accessed 12 Aug. 2024.

“Overview - SSRI Antidepressants.” Nhs.Uk, 15 Feb. 2021,
<https://www.nhs.uk/mental-health/talking-therapies-medicine-treatments/medicines-and-psychiatry/ssri-antidepressants/overview/>.

Psychotherapies - National Institute of Mental Health (NIMH).
<https://www.nimh.nih.gov/health/topics/psychotherapies>. Accessed 12 Aug. 2024.

Pulgar, Victor M. “Direct Electric Stimulation to Increase Cerebrovascular Function.” *Frontiers in Systems Neuroscience*, vol. 9, Mar. 2015. DOI.org (Crossref),
<https://doi.org/10.3389/fnsys.2015.00054>.

Rădulescu, Anca, et al. “Global and Local Excitation and Inhibition Shape the Dynamics of the Cortico-Striatal-Thalamo-Cortical Pathway.” *Scientific Reports*, vol. 7, no. 1, Aug. 2017, p. 7608. DOI.org (Crossref), <https://doi.org/10.1038/s41598-017-07527-8>.

Riemann, Bryan L., and Scott M. Lephart. “The Sensorimotor System, Part I: The Physiologic Basis of Functional Joint Stability.” *Journal of Athletic Training*, vol. 37, no. 1, 2002, pp. 71–79.

Rolls, Edmund T., et al. “The Orbitofrontal Cortex: Reward, Emotion and Depression.” *Brain Communications*, vol. 2, no. 2, July 2020, p. fcaa196. DOI.org (Crossref),
<https://doi.org/10.1093/braincomms/fcaa196>.

Safadi, Ziad, et al. “Functional Segmentation of the Anterior Limb of the Internal Capsule: Linking White Matter Abnormalities to Specific Connections.” *The Journal of Neuroscience*, vol. 38, no. 8, Feb. 2018, pp. 2106–17. DOI.org (Crossref),
<https://doi.org/10.1523/JNEUROSCI.2335-17.2017>.

Schermer, Maartje. “Ethical Issues in Deep Brain Stimulation.” *Frontiers in Integrative Neuroscience*, vol. 5, 2011. DOI.org (Crossref), <https://doi.org/10.3389/fnint.2011.00017>.

Senova, Suhan, et al. “Deep Brain Stimulation for Refractory Obsessive-Compulsive Disorder: Towards an Individualized Approach.” *Frontiers in Psychiatry*, vol. 10, Dec. 2019, p. 905. DOI.org (Crossref), <https://doi.org/10.3389/fpsy.2019.00905>.

Swierkosz-Lenart, Kevin, et al. “Therapies for Obsessive-Compulsive Disorder: Current State of the Art and Perspectives for Approaching Treatment-Resistant Patients.” *Frontiers in Psychiatry*, vol. 14, Feb. 2023, p. 1065812. DOI.org (Crossref),
<https://doi.org/10.3389/fpsy.2023.1065812>.

The Clinical Global Impressions Scale - PMC. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2880930/>. Accessed 12 Aug. 2024.

Torrico, Tyler J., and Sara Abdijadid. “Neuroanatomy, Limbic System.” *StatPearls*, StatPearls Publishing, 2024. PubMed, <http://www.ncbi.nlm.nih.gov/books/NBK538491/>.

Transcranial Magnetic Stimulation - Mayo Clinic.

<https://www.mayoclinic.org/tests-procedures/transcranial-magnetic-stimulation/about/pac-20384625>.
Accessed 12 Aug. 2024.

Tyagi, Himanshu, et al. “A Randomized Trial Directly Comparing Ventral Capsule and Anteromedial Subthalamic Nucleus Stimulation in Obsessive-Compulsive Disorder: Clinical and Imaging Evidence for Dissociable Effects.” *Biological Psychiatry*, vol. 85, no. 9, May 2019, pp. 726–34. DOI.org (Crossref), <https://doi.org/10.1016/j.biopsych.2019.01.017>.

Uhlhaas, Peter J., and Wolf Singer. “High-Frequency Oscillations and the Neurobiology of Schizophrenia.” *Dialogues in Clinical Neuroscience*, vol. 15, no. 3, Sept. 2013, pp. 301–13.

The Interaction of Soil-Transmitted Helminth Infections, Geophagy, and Pregnancy Outcomes in Sub-Saharan Africa By Minseo Ryu

Abstract

This project aims to identify the interactions between soil-transmitted helminth infections (STH), geophagy, and pregnancy outcomes in Sub-Saharan Africa. STH infections, a type of neglected tropical disease, affect billions worldwide. Transmitted through contaminated soil or water, STH disproportionately impacts poor communities with limited sanitation and lack of clean resources, especially low and middle-income areas around Sub-Saharan Africa. Pregnant women are particularly vulnerable, especially those practicing geophagy to protect against toxins, access micronutrients, and combat hunger. This research highlights the need for further investigation into the relationship between STH infections, geophagy, and pregnancy outcomes. It also emphasizes the importance of developing culturally sensitive interventions to improve maternal and fetal health in Sub-Saharan Africa.

Introduction

Pregnancy is a crucial period for both mother and infant, with pregnant women being particularly vulnerable to certain diseases and infections such as STH. A commonly neglected tropical disease, STH are parasitic intestinal worms that affect mainly underprivileged communities worldwide. STH is commonly passed through direct ingestion or contact with eggs in soil, water, and food contaminated with fecal matter.

The three most common types of STH infection include *Ascaris lumbricoides* (roundworm), *Trichuris trichiura* (whipworm), and *Necator americanus* and *Ancylostoma duodenale* (hookworms), which most commonly infect pregnant women (Hotez et al.). Each type of STH infection varies by lifecycle, symptoms, transmission, and size. Significantly affecting low and middle-income areas around Sub-Saharan Africa, STH impacts poor communities with limited sanitation and lack of clean resources.

Studies have shown connections between STH infections and the practice of geophagy, which is the craving for non-food materials like clay, ice, chalk, starch, and sand. Geophagy, the practice of eating earth—particularly chalk or clay—has been practiced for two million years, not only by humans but also by primates and other vertebrates. Although the reason why humans practice geophagy is unknown, some hypotheses include protection against toxins and harmful pathogens, access to micronutrients, and ease of hunger (Miao et al.). Indeed, in Sub-Saharan Africa and South Asia, geophagy is most commonly practiced by pregnant women. Cultural, spiritual, and traditional beliefs contribute to the practice, as it is often associated with different rituals, spirits, and traditional health ideals (Kwon et al.).

Pregnant women tend to experience earth cravings during their iron deficiency periods and engage in these practices during their second and third trimesters in areas of Sub-Saharan Africa. This paper aims to uncover an additional connection between STH, geophagy, and pregnancy by using peer-reviewed research papers, books, and resources to investigate the

potential interactions between these three elements. In addition to addressing STH in pregnant women, this research will also discuss the interventions that have been put into place to address STH infections and opportunities for further intervention.

What Are Soil-Transmitted Helminth Infections?

STH are parasitic intestinal worms that affect vulnerable communities worldwide. They are particularly prominent in Sub-Saharan Africa, China, East and South Asia, and India (Brooker et al.). The most common type of STH is roundworm, also called *Ascaris lumbricoides*. *Ascaris lumbricoides* are the largest intestinal nematodes measuring approximately 250-300 millimeters in length. Their life cycle begins when eggs and larvae enter the human body upon ingestion of embryonated eggs in food, water, or soil contaminated with human feces. Then, larvae enter the bloodstream, spread to the lungs, and penetrate the alveoli, causing pulmonary symptoms. When the larvae migrate through the bloodstream or the lymphatic system, symptoms include coughing, chest discomfort, and wheezing. Once the larvae mature into adult worms in the intestines, common symptoms include nausea, diarrhea, vomiting, weight loss, and abdominal pain (Gardner and Gardner). After ascending to the trachea and descending into the gastrointestinal tract, the adult parasite releases its eggs in the small intestine. These eggs are then excreted in feces, which can contaminate soil and infect new hosts. This cycle takes around 2-3 months, and adult worms within an infected individual can remain for 1-2 years.

Another common type of STH is whipworm, also called *Trichuris trichiura*. *Trichuris trichiura* is a nematode measuring about 30-50 millimeters in length. The life cycle begins when an individual ingests embryonated eggs in food, water, or soil contaminated with human feces. After the eggs hatch in the small intestine and release larvae, the larvae transform into adult worms that eventually mature and produce thousands of eggs in the colon. These adult worms live in the cecum and ascending colon. Eventually, these eggs are passed into human feces and form embryos. The cycle begins again when embryonated eggs are ingested again by humans. The cycle takes around 1-3 months and adult worms typically live in an infected individual for 1-3 years. Without treatment, symptoms such as diarrhea, iron-deficient anemia, and weight loss will persist in those infected (Viswanath et al.).

The last most common type of STH is hookworms, also called *Necator americanus* and *Ancylostoma duodenale*. Hookworms are roundworms measuring 7-13 millimeters in length. Transmitted through skin penetration, the life cycle of these worms begins as eggs in human feces. The eggs hatch into larvae in warm soil, eventually maturing, and can survive for several weeks in the soil. These filariform larvae penetrate human skin and infect humans. They first enter the bloodstream, move to the heart, lungs, pharynx, stomach, and eventually into the small intestines. In the small intestines, the larvae mature into adult worms that lay thousands of eggs, eventually defecating in human feces. The cycle takes around 4-7 weeks and adult worms typically live in an infected individual for 1-2 years. Common symptoms include diarrhea, abdominal pain, dyspepsia, and fever (Ghodeif and Jain; Hoagland and Schad; Khurana et al.; Chai et al.).

STH Infection's Impact On Pregnant Women

STH infections, specifically hookworms, significantly affect pregnant women in Sub-Saharan Africa. A study completed among pregnant populations estimated around 30% of women across multiple countries in Sub-Saharan Africa are infected with STH during pregnancy. Additionally, compared to non-pregnant women, pregnancy has been associated with increased STH infections (Garrison et al.).

STH infections cause anemia in pregnant women, increase maternal mortality, low birth weight, and greater risk of preterm birth (Mengist et al.). Anemia, characterized by low hemoglobin levels, can impair egg maturation and reduce the chances of a successful pregnancy (Alene and Dohe). Helminth-induced anemia also affects insulin resistance and blood glucose, contributing to gestational diabetes. Severe anemia can also develop when a condition, such as malaria or existing nutritional deficiencies, occurs in addition to the STH. For example, a study exploring the effect of comorbidities in Ghana has shown that coinfection with helminths and malaria was “associated with low birth weight, preterm delivery, and small birth weight for gestational age” (Alene and Dohe). Another study in Nigeria also supported these results, suggesting that co-infection of helminths and malaria contributed more to lower birth weight than malaria infection alone did (Alene and Dohe). These studies suggest that comorbidities exacerbate the effect of STH infections if additional infections are circulating in the body. Thus, an infected individual with multiple diseases and pre-existing health factors may experience more negative health outcomes compared to an infected individual infected with only one particular disease. In both cases, pregnant women experience different health outcomes and symptoms than non-pregnant people do.

Studies have shown that STH infections may bring immunological changes that might alter the immune system of the pregnant woman and affect fecundity, the ability to produce more healthy offspring. In a study exploring the connection between helminth infections and fecundity, the findings reported that certain STH species, like *A. lumbricoides* and hookworm, are associated with immunological changes. Continuing the discussion on the impact of STH infections on immune systems, T-cells are a type of white blood cells that help the immune response. In the case for STH, *A. lumbricoides* is only associated with type 2 T-cell responses, while the hookworm is associated with both type 1 and type 2 T-cell responses (Blackwell et al.). It has been theorized that type-2 inducing helminth infections may positively influence pregnancy in infected mothers, leading to successful pregnancy and immune tolerance (Chetty et al.). This finding was supported by another study that found *A. lumbricoides* may be beneficial for pregnancy outcomes and promote conception rather than suppress fecundity. This contrasts with the widely-held theory that diseases and parasites decrease reproduction (Blackwell et al.).

Beyond the effects of STH directly on pregnant women, maternal to fetal transmission of STH infections may occur. Studies show that helminth antigens are transferred to infants *in utero*, and infants show alterations in immunity. Infants may experience increases in type 2 T-cell responses due to helminth antigens, crucial for humoral immune responses, and decreases in type

1 T-cell responses, which identify and eradicate viruses and bacteria, to non-helminth antigens (Blackwell).

In summary, pregnant women are particularly susceptible to STH infections; these co-morbidities and different health risk factors may bring even worse symptoms and responses to STH-infected mothers (Blackwell; Mpairwe et al.).

Geophagy

There are different aspects—from fetal development, nutritional intake, malaria, nutritional stress, and anemia—that influence the susceptibility to STH during pregnancy. However, one hypothesized explanation for this increased vulnerability involves a form of pica, called geophagy. The term 'pica' originates from the Latin word for 'magpie,' a bird known for its indiscriminate eating habits. The condition was formally named in the sixth century A.D (Johnson). Geophagy is not characterized by a mere desire to consume non-foods, but it is an uncontrollable and strong craving for non-food materials like clay, chalk, and soil. It is often associated with mental health disorders, nutritional deficiencies, and pregnancy.

There are many different hypothesized reasons as to why individuals engage in this practice. Some believe that geophagy develops to palliate pregnancy-related symptoms such as nausea, providing therapeutic benefits; they believe that practicing geophagy protects against toxins and harmful pathogens, aids digestion, provides micronutrients, acts as an antacid, and prevents diarrhea (“Pica”; Bonglaisin et al.). Others believe that geophagy is an outlet to express gender identities and a way for women to belong within the household (Geissler). It is seen as a continuity of the lineage and family ancestry since earth materials symbolize the graves of ancestors and are seen as elements that bring people together with life-giving forces (Geissler; Jefferds et al.). These different aspects associated with the cultural practice of geophagy may influence a pregnant woman's decision to engage in such practices.

Geophagy is most commonly practiced by pregnant populations in areas around Sub-Saharan Africa. It has been reported that the prevalence of geophagy among pregnant women ranged from 65% in Kenya, 46% in Ghana, 42% in Namibia, to 28% in Tanzania (Kawai et al.). In Southwest Ethiopia, there was a statistically significant association between STH infections and geophagia: pregnant and lactating women reported to consume earth during gestation (Getachew et al.).

Although this is a universal trend in Sub-Saharan Africa, the type of soil preferred by pregnant women varies. In Northern Uganda, for example, soils are collected and consumed from riverbanks, burnt bricks, wells, markets, termite mounds, and swamps. In areas of Kenya, pregnant women prefer clay solids because of its fine texture and tendency to dissolve easily in the mouth (Geissler et al.). Either way, those who practice geophagy seek particular types of soil and consume them based on availability and region; soils that contain high amounts of clay, air-dried, baked, and smoked are eaten depending on the region (Huebl et al.).

The prevalence of geophagy increases as pregnant women face the second and third trimesters, essential for the developmental and physiological changes of both the mother and the

fetus. As the fetus matures, more nutritional demands, food cravings, and greater demand compel women to practice geophagy to combat hormonal fluctuations and nutritional deficiencies (Getachew et al.; Allport). For instance, pregnancy hormones can heighten their sense of smell and turn on olfactory triggers, provoking pregnant women to engage in this practice (Cameron; Kortei et al.). Furthermore, as the pregnancy stage progresses to the second and third trimesters, the demand for iron increases as the body tries to support both the fetus and the mother (Luke).

While pregnant women practice geophagy for many different reasons, the precise information on the actual impact on pregnancy and the fetuses is still limited, and available information likely underreports prevalence from these populations. One reason that could contribute to underreported or inaccurate data is that the practice of geophagy differs significantly from the hygiene practices linked to Western medicine. Thus, pregnant women may perceive the practice as too normalized or feel ashamed to disclose their participation (Bonglaison et al.; Huebl et al.).

What Is The Connection Between Geophagy And STH?

STH infections are commonly transmitted through the ingestion of food, skin penetration, water, or soil contaminated with human feces. When pregnant women practice geophagy, especially during the second and third trimesters, they engage in a practice that puts them at risk for infections like *Ascaris lumbricoides* (roundworm) and *Trichuris trichiura* (whipworm). This is backed by studies indicating that soil consumption can expose pregnant women to accidental ingestion of STH, lead poisoning, and heavy metals (Blum et al.; Odongo et al.). Direct consumption of soil increases the risk of ingesting STH eggs, which, in turn, heightens the likelihood of pregnant women becoming infected by STH infections. As a result, they may be more likely to face side effects of infection including low birth rate and preterm delivery.

The effect of geophagia and its link with STH infections still varies and is being debated by healthcare professionals. While some studies report a strong association between geophagia in pregnant women and STH infections like *Ascaris lumbricoides* and *Trichuris trichiura*, other studies report a weaker association (Getachew et al.; Young et al.). These inconsistencies could stem from various factors including the environment of the conducted research, the prevalence of STH based on each region, the lack of accepted techniques or methodologies, and inadequate research on pregnant women in these regions around Sub-Saharan Africa.

Preventing STH In Pregnant Populations

It is essential to keep pregnant women safe from infection and informed during a particularly vulnerable period using water, sanitation, and hygiene (WASH) practices. WASH are practices labeled as crucial to the well-being of human health and well-being (Hutton and Chase). These practices are particularly important to uphold public health and protect the lives of those living in underprivileged and low-income areas with limited access to clean water, proper sanitation, and available hygiene practices. Clean water, hand washing, personal cleanliness, reliable sanitation, and proper hygienic practices aim to improve individual well-being and stop

the spread of infectious diseases (Hutton and Chase; Shrestha et al.; Zerbo et al.). Regardless of the form, these practices are essential to prevent the spread of infectious diseases, germs, and illnesses.

WASH Practices

Currently, national deworming campaigns and health education campaigns are in place to combat STH infections (Adriko et al.). Wearing shoes, engaging in hand washing practices before eating and after defecating, and using soap are WASH intervention methods proven to reduce STH infections (Strunz et al.). Yet, millions of people around the world still lack these basic services and fail to engage in clean practices crucial for health and protection. In Africa alone, 411 million people lack access to basic water services, 779 million people lack access to basic sanitation, and 839 million lack basic hygiene as of 2022 (Mustapha et al.). This lack of access to important services drives both the general population and pregnant women around this area to practice open defecation, drink unsafe water, and bathe and wash clothes in contaminated water, thereby heightening exposure to these different infections (Prüss-Ustün et al.; Tseole et al.).

In particular, people living in poor rural areas have less access and lower coverage of WASH practices compared to those in urban areas. To illustrate: in Ethiopia, 75% of the urban population has access to on-premises water, while only 5% of its rural population has water piped in their homes. In the Democratic Republic of the Congo, 1% of its rural populations have access to on-premises water. Only 2% of rural populations in Chad have water running in their homes (“Water Scarcity in Sub-Saharan Africa”). Generally, 40% of urban populations in Sub-Saharan Africa used better sanitation, while only 23% of rural populations did (“Barriers to Water, Sanitation, and Hygiene”; Lynn et al.). Sanitation services and formal water supply have always favored urban and wealthier populations than rural areas (“Water Scarcity in Sub-Saharan Africa”). Consequently, pregnant women living in rural areas experience this disparity and suffer from inadequate access to WASH practices and resources than pregnant women residing in urban and wealthier areas.

Mass Drug Administration

Mass drug administration (MDA) is a possible intervention to control tropical diseases like STH infections. It involves providing treatment to all populations in an area, regardless of the circumstances and disease status. MDA is frequently labeled as a campaign-style strategy to administer drugs to high-risk populations and vulnerable areas (Chong et al.). These drugs are very effective and minimize the severe effects of STH infections. In particular, the World Health Organization (WHO) recommends MDA be given in the form of two anthelmintic medications: mebendazole or albendazole (Ng’etich et al.). Mebendazole and albendazole are inexpensive, produce minor side effects, and are effective.

Delving deeper into these two anthelmintic medications, mebendazole works by preventing worms from absorbing glucose, an essential component for the survival of STH (Chai

et al.). Albendazole is another effective treatment option for STH infections. Similar to the mebendazole, albendazole kills the worm by preventing it from absorbing sugar or glucose. (Chai et al.). For pregnant populations, a single 500 mg mebendazole dose or a single albendazole 400 mg dose is recommended after the first trimester, usually after a diagnosis (“Infectious Diseases”).

While these effective preventive medication options are available, they are less available to females compared to males; children, too, may not be reached by the deworming programs in certain areas around Sub-Saharan Africa due to the lack of equitable educational status, family size, and household media exposure. Consequently, this inequity hinders pregnant populations from receiving preventive medication options (Wong et al.).

Policy And Governmental Interventions

The control of these diseases and infections has been labeled as a priority among many donors, government, and agencies providing funding and aid for STH control (Tchuenté). In areas around Ethiopia, international partners support deworming programs financially, and the government continues to strengthen its partnership with international, national, and regional partners for a safer and healthier environment (Negussu et al.). With the support from these partners, government ministries have even launched strategies focusing on deworming among school-based health activities.

Indeed, significant progress and measures has been made, but the level of interventions in Sub-Saharan Africa is still lacking, and communities and pregnant populations living on the continent continue to suffer from STH infections.

Social And Behavior Change Communication

Social and Behavior Change Communication (SBCC) is a framework used to promote health changes in a community through communication strategies and educational materials. Some behavior change interventions include improving health literacy, enhancing partnerships with policymakers to create stronger guidelines, and building sustainable models for the future (Chauhan et al.). In regions of Sub-Saharan Africa, implementation of this framework is particularly weak, possibly due to differences in language barriers, dialects, inadequate WASH infrastructures, mass drug administration (MDA), and distribution points throughout the region (Hoefle-Bénard and Salloch). In Southern Ethiopia for instance, MDA coverage struggles to reach all populations due to the lack of awareness about STH infections and inaccessibility to reach drug distribution points. Depending on the location of these distribution points, the availability of effective anthelmintic drugs differs significantly. Some pregnant populations living closer to distribution points may have better access to prevention drugs than others who have to travel longer to access them (Asfaw et al.).

Neglected Tropical Disease Control Programs

In Sub-Saharan Africa, the implementation of WHO's NTD control programs, which aim to eliminate and control neglected tropical diseases like STH infections with regional collaborations, are difficult to maintain due to programmatic, political, and social differences (Akinsolu et al.; Belay et al.). For instance, the educational backgrounds of pregnant women differ significantly; women with higher educational backgrounds are more likely to take their child to receive deworming medication compared to women with no such education (Belay et al.). Depending on these differences, the effectiveness of NTD control programs may differ and are harder to maintain for less educated pregnant populations unaware of control programs (Mustapha et al.).

Discussion

This project aims to bring more awareness around the different ways in which STH infections and the cultural practice of geophagy impact pregnancy and pregnancy outcomes in Sub-Saharan Africa. Its goal is to bring more recognition toward the health risks of STH infections and advocate for better prevention methods and interventions in Sub-Saharan Africa. The data collected from a thorough literature review suggests that pregnant women are at higher risk of STH infections due to the cultural practice of geophagy.

Clash Between WASH And STH

Prevention of STH infections is particularly difficult to address due to the cultural component of geophagy both from an infectious view and WASH intervention standpoint. Currently, WASH practices clash with the cultural practice of geophagy. WASH practices emphasize basic hygiene and sanitation practices, but geophagy directly contradicts this practice as it involves the direct consumption of potentially harmful earth materials. Consequently, western interventions, like WASH initiatives, can overlook the importance of the cultural practice of geophagy among different regions around Sub-Saharan Africa. Research indicates that pregnant individuals may feel ashamed and avoid seeking treatment options.

While WASH practices must continue to be emphasized in vulnerable areas around Sub-Saharan Africa and other regions around the world, it is also crucial to recognize and find ways to disseminate culturally sensitive health information. As a public health community, there must be more effective ways to address and communicate about cultural practices that could be harmful to health. WASH practices must serve as an intervention system that addresses differences and encompasses the needs of populations around the world. Implementing ineffective or uninformed interventions does not positively promote change in the region; rather, it hinders and disrespects the culture and identity of those living in the area.

Indeed, implementing efficient intervention methods that account for these diverse cultures and approaches to pregnancy will be challenging. However, with careful tailoring to individual countries and tribal differences, the public health community can make gradual but meaningful progress to find more suitable treatment options for pregnant populations in each area.

More Targeted Action

In addition to more culturally-sensitive interventions, stronger implementation of SBCC interventions, NTD control programs, MDA, preventive medications, and deworming programs are needed. While these interventions are present in Sub-Saharan Africa and have made significant progress, they are inadequate to fully address the unique needs of pregnant populations. The inequalities influencing this—wealth status, distribution points, misconceptions of drug safety, lack of sanitation campaigns, and shortage of health education messages—need to improve for the betterment of these populations as a whole.

Most importantly, better funding and aid programs must be implemented in Sub-Saharan Africa to control STH infections in the longer term. Currently, funding and aid programs through U.S. global health assistance are temporary and not sustainable (Ingram et al.). These funding mechanisms often fail to support a stable infrastructure that can support itself even after the grant period (Ingram et al.).

In conclusion, there are many programs and organizations providing funding and aid to countries across Sub-Saharan Africa, but more permanent intervention is necessary to ensure control of STH infections in pregnant populations around Sub-Saharan Africa.

Limitations

This is a literature review based on secondary research. The biggest limitation of this project was the inability to observe first-hand the impacts of STH infections and geophagy on pregnant populations. While there have been studies looking at STH infections, geophagy, and pregnant populations, conclusions varied based on the control group, environment, research methodology, and findings. Furthermore, pregnant women are a protected group due to their vulnerabilities and health needs. This factor alone limits the amount of necessary research available to fully address STH infections and health concerns in this population. Considering these limitations, this paper reflects the available data in the best way possible.

Conclusion

STH infections are one of many neglected tropical diseases and pose major risks to pregnant populations worldwide. While STH infections and the cultural practice of geophagy negatively affect pregnancy populations in Sub-Saharan Africa, there must be further research done to fully address and understand the link between STH infections, geophagy, and pregnancy risks. A greater and more targeted focus on eliminating the risks of STH is needed to raise awareness of intervention practices and address the loopholes in implementation within each region across Sub-Saharan Africa. While doing so, it is crucial to recognize the existing practices in the region and remain respectful when educating communities on the potential health risks involved in their respective practices.

Works Cited

- Adriko, Moses, et al. "Impact of a National Deworming Campaign on the Prevalence of Soil-transmitted Helminthiasis in Uganda (2004-2016): Implications for National Control Programs." *PLoS Neglected Tropical Diseases*, vol. 12, no. 7, July 2018, p. e0006520. <https://doi.org/10.1371/journal.pntd.0006520>.
- Akinsolu, Folahanmi T., et al. "Community Perception of School-based Mass Drug Administration Program for Soil-transmitted Helminths and Schistosomiasis in Ogun State, Nigeria." *PLoS Neglected Tropical Diseases*, vol. 17, no. 7, July 2023, p. e0011213. <https://doi.org/10.1371/journal.pntd.0011213>.
- Alene, Kefyalew Addis, and Abdulahi Mohamed Dohe. "Prevalence of Anemia and Associated Factors Among Pregnant Women in an Urban Area of Eastern Ethiopia." *Anemia*, vol. 2014, Jan. 2014, pp. 1–7. <https://doi.org/10.1155/2014/561567>.
- Allport, Susan. *Women Who Eat Dirt*. content.ucpress.edu/chapters/11342.allport.pdf.
- Asfaw, Mekuria Asnakew, et al. "Evaluating Equity and Coverage in Mass Drug Administration for Soil-Transmitted Helminth Infections Among School-Age Children in the Hard-to-Reach Setting of Southern Ethiopia." *Pediatric Health Medicine and Therapeutics*, vol. Volume 12, July 2021, pp. 325–33. <https://doi.org/10.2147/phmt.s316194>.
- Barriers to Water, Sanitation, and Hygiene in Sub-Saharan Africa: A Mini Review*. [ouci.dntb.gov.ua/en/works/4rGj8pJl](https://www.who.int/publications/i/item/9789241550000).
- Belay, Daniel Gashaneh, et al. "Deworming Among Preschool Age Children in sub-Saharan Africa: Pooled Prevalence and Multi-level Analysis." *Tropical Medicine and Health*, vol. 50, no. 1, Oct. 2022, <https://doi.org/10.1186/s41182-022-00465-w>.
- Blackwell, Aaron D. "Helminth Infection During Pregnancy: Insights From Evolutionary Ecology." *International Journal of Women S Health*, vol. Volume 8, Nov. 2016, pp. 651–61. <https://doi.org/10.2147/ijwh.s103529>.
- Blackwell, Aaron D., et al. "Helminth Infection, Fecundity, and Age of First Pregnancy in Women." *Science*, vol. 350, no. 6263, Nov. 2015, pp. 970–72. <https://doi.org/10.1126/science.aac7902>.
- Blum, Winfried E. H., et al. "Does Soil Contribute to the Human Gut Microbiome?" *Microorganisms*, vol. 7, no. 9, Aug. 2019, p. 287. <https://doi.org/10.3390/microorganisms7090287>.
- Bonglaisin, Julius Nsawir, et al. "Geophagia: Benefits and Potential Toxicity to human—A Review." *Frontiers in Public Health*, vol. 10, July 2022, <https://doi.org/10.3389/fpubh.2022.893831>.
- Brooker, S., et al. "Global Epidemiology, Ecology and Control of Soil-Transmitted Helminth Infections." *Advances in Parasitology/Advances in Parasitology*, Jan. 2006, pp. 221–61. [https://doi.org/10.1016/s0065-308x\(05\)62007-6](https://doi.org/10.1016/s0065-308x(05)62007-6).
- Cameron, E. Leslie. "Pregnancy and Olfaction: A Review." *Frontiers in Psychology*, vol. 5, Jan. 2014, <https://doi.org/10.3389/fpsyg.2014.00067>.

- Chai, Jong-Yil, et al. “Albendazole and Mebendazole as Anti-Parasitic and Anti-Cancer Agents: An Update.” *Korean Journal of Parasitology*, vol. 59, no. 3, June 2021, pp. 189–225. <https://doi.org/10.3347/kjp.2021.59.3.189>.
- Chauhan, Bhupendrasinh F., et al. “Behavior Change Interventions and Policies Influencing Primary Healthcare Professionals’ Practice—an Overview of Reviews.” *Implementation Science*, vol. 12, no. 1, Jan. 2017, <https://doi.org/10.1186/s13012-016-0538-8>.
- Chetty, Alisha, et al. “Impact of Helminth Infections on Female Reproductive Health and Associated Diseases.” *Frontiers in Immunology*, vol. 11, Nov. 2020, <https://doi.org/10.3389/fimmu.2020.577516>.
- Chong, Nyuk Sian, et al. “Modelling the Ability of Mass Drug Administration to Interrupt Soil-transmitted Helminth Transmission: Community-based Deworming in Kenya as a Case Study.” *PLoS Neglected Tropical Diseases*, vol. 15, no. 8, Aug. 2021, p. e0009625. <https://doi.org/10.1371/journal.pntd.0009625>.
- Gardner, Scott L., and Sue Ann Gardner. “Concepts in Animal Parasitology.” *DigitalCommons@University of Nebraska - Lincoln*, digitalcommons.unl.edu/zeabook/160.
- Garrison, Amanda, et al. “Soil-transmitted Helminth Infection in Pregnancy and Long-term Child Neurocognitive and Behavioral Development: A Prospective Mother-child Cohort in Benin.” *PLoS Neglected Tropical Diseases*, vol. 15, no. 3, Mar. 2021, p. e0009260. <https://doi.org/10.1371/journal.pntd.0009260>.
- Geissler, P. W., et al. “Geophagy, Iron Status and Anaemia Among Pregnant Women on the Coast of Kenya.” *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 92, no. 5, Sept. 1998, pp. 549–53. [https://doi.org/10.1016/s0035-9203\(98\)90910-5](https://doi.org/10.1016/s0035-9203(98)90910-5).
- Geissler, P. Wenzel. “The Significance of Earth-Eating: Social and Cultural Aspects of Geophagy Among Luo Children.” *Africa*, vol. 70, no. 4, Nov. 2000, pp. 653–82. <https://doi.org/10.3366/afr.2000.70.4.653>.
- Getachew, Mestawet, et al. “Soil-Transmitted Helminthic Infections and Geophagia Among Pregnant Women in Jimma Town Health Institutions, Southwest Ethiopia.” *Ethiopian Journal of Health Sciences*, vol. 31, no. 5, Sept. 2021, <https://doi.org/10.4314/ejhs.v31i5.16>.
- Ghodeif, Alhassan O., and Hanish Jain. “Hookworm.” *StatPearls - NCBI Bookshelf*, 15 June 2023, www.ncbi.nlm.nih.gov/books/NBK546648.
- Hoagland, K. E., and G. A. Schad. “*Necator Americanus* and *Ancylostoma Duodenale*: Life History Parameters and Epidemiological Implications of Two Sympatric Hookworms of Humans.” *Experimental Parasitology*, vol. 44, no. 1, Feb. 1978, pp. 36–49. [https://doi.org/10.1016/0014-4894\(78\)90078-4](https://doi.org/10.1016/0014-4894(78)90078-4).
- Hoefle-Bénard, Juliette, and Sabine Salloch. “Mass Drug Administration for Neglected Tropical Disease Control and Elimination: A Systematic Review of Ethical Reasons.” *BMJ Global Health*, vol. 9, no. 3, Mar. 2024, p. e013439. <https://doi.org/10.1136/bmjgh-2023-013439>.

- Hotez, Peter J., et al. "Helminth Infections: Soil-transmitted Helminth Infections and Schistosomiasis." *Disease Control Priorities in Developing Countries - NCBI Bookshelf*, 2006, www.ncbi.nlm.nih.gov/books/NBK11748.
- Huebl, Lena, et al. "Geophagy in Northern Uganda: Perspectives From Consumers and Clinicians." *American Journal of Tropical Medicine and Hygiene*, vol. 95, no. 6, Dec. 2016, pp. 1440–49. <https://doi.org/10.4269/ajtmh.15-0579>.
- Hutton, Guy, and Claire Chase. "Water Supply, Sanitation, and Hygiene." *The World Bank eBooks*, 2017, pp. 171–98. https://doi.org/10.1596/978-1-4648-0522-6_ch9.
- Ingram, George, et al. "Top Five Reasons Why Africa Should Be a Priority for the United States." *Brookings*, 3 Apr. 2013, www.brookings.edu/articles/top-five-reasons-why-africa-should-be-a-priority-for-the-unit-ed-states.
- Jefferds, Maria Elena D., et al. "Formative Research Exploring Acceptability, Utilization, and Promotion in Order to Develop a Micronutrient Powder (Sprinkles) Intervention Among Luo Families in Western Kenya." *Food and Nutrition Bulletin*, vol. 31, no. 2_suppl2, June 2010, pp. S179–85. <https://doi.org/10.1177/15648265100312s210>.
- Johnson, Bruce E. "Pica." *Clinical Methods - NCBI Bookshelf*, 1990, www.ncbi.nlm.nih.gov/books/NBK255/#:~:text=Pica%20is%20the%20compulsive%20eating,objects%20to%20satisfy%20its%20curiosity.
- Kawai, Kosuke, et al. "Geophagy (Soil-eating) in Relation to Anemia and Helminth Infection Among HIV-Infected Pregnant Women in Tanzania." *PubMed Central (PMC)*, 1 Jan. 2009, www.ncbi.nlm.nih.gov/pmc/articles/PMC7893611.
- Khurana, Sumeeta, et al. "Diagnostic Techniques for Soil-Transmitted Helminths – Recent Advances." *Research and Reports in Tropical Medicine*, vol. Volume 12, Aug. 2021, pp. 181–96. <https://doi.org/10.2147/rrtm.s278140>.
- Kortei, Nii Korley, et al. "Elemental Minerals and Microbial Compositions as Well as Knowledge and Perceptions Regarding Kaolin (Clay) Consumption by Pregnant Women in the Ho Municipality of Ghana." *Pan African Medical Journal*, vol. 34, Oct. 2019, <https://doi.org/10.11604/pamj.2019.34.113.17394>.
- Kwon, Dayoon, et al. "Association of Pica With Cortisol and Inflammation Among Latina Pregnant Women." *American Journal of Human Biology*, vol. 36, no. 5, Dec. 2023, <https://doi.org/10.1002/ajhb.24025>.
- Luke, Barbara. "Nutrition During Pregnancy: Part I, Weight Gain; Part II, Nutrient Supplements." *JAMA*, vol. 265, no. 2, Jan. 1991, p. 281. <https://doi.org/10.1001/jama.1991.03460020139044>.
- Lynn, Mary K., et al. "Soil-Transmitted Helminths in the USA: A Review of Five Common Parasites and Future Directions for Avenues of Enhanced Epidemiologic Inquiry." *Current Tropical Medicine Reports*, vol. 8, no. 1, Jan. 2021, pp. 32–42. <https://doi.org/10.1007/s40475-020-00221-2>.

- Mengist, Hylemariam Mihiretie, et al. “Intestinal Helminthic Infection and Anemia Among Pregnant Women Attending Ante-natal Care (ANC) in East Wollega, Oromia, Ethiopia.” *BMC Research Notes*, vol. 10, no. 1, Sept. 2017, <https://doi.org/10.1186/s13104-017-2770-y>.
- Miao, Diana, et al. “A Meta-analysis of Pica and Micronutrient Status.” *American Journal of Human Biology*, vol. 27, no. 1, Aug. 2014, pp. 84–93. <https://doi.org/10.1002/ajhb.22598>.
- Mpairwe, H., et al. “Pregnancy and Helminth Infections.” *Parasite Immunology*, vol. 36, no. 8, Aug. 2014, pp. 328–37. <https://doi.org/10.1111/pim.12101>.
- Mustapha, Adetoun, et al. “Pathway to Equity in Water, Sanitation, and Hygiene (WASH) in Africa: Challenges and Opportunities.” *PubMed*, vol. 8, no. 4, Aug. 2024, p. e315. <https://doi.org/10.1097/ee9.0000000000000315>.
- . “Pathway to Equity in Water, Sanitation, and Hygiene (WASH) in Africa: Challenges and Opportunities.” *PubMed*, vol. 8, no. 4, Aug. 2024, p. e315. <https://doi.org/10.1097/ee9.0000000000000315>.
- Negussu, Nebiyu, et al. “Ethiopia Schistosomiasis and Soil-Transmitted Helminthes Control Programme: Progress and Prospects.” *PubMed Central (PMC)*, 2017, www.ncbi.nlm.nih.gov/pmc/articles/PMC5582635.
- Ng’etich, Annette Imali, et al. “Anthelmintic Resistance in Soil-transmitted Helminths: One-Health Considerations.” *Parasitology Research*, vol. 123, no. 1, Dec. 2023, <https://doi.org/10.1007/s00436-023-08088-8>.
- Odongo, A. O., et al. “Heavy Metals and Parasitic Geohelminths Toxicity Among Geophagous Pregnant Women: A Case Study of Nakuru Municipality, Kenya.” *Environmental Geochemistry and Health*, vol. 38, no. 1, Mar. 2015, pp. 123–31. <https://doi.org/10.1007/s10653-015-9690-3>.
- “Pica.” *PubMed*, 1 Jan. 2024, pubmed.ncbi.nlm.nih.gov/30335275.
- “Principles and Practice of Pediatric Infectious Diseases.” *Elsevier Shop*, shop.elsevier.com/books/principles-and-practice-of-pediatric-infectious-diseases/long/978-0-323-75608-2.
- Prüss-Ustün, Annette, et al. “Burden of Disease From Inadequate Water, Sanitation and Hygiene for Selected Adverse Health Outcomes: An Updated Analysis With a Focus on Low- and Middle-income Countries.” *International Journal of Hygiene and Environmental Health*, vol. 222, no. 5, June 2019, pp. 765–77. <https://doi.org/10.1016/j.ijheh.2019.05.004>.
- Shrestha, Akina, et al. “Water, Sanitation, Hygiene Practices, Health and Nutritional Status Among Children Before and During the COVID-19 Pandemic: Longitudinal Evidence From Remote Areas of Dailekh and Achham Districts in Nepal.” *BMC Public Health*, vol. 22, no. 1, Nov. 2022, <https://doi.org/10.1186/s12889-022-14346-8>.
- Strunz, Eric C., et al. “Water, Sanitation, Hygiene, and Soil-Transmitted Helminth Infection: A Systematic Review and Meta-Analysis.” *PLoS Medicine*, vol. 11, no. 3, Mar. 2014, p. e1001620. <https://doi.org/10.1371/journal.pmed.1001620>.

- Tchuenté, L. a Tchuem. “Control of Soil-transmitted Helminths in sub-Saharan Africa: Diagnosis, Drug Efficacy Concerns and Challenges.” *Acta Tropica*, vol. 120, Sept. 2011, pp. S4–11. <https://doi.org/10.1016/j.actatropica.2010.07.001>.
- Tseole, Nkeka P., et al. “Barriers and Facilitators to Water, Sanitation and Hygiene (WaSH) Practices in Southern Africa: A Scoping Review.” *PLoS ONE*, vol. 17, no. 8, Aug. 2022, p. e0271726. <https://doi.org/10.1371/journal.pone.0271726>.
- Viswanath, Avinash, et al. “Trichuris Trichiura Infection.” *StatPearls - NCBI Bookshelf*, 14 Aug. 2023, www.ncbi.nlm.nih.gov/books/NBK507843.
- “Water Scarcity in Sub-Saharan Africa Means Children Are Having to Go to Wells Instead of to School. How Big Is the Problem?” *World Economic Forum*, 8 Nov. 2022, www.weforum.org/agenda/2022/09/water-accessibility-divide-sub-saharan-africa-visualised.
- Wong, Matthew Tze Jian, et al. “Soil-transmitted Helminthic Vaccines: Where Are We Now?” *Acta Tropica*, vol. 239, Mar. 2023, p. 106796. <https://doi.org/10.1016/j.actatropica.2022.106796>.
- Young, Sera L., et al. “Geophagia Is Not Associated With Trichuris or Hookworm Transmission in Zanzibar, Tanzania.” *Transactions of the Royal Society of Tropical Medicine and Hygiene*, vol. 101, no. 8, Aug. 2007, pp. 766–72. <https://doi.org/10.1016/j.trstmh.2007.04.016>.
- Zerbo, Alexandre, et al. “Water Sanitation and Hygiene in Sub-Saharan Africa: Coverage, Risks of Diarrheal Diseases, and Urbanization.” *Journal of Biosafety and Biosecurity*, vol. 3, no. 1, June 2021, pp. 41–45. <https://doi.org/10.1016/j.job.2021.03.004>.

The Ideal Ultimate Frisbee Throw: Grid search for the perfect pass

By Anthony Ding

Abstract

In 1968, the idea of Ultimate Frisbee was introduced by Joel Silver. In this short time, it has slowly risen to popularity, but not to the level of other, more established sports. As such, the depth of research that sports like basketball or football athletes can benefit from are far greater than that of ultimate frisbee players. This study aims to assist in remedying that, and answered a central question: What would the ideal short-to-mid range pass look like in Ultimate Frisbee? Throws were simulated and given a rating, a single quantitative measure of the quality of the throw. Since this area has not been investigated in too much depth, this study used broad ranges for initial values of speed, pitch, and spin speed. This work provides a methodology for the analysis of frisbee throws, setting a foundation for future studies, which may look into more specific parameters and complex factors.

Introduction

Ultimate Frisbee (UF) is a relatively new sport, only a few decades old. The main goal is to pass the frisbee from one player to another until the end zone is reached, and the scoring team receives a point. Possibly the largest part of UF is passing the frisbee from one player to another, since a player holding the frisbee is not allowed to take any steps. Players run either to guard the player they have marked, or to get to a more open position where the disc can be more easily passed to them (Krustrup & Mohr, 2015). The large amount of running can thus be attributed to the goal of completing good passes. This leads to good passes being an integral part of playing well.

In part due to being a newer and relatively less popular sport, many aspects of UF have not been studied as much as in other sports such as football or basketball. There have been some studies analyzing the flight of the disc in disc golf, and how factors such as weight impact the way the disc flies, but the depth of this research does not apply to UF (Immonen, 2022). Some studies have looked into factors of the flight of a disc in UF, by building a robotic arm (Junge & Hughes, 2023), investigating the flight with Newtonian mechanics (Landell-Mills, 2019), or modelling the flight of the frisbee using MatLab (Hummel, 2003).

Simulations have also been made studying the fluid dynamics around the disc, rather than just the flight path of the disc itself (Peng & Yang, 2021). Studies have also been done using real world experimental tests, with values being read by sensors/ These values are then studied to reveal how accurate they are (Lorenz, 2005). Wind tunnels were also used in certain studies, to study the disc in a very controlled environment (Koyanagi et al., 2012). Even the flight of wobbly frisbees were analyzed using a triaxial gyroscope (Weizman et al., 2020).

This study attempts to identify ideal releases for short to mid range passes in UF, answering the question of what an ideal frisbee throw would consist of and show a methodology in which these parameters can be obtained.

In the methods section, the process for gathering and analyzing data will be discussed. In depth details about the output from simulations and the analysis of this data will be presented in the results section, and an explanation of how different factors of a throw impact the trajectory and subsequent relevancy of this data will be discussed in the discussion section. The conclusion will present the final findings of the study, as well as present limitations and directions for future study.

Methods

The data for this study is the result of simulations, which were programmed in MatLab and analyzed using custom software in Python. The graphs from Matlab will illustrate how differing initial conditions, such as release speed or spin impact a 5-meter pass by showing their relationship with final speed, final height, and others.

It is first important to establish what an ideal short-to-mid range pass would consist of. For the purposes of this study, passes over a distance of 5 meters will be analyzed. Differences in height of the defender and attacker will not be considered since the advantage could go either way in a real game, so the target height will be around chest level at ~1.3 meters. Due to the short ranged nature of this pass, slower passes will be preferred. This will also allow for the disc to curve more and create more dynamic passes that are more difficult to intercept. The defender is also assumed to be a meter in front of the attacker trying to catch the disc, and the closest distance that the path of the frisbee comes to the defender is also measured.

The flight path of the discs were modeled using a Matlab simulation code extract from (Hummel, 2003). The simulation takes in a set of initial parameters, such as initial speed along with physical constants like rotational inertia of the standard frisbee and acceleration due to gravity.

The simulation takes into account most basic forces encountered by the disc as it travels. These factors include constants such as acceleration due to gravity, but factors that vary as the disc travels such as aerodynamic forces. These forces act unevenly on the frisbee due to angle, spin, and other factors, leading to the center of pressure being different from the center of mass. This can be compensated for with moments.

Drag is a large part of the consideration in the flight of frisbees. In the simulation, it was modelled using the formula for drag

$$D = \frac{1}{2} \rho v^2 A C_D, \quad (1)$$

where D is the magnitude of drag force, C_D is the drag coefficient, A is the frontal area of the disc projected in the direction of the velocity, v is the linear velocity, and ρ is the fluid density. C_D and A is partially determined by the angle of attack for frisbees. As such, varying the pitch of the disc will result in varying values of drag coefficient, and thus drag force. Additionally, drag force is also proportional to the velocity squared, meaning that varying initial speed would also impact drag force. As such, altering these factors would allow observations of the impact of drag on frisbee flight. C_D can be calculated using the formula

$$C_D = C_{D0} + C_{D\alpha}(\alpha - \alpha_0)^2, \quad (2)$$

where C_D is the drag coefficient, C_{D0} is the form drag, or minimum drag, $C_{D\alpha}$ is the induced drag, and α is the angle of the frisbee with the horizontal (McCormick, 1995).

Another element of frisbee flight is lift, which fundamentally operates similarly to drag. It follows the equation

$$L = \frac{1}{2} \rho v^2 A C_L \quad (3)$$

where L is the magnitude of lift, C_L is the lift coefficient, A is the frontal area of the disc projected in the direction of the velocity, v is the linear velocity, and ρ is the fluid density. The force is generated due to Newton's third law, as the air pushes back up against the frisbee while the frisbee pushes the air in front downwards to create a path to travel along. Similarly to drag, C_L and A can be altered for frisbees by adjusting the angle of attack. Velocity is a direct factor into the equation for the magnitude of lift, and again altering initial velocity and initial pitch of the disc will alter the lift force throughout the flight of the disc, and subsequently its final height. It should also be noted that the above equation is only an approximation for the lift of a frisbee, as the unique shape of a frisbee creates many irregularities in air flow. However, data of pressure around a frisbee has been collected and studied in the past (Potts & Crowther, 2002).

Since lift force is directed in the vertical direction, any alterations to it throughout the flight of the frisbee will have effects upon the final height of the frisbee. Since the final height is one of the variables that are used to determine the rating of the throw, changes made that impact lift will subsequently affect the rating of the throw.

One factor that could impact these variables is wind. Due to variance in wind speed and direction, it would not be possible to account fully for wind in a universal scenario. Difference in wind speed and direction would impact the relative velocity of the frisbee compared to the air around it, impacting both drag and lift as mentioned above. As such, this study will neglect the effects of wind and assume there is no wind during the throw for simplicity and more universal application.

This data is used to generate a flight path for the frisbee. However, due to the nature of simulations, it will become less accurate over longer distances. As such, a short pass distance of 5 meters was used. Different values of initial speed, pitch, and spin were processed in the simulation through the use of nested while loops. The range for the values chosen is very high due to this study aiming to find a general target for initial parameters. The data from the simulation was then processed through a custom Python analysis, which gave a rating to each throw. The function

$$rating = - height_norm - speed_norm + 2 \quad (4)$$

was used to determine an overall rating for the pass, where rating is the final rating assigned to the throw, $height_norm$ is the normalized value of height about 1.3, and $speed_norm$ is the normalized speed value. The addition of the constant 2 makes sure that rating will always have a

positive value. It should be noted that distance from the interceptor was omitted from this formula as the distance was not significant enough to have any effect. This will be discussed further in the discussion section.

The output $height_norm$ was found using the equation

$$height_norm = (|1.3 - height| - min) / (max - min) \quad (5)$$

where height refers to the final height of the disc, min refers to the minimum value of $|1.3 - height|$ out of all the throws, and max refers to the maximum value of $|1.3 - height|$ in all the throws.

The outputs $speed_norm$ and int_dist_norm were found using standard normalization, through the equations

$$speed_norm = (speed - min_speed) / (max_speed - min_speed) \quad (6)$$

$$int_dist_norm = (int_dist - min_int_dist) / (max_int_dist - min_int_dist) \quad (7)$$

where speed refers to the final speed of the disc, min_speed refers to the minimum value of speed out of all the throws, and max_speed refers to the maximum value of speed in all the throws. int_dist refers to the minimum distance between the path of the frisbee and the interceptor, min_int_dist refers to the minimum value of this distance out of all throws, and max_int_dist refers to the maximum value of this distance.

A set of initial parameters for the disc were given, with varying levels of initial speed, spin, and the initial pitch angle of the disc at release. A very large range in all parameters are covered, in order to find generally how the disc should be thrown. These variables were then entered into a MatLab simulation, which modeled the flight of the disc (Hummel, 2003). This model was chosen because it modeled the flight of the disc accurately in short ranges, which was the focus for this study. It was also available in its entirety in a published thesis.

The flight of the discs for the varying starting variables were then exported from MatLab so they could be analyzed. The data was imported to Python, and we considered the previously mentioned scenario where the receiver is 5m away, with an interceptor a meter in front of the receiver. Horizontal speed and height of the disc were also measured at this point as final speed ($speed$) and final height ($height$). The closest distance between the interceptor and the flight path of the frisbee was also measured, and recorded as interceptor distance (int_dist). Finally, the parameters final speed, final height, and interceptor distance were normalized using min-max normalization and used as inputs for the rating equation, which describes how good the pass for that simulation was.

Results

In Fig. 1 it can be seen that as initial speed ($speedo$) or pitch ($thetao$) is increased, the other must be decreased slightly in order to achieve the best possible throw at that speed or pitch. Through observing the graph when $thetao = 0.3$, it is possible that an even higher rating could be achieved by throwing at that angle with a lower speed.

Fig. 2 depicts a similar pattern, but presented in the form of a heatmap. It clearly illustrates that it is ideal to throw by angling the disc further above the horizontal plane, and releasing at a lower speed. However, it is worth noting that the throws with the highest rating in this study were not the throws with the *lowest* initial speed and *highest* initial pitch, instead having a slightly higher speed and lower pitch.

Fig. 3 depicts the effects of *thetao* and *speedo* upon the final normalized height. From the heatmap, it can be clearly seen that there are a range of values for *thetao* and *speedo* which correlate with each other which allow for the frisbee to reach the target height.

A very clear pattern is shown in Fig. 4, with a lower value of *thetao* or higher value of *speedo* resulting in a higher final speed. There is a consistent increase in final speed with every marginal decrease in *thetao*, and similarly with every marginal increase of *speedo*.

Fig. 5 shows the relationship *thetao* has upon final speed and rating. It is shown that at higher values for *thetao*, rating values have a wider range. Additionally, higher values of *thetao* tend to produce generally lower values of final speed.

Fig. 6 depicts the impact of *thetao* and *speedo* upon the final height, with a line indicating the target height of 1.3m. Higher values of *thetao* and *speedo* seem to result in a higher final height for the disc. Points that lie farthest from the target height seem to be those with higher or lower values for *both thetao* and *speedo*.

Fig. 7 places emphasis on the speed of the disc, looking into the impact of *speedo* upon the normalized final speed value and the subsequent effect upon rating. As shown in the graph, higher *speedo* values correlate to higher *speed_norm* values and lower rating. It should be noted that *speed_norm* is a direct factor in the calculation for rating.

Fig. 8 depicts a graph of *gd* and rating. There is no clear pattern.

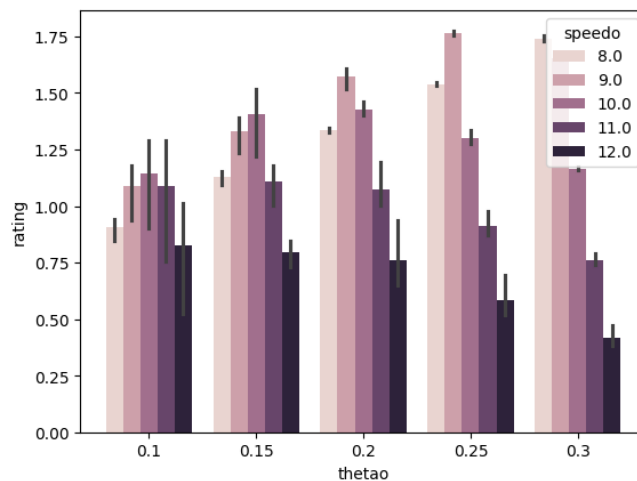


Fig. 1 Influence of initial pitch and initial speed on rating. *speedo* is measured in meters/second, *thetao* is measured in radians above the horizontal axis.

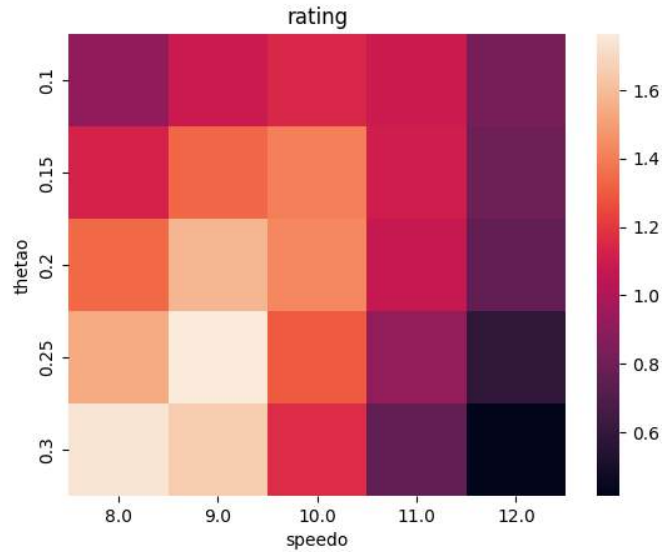


Fig. 2 Alternate perspective of Figure 1 in the form of a heatmap. Illustrates the ideal maximization of rating. *speedo* is measured in meters/second, *thetao* is measured in radians above the horizontal axis.

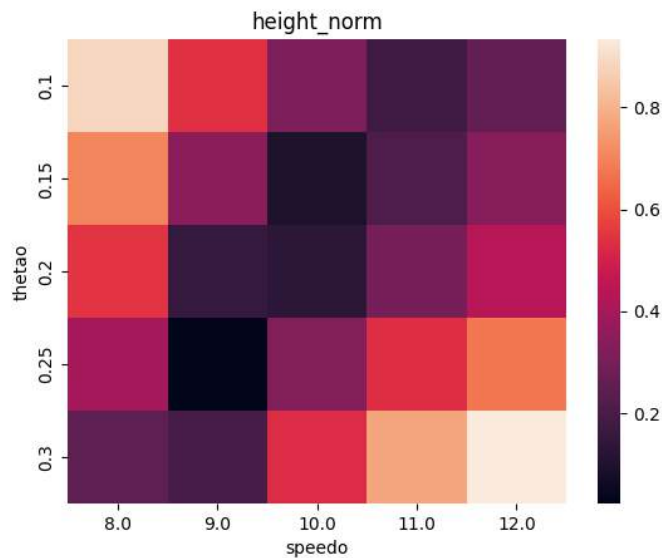


Fig. 3 Relationship between initial pitch/initial speed and the normalized height value. *speedo* is measured in meters/second, *thetao* is measured in radians above the horizontal axis. *height_norm* is the normalized value of height about the value of 1.3 meters.

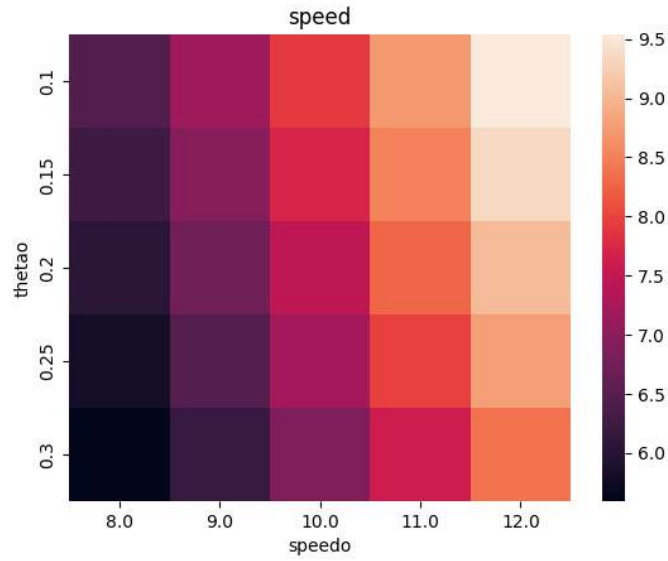


Fig. 4 Influence of *speedo/thetao* upon final speed. *speedo* and final speed are measured in meters/second, *thetao* is measured in radians above the horizontal axis.

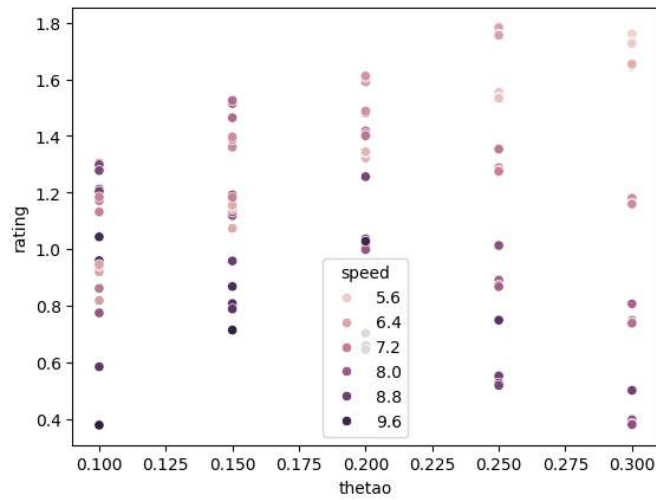


Fig. 5 Overall impact of initial pitch upon rating and final speed. Final speed is measured in meters/second, *thetao* is measured in radians above the horizontal axis.

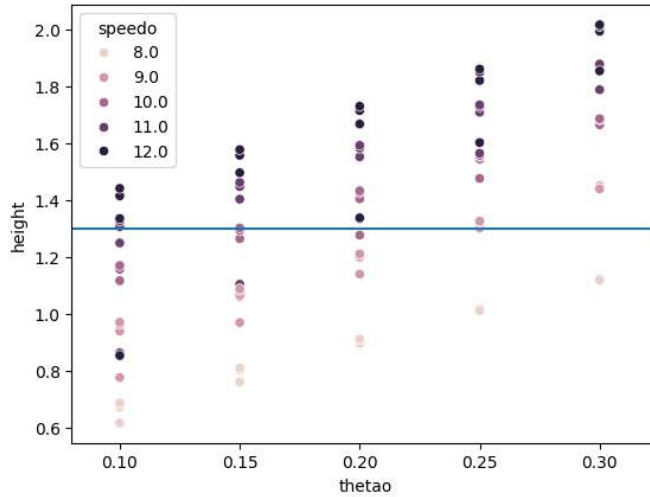


Fig. 6 Impact of initial pitch and initial speed upon final height. *speedo* is measured in meters/second, *thetao* is measured in radians above the horizontal axis, height is measured in meters. The blue horizontal line indicates the target height of 1.3m

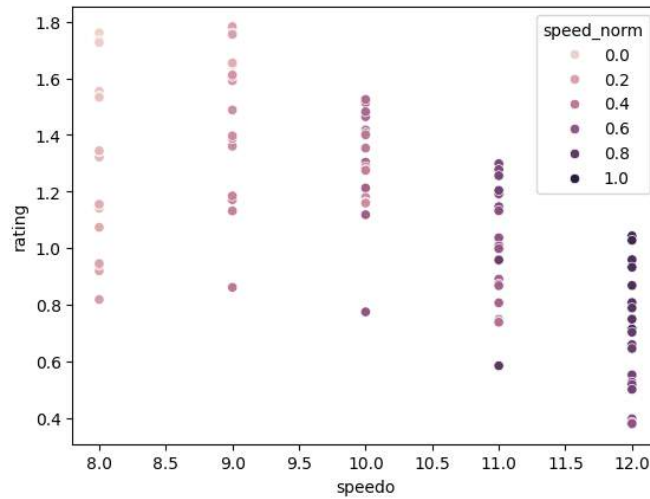


Fig. 7 Impact of initial speed upon rating and normalized final speed. *speedo* is measured in meters/second. Focus is placed upon speed, showing the impact of initial speed upon normalized final speed and final rating

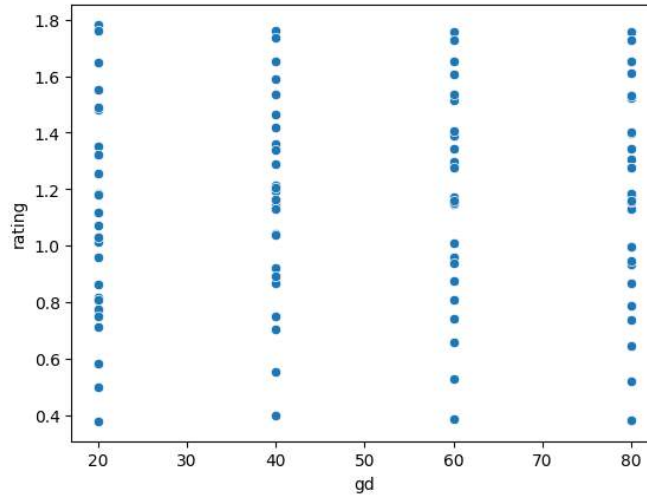


Fig. 8 Impact of initial spin (gd) upon rating. gd is measured in rad/sec

Discussion

For a pass in UF of 5 meters, Fig. 1 and Fig. 2 show that the ideal initial parameters are an initial velocity of $\sim 9\text{m/s}$ and an initial pitch of around 0.25 radians above the x-axis. Fig. 1 also displays an interesting pattern on the effects of *speedo* and *thetao* in conjunction. An equilibrium must be found between *speedo* and *thetao* in order to find a throw with the ideal final speed and height, since an increase in one input means the other must decrease. This is due to the disc experiencing more drag at a higher pitch due to having increased frontal area, meaning it must be thrown at a higher initial velocity to reach the target at the ideal height. This is further explored in Fig. 3, which shows a clear set of corresponding values of *speedo* and *thetao* that result in the ideal final height.

In Fig. 4, the direct impact that initial speed and initial pitch have on final speed can be seen. As such, in order to achieve a slow throw, The disc should be thrown slowly at an angle. However, this data must be considered in conjunction with data from Fig. 3 as both final speed and final height are factors in the final rating of the throw. It can be observed that the values of initial speed and initial pitch must line up with each other in order to achieve the desired height. As mentioned previously, this is primarily due to the increased frontal area of the disc as *thetao* increases, as well as the fact that more lift is generated with a higher pitch.

A pattern can be observed in Fig. 5, with there generally being more variability in the rating of throws as the initial pitch increases. This may be due to speed having more impact at a larger angle since more air resistance is encountered. This is further supported by the fact that the rating value seems to be more affected by the final speed as the initial pitch increases.

Fig. 6 depicts a very clear relationship, with all but a few outliers showing that the final height of the disc increases as initial pitch and speed increase. The line at 1.3m shows the target height, however and shows the values of initial speed and initial pitch, which, in conjunction, bring the frisbee to the correct height.

In Fig. 7, there is a very clear, logical correlation between initial and final speed. It makes logical sense that a frisbee that is thrown faster will still be travelling faster when it reaches its destination. However, this graph also shows an interesting correlation between initial speed and the final rating. While generally a slower speed results in a high rating, this does not hold true at very low initial speeds. This is likely due to the throws all being below the target height, which decreases rating.

As mentioned above, the spin of the disc was not considered as a factor. Fig. 8 clearly depicts this, showing the lack of effect that gd has upon rating. This persists even over the extremely large range of gd values tested. The only final value that gd had any noticeable impact upon was the distance from the interceptor, but that was not considered as the distances were only a few centimeters at most.

It should be noted that this distance is minimal due to the input parameters not including rotation around the x-axis, which is where the vast majority of the curve in a frisbee's flight path comes from. Trying to also simulate this would also add many layers of complexity, since the tilting of the disc will make it far more affected by wind, yet another variable not considered in this study.

Conclusion

This study aimed to find a set of initial parameters that would lead to the ideal pass in UF over a distance of 5 meters. A set of initial parameters were fed into a model, which provided flight paths of the frisbee. This data was then processed through a custom analysis program, and evaluated using a new method of evaluation that was proposed in this study. This resulted in a single value that would give a quantitative, objective evaluation on each pass.

Since the variables used in this study were over a wide range and not very specific, general values were found that would give a good pass at 5 meters. The parameters for the best throws were at a starting pitch of 0.25 radians above the x-axis and at a speed of 9.0m/s. It was also found that the initial spin of the disc did not have much impact on the final speed and final height of the throws, even at greatly differing values.

The biggest concern about this study is whether or not the data will actually be accurate during an actual game, since it doesn't account for any wind or any other external factors. These are natural limitations when using a general simulation to model throws. It also does not factor in any attempt to curve the disc, which is quite a large part of passing in UF. This study will provide a solid baseline about UF throws, allowing for further exploration of more complex factors.

Works Cited

- Hummel, S. A. (2003). *Frisbee Flight Simulation and Throw Biomechanics*. University of California, Davis.
- Immonen, E. (2022). Optimal design for disc golf by computational fluid dynamics and machine learning. *Structural and Multidisciplinary Optimization*, 65.
<https://doi.org/10.1007/s00158-021-03107-7>
- Junge, K., & Hughes, J. (2023). Exploring Dynamically Controlled Frisbee Throws Using a Highly Compliant Robotic Arm. *2023 IEEE International Conference on Soft Robotics (RoboSoft)*, 1–6. <https://doi.org/10.1109/RoboSoft55895.2023.10122037>
- Koyanagi, R., Seo, K., Ohta, K., & Ohgi, Y. (2012). A computer simulation of the flying disc based on the wind tunnel test data. *Procedia Engineering*, 34, 80–85.
<https://doi.org/10.1016/j.proeng.2012.04.015>
- Krustrup, P., & Mohr, M. (2015). Physical Demands in Competitive Ultimate Frisbee. *Journal of Strength and Conditioning Research*, 29(12), 3386–3391.
<https://doi.org/10.1519/JSC.0000000000000989>
- Landell-Mills, N. (2019). How frisbees fly according to Newtonian physics.
- Lorenz, R. D. (2005). Flight and attitude dynamics measurements of an instrumented Frisbee. *Measurement Science and Technology*, 16(3), 738.
<https://doi.org/10.1088/0957-0233/16/3/017>
- McCormick, B. W. (1995). *Aerodynamics, Aeronautics and Flight Mechanics* by Barnes McCormack (2nd ed.).
https://www.academia.edu/39803770/Aerodynamics_Aeronautics_and_Flight_Mechanics_by_Barnes_McCormack
- Peng, Y., & Yang, J. (2021). Computational Fluid Dynamics Study of A Flying Frisbee. *2021 4th International Symposium on Traffic Transportation and Civil Architecture (ISTTCA)*, 302–306. <https://doi.org/10.1109/ISTTCA53489.2021.9654449>
- Potts, J., & Crowther, W. (2002). Frisbee(TM) Aerodynamics. In *20th AIAA Applied Aerodynamics Conference*. American Institute of Aeronautics and Astronautics.
<https://doi.org/10.2514/6.2002-3150>
- Weizman, Y., Tan, A. M., & Fuss, F. K. (2020). Measurement of Flight Dynamics of a Frisbee Using a Triaxial MEMS Gyroscope. *Proceedings*, 49(1), Article 1.
<https://doi.org/10.3390/proceedings2020049066>

The Cause of the Politicization of the Judiciary and How It Threatens Democracy in South Korea By Taeun Kim

Abstract

The original meaning and function of politics are diminishing in South Korea due to the deepening political polarization in the society. This polarization is not only confined to the political elite but has extended to the general populace, creating a stark division within the country. The judiciary and politics are making concerted efforts to influence each other, potentially harming the foundations of democracy by going against the principle of the separation of powers. This research identifies the relationship between political polarization, judicialization of politics, and politicization of the judiciary in South Korea by discussing the actions and interactions of the Korean governments and the National Assembly across different historical periods, providing an understanding of how these forces have changed and their impact on the democratic stability of the nation.

Keywords: Politicization of Judiciary, Judicialization of Politics, Political Polarization, Law and Politics, Democracy, Separation of Powers, Supreme Court, Media

Introduction

When Lee Jae-myung of the Democratic Party of Korea said late last year that he was concerned about the "judicialization of politics" over the prosecution of politicians of his party, Justice Minister Han Dong-hoon responded by saying that the problem was the "politicization of the judiciary," which politicians use the judiciary to defend themselves against crimes. Such speeches indicate the rise of interest in the relationship between the judiciary and politics in South Korea. It is considered that the judicialization of politics leads to the politicization of the judiciary (Torre 2016), understood as a more urgent topic in South Korea (한규섭, 2023). How can the cause of the politicization of the judiciary be used to explain the threat to democracy that South Korea is facing? While multiple previous research studies have focused on examining the relationship between the judicialization of politics and its effect on Korean democracy, it is not commonly understood how the politicization of the judiciary is potentially or already threatening Korean democracy.

Therefore, this research focuses on how political polarization leads to the judicialization of politics and its relationship with the politicization of the judiciary. Furthermore, it aims to elaborate on the effect the politicization of the judiciary has on democracy and how it exclusively influences Korea.

Theoretical background & Political polarization

Democracy is based on the principle of majority rule, and the members of representative institutions are elected by citizens eligible to vote in the nation's election. Therefore, the legislative, executive, and judicial branches cannot change the country's fundamentals, which the

Constitution permanently establishes. Under the principle of separation of powers, the classic role of the courts when dealing with politics is to declare what the law is and to order the correction of violations of constitutional norms or individual rights when the political branches violate them. Judicial independence does not mean its isolation from politics but rather a position in politics that allows it to pursue constitutional values and truth without being influenced by politics.

Political Polarization in the Political Arena

Recently, however, the benefit of the principle of majority has been diminishing in South Korea; political polarization toward the extremes is occurring among politicians in the National Assembly. Table 1 is shown below to explain this situation in numerical statistics.

	Number of Received Bills	Number of Processed Bills	Percentage of Processed Bills
16th National Assembly (2000.05.30~2004.05.29)	2,507	1,753	69.9%
17th National Assembly (2004.05.30~2008.05.29)	7,489	4,328	57.8%
18th National Assembly (2008.05.30~2012.05.29)	13,913	7,612	54.7%
19th National Assembly (2012.05.30~2016.05.29)	17,822	8,010	44.9%
20th National Assembly (2016.05.30~2020.05.29)	24,141	9,139	37.9%
21th National Assembly (2020.05.30~2024.05.29)	25,830	9,478	36.7%

Table 1

The percentage of processed bills throughout the past years has consistently decreased while the number of received bills is increasing. This means that while the problems the citizens face are growing daily, the political arena is fighting between their ideals, setting the citizens' needs aside. Politicians, who are supposed to put the citizens' lives and their benefits as the utmost priority as their representatives when voting about legislation, are instead voting based on their partisan interests; they tend not to even look at legislation that clashes with their partisan's ideology. This situation fails to fulfill Article 1, Section 2 of the Constitution of the Republic of Korea, which states, "The sovereignty of Korea resides in its citizens, and all power comes from the citizens."

Example of the South Korea's President Yoon Suk Yeol

Article 53 of the Constitution of the Republic of Korea gives the president the power to veto bills passed by the National Assembly. Although the president has these powers, it isn't easy to see them proceeding with the vetoes because the presidential veto intends to ensure that the passed bill is aligned with the Constitution, not with the president's political stance, and because their reputation towards the citizens might be damaged. This trend can be seen in the table below.

	Presidential Tenure	Number of Presidential Vetoes	Avg. Number of Presidential Vetoes
Rhee Syng-man	11 years 11 months	45	10.0
Yun Po-sun	1 year 7 months	0	
Park Chung-hee	15 years 10 months	5	
Choi Kyu-hah	8 months	0	
Chun Doo-hwan	6 years 10 months	0	
Roh Tae-woo	5 years	7	2.3
Kim Young-sam	5 years	0	
Kim Dae-jung	5 years	0	
Roh Moo-hyun	5 years	6	
Lee Myung-bak	5 years	1	
Park Geun-hye	4 years	2	
Moon Jae-in	5 years	0	
Yoon Suk-yeol	2 years 2 months	15	

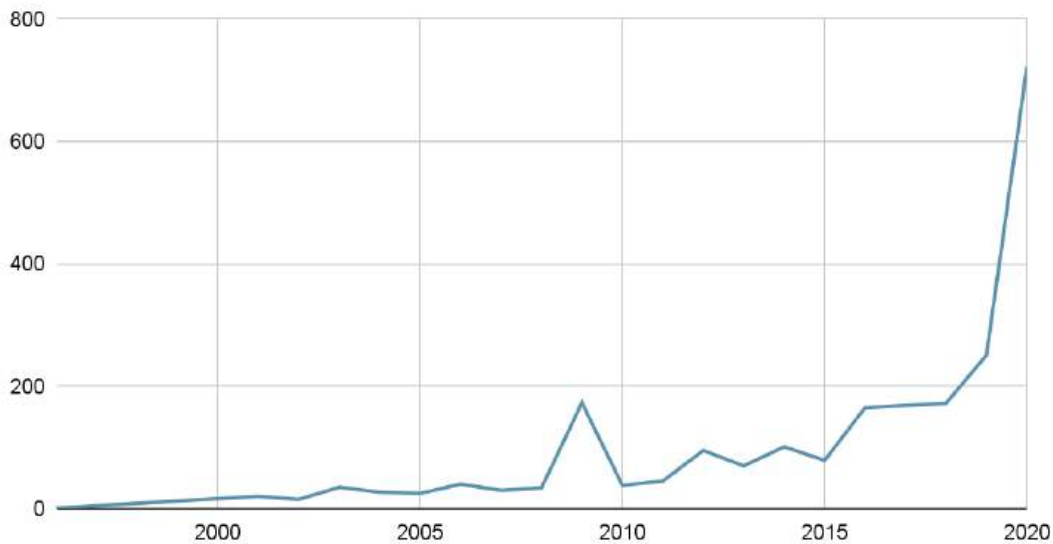
Table 2

From Rhee Syng-man to Chun Doo-hwan, when Korea wasn't yet democratized, the average number of vetoes each president used was ten times. However, only 2 out of 5 presidents vetoed, each vetoing 45 and 5 times during over ten years of presidency. Furthermore, 3 out of 7 presidents that came to power after the 1987 democratization, starting from Roh Tae-woo, did not veto a single time, while the other four did, making it an average of 2.3 times per president. The current South Korean President, Yoon Suk-yeol, has vetoed 15 times until now. He is still in his 2nd year of the term, with three years remaining. Considering this fact, President Yoon has made unusually excessive amounts of vetoes compared to the former presidents; he made the most since the 1987 democratization in his 2nd year.

Some of the bills he vetoed require detailed investigations into his wife's stock control crime, the Itaewon crowd collapse, the government's weak countermeasures, and the particular prosecutor act of the People Power Party's—his party—crime on 5 billion won bribery. This clearly shows President Yoon's intentions to nullify bills that go against his and his party's political ideologies and stances, which leads to even worse political polarization.

Political Polarization in the Public

The political polarization that started in the political arena has now begun to be seen among citizens and the public. The National Assembly Futures Institute's Futures Brief Vol.23-09 has acknowledged that political polarization in public within South Korea has become a severe problem since 2009 and is getting even worse yearly.



Graph 1 Frequency Counts of Public Articles that Discuss Political Polarization
Recreated graph of the National Assembly Futures Institute: South Korea's Political Polarization: 13 Typological Characteristics

Judicialization of politics

The judicialization of politics refers to the reliance on courts and judicial means to address core moral predicaments, public policy questions, and political controversies (Hirschl 2013). It is thought to be one of the leading causes of the politicization of the judiciary; politicians rely on the judiciary to find the solvency of their problems, and this incentivizes politicians to make efforts to influence the judiciary so they can use the law as a shield to defend themselves and as a weapon at the same time to attack other politicians and gain an advantageous position in the arena.

As the political polarization is getting worse and worse, the discussions and conversations for the better future of the country and its citizens are decreasing. Instead of going

through the complicated and stressful process of bringing up decent ideas for legislation, they are willing to increase their presence by suing and accusing the opponents, which is comparatively more straightforward. For instance, during the 2024 General Election, 1681 people have been sued by politicians or enforced by the police department, which is an increase of 24.5% compared to the 2020 General Election.

One of the specific examples is when the democratic party sued Han Dong-hoon, the Chair of the Emergency Committee of the conservative party, right before the General Election because he criticized one of the democratic party's nominees for the election, saying he was a "Sichuan King," claiming that he spread fake information. Right after, the conservatives counter-sued Lee Jae-myung, the head of the democratic party, claiming that he spread fake news to get Han punished during the election simply. These legal cases rooted in political polarization mute essential discussions for society, deplete the quality of citizens' lives and undermine the authority of parliament and the legitimacy of politics.

Politicization of the judiciary

The politicization of the Judiciary refers to "the process of courts and judges becoming influenced by political factors and considerations in their decision-making" (Miller 2019). People with power, especially the demagogues, tend to weaponize these unstable social atmospheres surrounding political polarization and make efforts to influence the Judiciary.

Example of the former United States President Donald J. Trump

A comparatively recent example related to the politicization of the judiciary is how former US president Donald Trump changed the attorney generals multiple times. During the four years of his presidency, from January 20th, 2017, to January 20th, 2021, he assigned six different attorney generals.

	Date of Office	End of Office
Sally Yates	January 20, 2017	January 30, 2017
Dana Boente	January 30, 2017	February 9, 2017
Jeff Sessions	February 9, 2017	November 7, 2018
Matthew Whitaker	November 7, 2018	February 14, 2019
William Barr	February 14, 2019	December 23, 2020
Jeffrey A. Rosen	December 24, 2020	January 20, 2021

Table 3

There may be unrevealed reasons behind the frequent change of attorney generals. However, because his presidency was one of the times when there were most changes in the seat,

it is assumed that Trump ultimately tried to control and influence the prosecutors. Sally Yates is a definite example. She was fired because she "betrayed the Department of Justice by refusing to enforce a legal order." She was named a traitor by Trump because she refused to defend his tightened travel policies, claiming it was unlawful, and was fired after ten days she was assigned the attorney general. There may be reputations arguing that since the prosecutors are a group that initially receives commands from the president, this makes it non-related to the politicization of the judiciary. However, suppose the prosecutors and the Minister of Justice cannot maintain stable neutrality in investigations and decisions. In that case, criminal justice, a significant principle in the judiciary, is also likely to be distorted.

How the politicization of the judiciary harms democracy

There are several reasons why the politicization of the judiciary is problematic.

The Essence of Liberal Democracy

The crux of liberal democracy is enabling citizens to check the government and maintain checks and balances within a country. Government powers have to be separated due to the principle of separation of powers under a democratic society (Bellamy 1996), and make the best efforts to listen to its citizens and ensure their maximum benefits since democracy is a sort of applied utilitarianism (Ely, 1990) and the advantage of democracy is in contributing to the well-being of the most significant number (Tocqueville, 1945).

However, the judiciary cannot be separated and independent; instead, it begins a puppet play by the governors. In that case, the principle of separation of powers will collapse, leading to a society lacking checks and balances. Without a stable balance between the powers, the borders dividing the powers will crumble, threatening the fundamentals of democracy.

Furthermore, the importance of the judiciary is not limited to the fact that it is 'one of the three.' Rule of Law is the factor that compromises social justice, enabling societies to be democratic, respect human rights, and organize economies so that they benefit the majority of the population (Waldron, 2011). The judiciary is the most crucial factor stabilizing and part of the Rule of Law. Unlike other powers, the judiciary is the one that protects the fundamental rights of the citizens from the government's inappropriate usage of force or other suppressions; it is the most potent shield the citizens have. When those in power hold this shield, it will instead function as a blade threatening the ones it should protect. This decreases the accountability and transparency of the government and the country, and the public will lose their trust in them, a complete failure of democracy.

The Exclusivity of Democracy in South Korea

Democracy is more than simply a system of government in Korea, as the table below explains.

Japanese Colonial Era	1910 ~ 1945
Trusteeship of Korea by the US & Soviet Union	1945 ~ 1948
Korean War	1950 ~ 1953
1st Military Dictatorship	1961 ~ 1979
2nd Military Dictatorship	1980 ~ 1987
6th Republic & Democratization	1987 ~

Table 4

As the table shows, it has not been long since Korea started functioning as a proper democratic country. From 1910 to 1945, Korea was colonized by the Japanese and could not make any independent decisions; Korean people participating in any of the governmental organizations and making choices were prohibited, including economy, foreign affairs, Korean culture, and more. Right after Korea achieved independence in August 1945, the nation was split into two parts, North and South, by the United States and the Soviet Union. Three years of an international war occurred on the peninsula, and the military regime took power as the battle ended. Thousands of people were killed to fight the regime and obtain democracy, and millions were sentenced to inappropriate punishments.

Korea is experiencing its first 30 years of democracy for the first time in 5000 years. The sole peninsula has its painful history of being oppressed, split, and blooded. It took them more than just efforts to achieve the democracy they have now, making democracy even more valuable. They know what happens when the judiciary is under the control of the government and has experienced its harsh consequences. Therefore, it can be concluded that democracy is an especially valuable settlement for Korea.

How the presidency may influence the judiciary

In Korea, the current president has the power to appoint 14 justices of the Supreme Court, whose terms are six years, with the consensus of the National Assembly. If so, to protect the basis of democracy, presidents must assign justices who will not be influenced by politics and provide fair sentences to everyone. However, it turns out that there is an existing relationship between the president who appointed the justice and his or her political color in the decision (한규섭, 2023).

Han's research quantitatively analyzes 325 en banc decisions of the Supreme Court from March 2006 to July 2023 to estimate the judgmental tendencies of each Supreme Court justice. A total of 50 justices, including the chief, were included in the analysis. The research applied the latent variable model, which clusters Supreme Court justices who consistently make similar rulings without subjective judgment so that they receive similar scores. The analysis is oriented

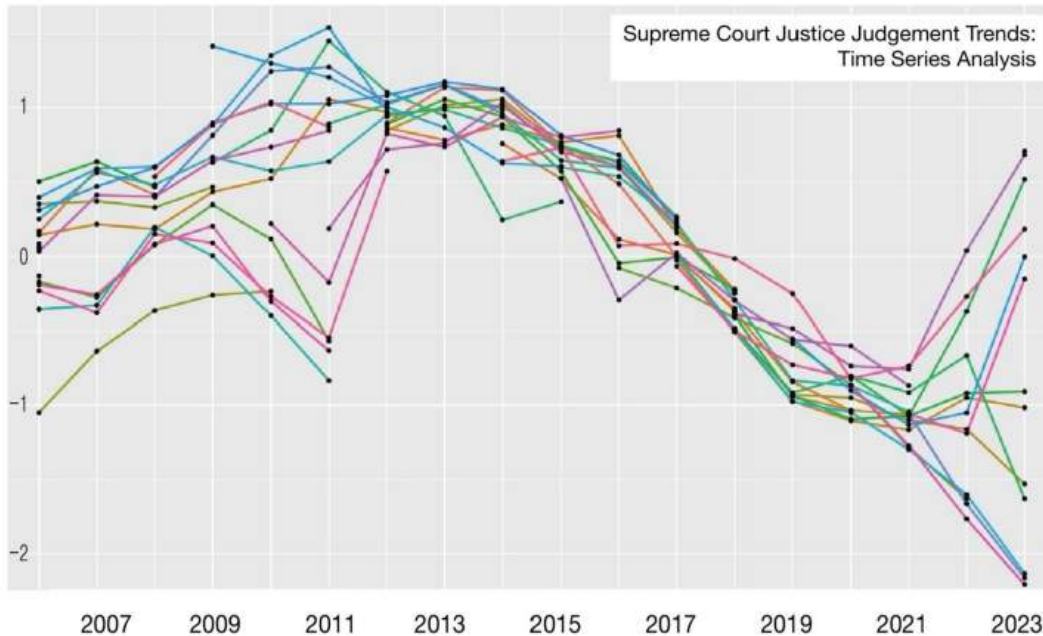
around the mean of the 50 Supreme Court justices, which is 0, with lower scores indicating more liberal leanings and higher scores meaning more conservative leanings.

Appointer	Number of Justices	Average	Standard Deviation
Kim Dae-jung	6	0.032	0.035
Roh Moo-hyun	12	-0.117	1.064
Lee Myung-bak	13	0.171	0.413
Park Geun-hye	5	0.099	0.607
Moon Jae-in	13	-0.358	0.574
Yoon Suk-yeol	1	1.003	

Table 5

Recreated table of The Seoul Shinmun Daily - 한규섭의 데이터 정치학; ‘정치의 사법화’가 촉발한 ‘사법의 정치화’... 법원 신뢰 추락 우려/서울대 언론정보학과 교수(정치커뮤니케이션)

This research by Han shows that the justice's tendencies of decisions shift towards the political ideologies of the president that appointed him or her. Additionally, due to the enhanced level of the judicialization of politics and politicization of the judiciary, it is inevitable for senior officials and presidents to consider the fact that they may be somehow related to the swirls, which incentivizes them to appoint justices that ensure 'loyalty' to them. The candidate's ideological clarity is a more critical factor than neutrality or expertise when deciding a justice.



Graph 2

Recreated graph of The Seoul Shinmun Daily - 한규섭의 데이터 정치학; ‘정치의 사법화’가 촉발한 ‘사법의 정치화’... 법원 신뢰 추락 우려/서울대 언론정보학과 교수(정치커뮤니케이션)

Conclusions and discussions

This research aims to identify the causes of the politicization of the judiciary in South Korea and its impact on the country's democracy. In conclusion, there is a positive correlation between political polarization and the judicialization of politics, suggesting that the politicization of the judiciary results from this relationship. Additionally, the study shows that the media's influence on political matters, particularly individual journalists' biases and reporting styles, may unexpectedly contribute to the extent of judicial politicization.

This paper is significant because it identifies political polarization as a critical driver behind the politicization of the judiciary. By establishing this link, the research opens avenues for future exploration of practical solutions that mitigate these issues. Furthermore, the study highlights the threats facing South Korean democracy, emphasizing the urgency of these challenges in light of the country's historical struggle to attain and preserve its democratic institutions. The research aims to raise public awareness about these threats and their potential consequences.

While this study tries to identify the risks to Korean democracy posed by the politicization of the judiciary and its relationship with the judicialization of politics, it acknowledges the need for further exploration from diverse perspectives. The paper exhibits the characteristics of a review article and only proposes developed solutions to the identified problems. Future research should focus on practical mechanisms for addressing these challenges

by delving deeper into relevant Korean laws or assessing public sentiment and current views toward the political landscape.

The ongoing political polarization in South Korea contributes to the uprising of political integrity, leading to political accusations among politicians prioritizing mobilizing their core supporters over fulfilling their duties. The focus of Korean politics has shifted towards populism and the exposure of corruption rather than pursuing policies that improve citizens' lives. The politicization of the judiciary is further enabled by a general lack of public engagement in political matters, particularly in the aspects that truly matter. This exemplifies the critical role that citizens must play in political processes. They must stay informed about significant political events and seek to understand the underlying purposes of these developments, ensuring that they cast responsible and informed votes. The media also plays a crucial role in this issue. Media must hold responsibility for their reports and ensure they broadcast facts and truths without allowing their political biases to shape the narrative. By doing so, they can contribute to a more informed public and a healthier democracy.

Works Cited

- “Attorneys General of the United States | AG | Department of Justice.” www.justice.gov, www.justice.gov/ag/historical-bios.
- Bellamy, Richard. “The Political Form of the Constitution: The Separation of Powers, Rights and Representative Democracy.” *Political Studies*, vol. 44, no. 3, Aug. 1996, pp. 436–56, <https://doi.org/10.1111/j.1467-9248.1996.tb00593.x>.
- Copp, David. “Explanation and Justification in Ethics.” *JSTOR*, 1990, www.jstor.org/stable/2380995.
- Della Torre, Lucia. “From the ‘Judicialization of Politics’ to the ‘Politicization of Justice’ in the UK and Switzerland.” *BORDER CROSSING*, vol. 6, no. 2, July 2016, pp. 118–35, <https://doi.org/10.33182/bc.v6i2.487>. Accessed 20 Dec. 2019.
- Ely, John. “Constitutional Interpretivism: Its Allure and Impossibility.” *53 Indiana Law Journal* 399 (1978), vol. 53, no. 3, Apr. 1978, www.repository.law.indiana.edu/ilj/vol53/iss3/2/.
- Hirschl, Ran. “The Judicialization of Politics.” *Oxford Handbooks Online*, Oxford University Press, 2011, <https://doi.org/10.1093/oxfordhb/9780199604456.013.0013>. Accessed 29 Oct. 2020.
- 정치의 사법화, 사법의 정치화. “정치의 사법화, 사법의 정치화.” Lawtimes.co.kr, 2024, www.lawtimes.co.kr/opinion/184590. Accessed 9 Aug. 2024.
- ‘정치의 사법화(司法化)’와 민주주의 박 恩 正. Vol. 51, no. 1, 2010, pp. 1–26, s-space.snu.ac.kr/bitstream/10371/71010/1/0x701936.pdf. Accessed 9 Aug. 2024.
- 정유민. 8 Apr. 2024, www.sedaily.com/NewsView/2D7TZPHGGI.
- Miller, Mark C., and Pol Sci. “Miller, Mark C. *Judicial Politics in the United States*. Westview. (‘Text’). Miller, Mark C., Editor. *Exploring Judicial Politics*. Oxford University Press. (‘Reader’). Carter and Burke. *Reason in Law*, 8 Th Edition. Longman.” Oxford University Press, Jan. 2015. Accessed 9 Aug. 2024.
- 오명숙. “[한규섭의 데이터 정치학] ‘정치의 사법화’가 촉발한 ‘사법의 정치화’... 법원 신뢰 추락 우려/서울대 언론정보학과 교수(정치커뮤니케이션).” *서울신문*, *서울신문*, 31 Dec. 2023, www.seoul.co.kr/news/editOpinion/opinion/politics-hgs/2024/01/01/20240101021001. Accessed 9 Aug. 2024.
- Park, Jonghyun. *The Judicialization of Politics in Korea*. 2008, manoa.hawaii.edu/aplpj/wp-content/uploads/sites/120/2011/11/APLPJ_10.1_park.pdf.
- Persson, Torsten, et al. “Separation of Powers and Political Accountability.” *The Quarterly Journal of Economics*, vol. 112, no. 4, Nov. 1997, pp. 1163–202, <https://doi.org/10.1162/003355300555457>.
- “The Growing Politicization of the US Supreme Court.” *Eagleton Political Journal*, eagletonpoliticaljournal.rutgers.edu/growing-politicization-of-the-us-supreme-court/.
- Waldron, Jeremy. “The Rule of Law and the Importance of Procedure.” *Getting to the Rule of Law*, Sept. 2011, pp. 3–31, <https://doi.org/10.18574/nyu/9780814728437.003.0001>. Accessed 29 Mar. 2023.
- 윤영미. “DBpia.” *헌법논총*, vol. 29, 2018, pp. 357–98, www.dbpia.co.kr/journal/articleDetail?nodeId=NODE09014695. Accessed 9 Aug. 2024.

The Evolution of French Profession Titles and Feminism By Xinyi Li

Abstract

French is a gendered language, and people would address each other differently according to their genders. Decades ago, however, most professional titles in French only had masculine forms. It was after the feminist movements that the feminine forms started to evolve, and became gradually accepted by society. This paper traces the history of this problem, including the underlying causes, the particulars during the movement, and the implications for gender-inclusive language in the future, including the feminization of job titles. Specifically, this paper will address gender in French and examine the reasons behind the absence of the neutral form and the subsequent establishment of the masculine form as the default gender. Then, connecting language use and the spread of social ideologies, this paper ties the evolution of the French language with the development of feminist ideas. With this as the basis to understand the entire movement, this paper continues to identify some important factors that resisted the change, and the activists' movements to counter it. This research concludes by examining the prospects for gender-inclusive languages in French within the context of non-binary gender.

Keywords: Feminization of Job Titles; French Language; Gender inclusivity; Linguistic Evolution; Feminist movements

1. Introduction

French is a Romance Language that, unlike English, distinguishes gender in its words as either masculine or feminine. Every noun inherently owns a grammatical gender (for example, the feminine article “la” should be used for the word “table”). This linguistic structure also extends to the adjectives and the other specific nouns that describe people by changing their form according to the actual gender of their subjects, as in the case of “actor” and “actress” in English. Naturally, as a prominent set of nouns with gender, professional titles in French incorporate both masculine and feminine forms.

This representation of both genders, however, is not the case throughout the long history. Females have been underrated in the job market, and it would be imaginable for a woman to be an engineer, an author, or a teacher in the past. Professions like those were historically dominated by men. Hence, only masculine forms of those words existed yet none for feminine ones, signifying the discrimination women faced in their workplace. It wasn't until recent decades with the burgeoning feminist movement advocating for female rights that the previous situation of men dominating the labor forces or the intelligentsia had broken. Since then, the percentage of women being educated and working has increased, gradually shattering the previously set biased standard. At this moment, addressing a female colleague using a masculine professional title would be improper. As a result, feminine forms of professional terms concurrently with feminist movements. In the beginning, such transformation was small-scale – it is limited to a few words with constrained impact. After the development of the Internet, however, people started to use

feminine forms of professional titles spontaneously. Despite disputes and struggles, they are now finally widely accepted by the public and even adopted by authorities.

The acceptance of the feminine forms of occupation titles marked an important transformation in society: the acceptance of females in the workplace. Thus, this event highlights the connection between the evolution of human languages and social ideologies. The human language has historically been used to marginalize females, but with time, it has changed and is now able to break past prejudices.

This paper will explore the transformation of feminine forms of professional words in French and examine its link with Feminism. In detail, this paper will focus on the history and its contents, examine the movements in word transformation, and predict the future developments in word forms. Throughout this exploration and the special lens of French, we aim to learn more about the crucial role that language could play in societal reforms and how those linguistic contribute to the goal of gender equality.

2. Gender in French

Before we look into the details of the issue related to the feminization of professional titles in French, it is important to introduce the foundation of this problem: the genders in French. In English, it is normal to introduce a female by saying, “She’s a teacher”, but doing the same thing in French, “Elle est professeur”, raises concerns. The reason lies deep in a key difference: English words are gender neutral, but there are two genders in French, feminine and masculine. Using masculine forms, such as “professeur” in the formal example, to describe females, makes them underrepresented as feminists believe.

It is well-established in linguistics that all Romance Languages are derived from Vulgar Latin, which is gendered, and French is one of the Romance Languages. The gender rules in Latin therefore apply in French. However, neutral words exist in Latin known as neuter words, but after evolution, these neutral terms became absorbed into either the masculine or the feminine category. After a systematic comparison and analysis of the Latin and Old French word roots, Polinsky and Everbroeck concluded that since neuter words are linguistically unstable and that in Late Latin, the changes in the pronunciation of words, leading to the indistinguishability of neuter words compared to their masculine and feminine forms in consonants and vowel length, caused the gradual assimilation of the three genders into two (Polinsky and Everbroeck).

3. Why the Two-Gendered System Became Problematic

The shift from a three-gendered system into a two-gendered one alone would not cause problems if suitable descriptions were used to address females correctly. It was the continued preference for men, as seen by the long-standing phenomenon in French which dates back to the 17th century and persists to recent decades – the marginalization of feminine words, and therefore the female gender. Even in situations when females are the majority, the default grammar rules still apply to masculine forms. For example, for a group of women and one man, the correct grammar would be using “ils”, or the masculine form of “they”, to describe them, instead of

“elles”, the feminine form; in that same scenario, the noun-adjective agreement in French also supports the usage of adjectives in their masculine forms, such as “contents”, the masculine plural form for “happy”, instead of “contentes”, the feminine plural form (Yaguello). Going towards the specific topic of the domination of masculine profession titles, for a long time in history, words such as "médecin" (doctor), "ingénieur" (engineer), and "avocat" (lawyer), in addition to the previous "professeur" example, did not have their feminine counterpart (Rhodes-Robinson).

Most people when would explain this male domination by mentioning the absence of female roles in the job market, as it is a general trend of barely educated women being housewives instead of participating in prestigious jobs. However, scholars like Éliane Viennot propose another perspective on the problem. She proposes that, back in the 16th century, feminine forms of professional titles owned an equal status as their masculine forms, as represented by a common term back then “autrice” (female author) (Viennot). As the 17th century dawned, however, dominated by the monarchy of the royals, major Grammarians such as Malherbe and the official French Language organization, l’Académie Française, began to impose rules that tended toward masculine forms to establish a masculine rule (Viennot). They reshaped the French language to purify and clarify it, and also centralized the male-dominated cultural and political power in France (Fitzsimmons).

4. Language and Social Ideologies

The main concern about the masculine–dominated French is that it underrepresented females, leading to the gradual degradation in female status. While the status of women in the earlier centuries may have been customary, beginning in the 20th century as feminist ideas started to gain traction, efforts were made to advocate for equal gender rights. Yet why would French feminists be specifically concerned about the underrepresentation of feminine forms in French?

It is important to understand the way language and ideologies interact. Language rules or phenomena can gradually influence people’s perception of society, as demonstrated by how the authorities back in the 17th century shaped the French to instill a mindset that favored males. Therefore, to break that long-standing mindset, breaking the implications that everyday language hints to its users is also a vital approach. In general, languages can reflect social ideologies and situations, for factors such as social classes and education opportunities could affect language use; on the other hand, languages can also be influenced by social events such as political and cultural movements (Rhodes-Robinson).

Hence, bringing changes to the French Language was a start for females to advocate for their rights, which was exactly what they did. Influenced by the major feminist movements, women started to enter previously male-dominated professions; this was the start of French profession title feminization, as it existed in everyday conversations.

5. Challenges Towards the Reform

5.1 The Official Opposition and L'Académie Française

To shift those activities from an unofficial, or even grassroots, status, to become officially promoted, advocacies from authorities are crucial. For a long time in history, however, authorities brought setbacks instead. This was the first challenge toward the feminization movement: the long-standing official opposition.

As the most authoritative organization that has been regarded as the guideline of the “correct French” since its establishment in 1635, l'Académie Française historically resisted actions that could raise the importance of feminine word forms, including feminizing job titles (Warren and Jellenik). It was because as an organization with a long history, l'Académie Française was conservative and feared that any changes to their current language would violate the tradition and cause complexity to the language. Additionally, the Académie believed the French language is closely connected to the national identity of France, and any linguistic changes, including the feminization of job titles, could harm the core French culture (Baron).

5.2 Difficulties in Use

Additionally, the actual imposing stage revealed challenges as well. A study of the discourse by French, Belgian, and Canadian media, including newspapers and magazines, in 2005 showed that even though feminine word forms were adapted with new forms constantly emerging, making it difficult for one standard word form to prevail. The study identified three ways to produce feminine forms linguistically, based on the masculine forms, yet the approaches to feminizing each word are hardly united even within the same publication (Baluta). Therefore, at that time, it could be hard for one linguistic form to be widely accepted in French-speaking society; it might also indicate difficulties for schools to select a standardized adopted language for teaching.

6. Feminization Efforts

6.1 Formal Unification

It could be impractical to want a government to accept an immature and random language system, as depicted previously in part 5.2, so first, actions were taken to unite the word usage and improve the situation. Scholars such as Skye Rhodes-Robinson from the 2000s conducted research and asked French speakers to write the feminine version of sentences their own way, and by combining their results with the prior language use, they proposed generic feminization rules that are most likely being widely accepted (Rhodes-Robinson). Meanwhile, public guidelines were also published. The French Ministry of Culture released an official document, *Femme, J'écris ton nom* to remind the public of the long-existing feminine forms that were neglected due to political and social reasons as an official and established rule (Becquer et al.). With such efforts, some of the most commonly used words gradually had their feminine forms developed and stably put into use.

6.2 Official acceptance

Individuals and organizations also worked hard to earn acceptance from authoritative organizations. A representative of the major advocates, Laurent Fabius from the “Commission de féminisation des noms de métiers et de fonctions”, issued a proposal sent directly to the government about the usage of feminine terms when addressing women (Baluta).

Inside l’Académie Française, supports were gradually granted. This process, as illustrated in part 4, had a strong link with the political status in France at that period. Before 1980, there were only eight females in the Académie; this underrepresentation of females in the organization could have contributed to the stubbornness of its attitude (Warren and Jellenik). Stepping into the 90s in the 20th century, female participation in politics became more significant, as seen by the appearance of the “feminine political persona”, which raised the importance of females’ voices (Burnett and Bonami). Two proposals for language change in 1986 and 1996 can prove this connection. Despite being similar in content, the two proposals received different results, as the former was rejected and the latter was accepted, and the basis of feminine ideology did not exist back in 1986 (Burnett and Bonami). Finally, the attitude of l’Académie Française softened in 2019, granting permission for the feminization of professional job titles (Warren and Jellenik). In addition to publishing a note stating that the Académie’s acceptance of a report on the feminization of job names by the Académie on its official website, it also posted the report that acknowledged the need for feminization despite its difficulty and the need to conform to flexibility and pragmatism during the gradual change (Académie française).

7. Public Attitude and Responses

The feminization movement until now earned mixed results, but strong evidence for further action prevails. For one thing, a study of three major French newspapers demonstrates that even though the usage of feminine forms increased, those changes were mostly limited to lower-ranking jobs, implying that the connection between feminine word usage and enhanced female status may be harder to establish than anticipated (Baluta). Disagreement towards this reform exists within the French-speaking public as well. While some expressed approval of the effort, others hesitated to accept the new feminine word forms (Van Compernelle). Some people compared it to some more pressing issues such as equal pay, raising doubts about whether such change in word forms could invoke actual changes in female status (Van Compernelle).

8. Future Developments

Job title feminization required further improvement and actions, but it is not the only question concerning politics and linguistics in French. As proposals and awareness of non-binary genders rise these days, the need for creating neutral forms also exists. Regarding this area, experts such as Yaguello express their concern: the neutral word form was abandoned long ago, and the current attempts may complicate the text, and adjectives, in addition to nouns, are also gendered, making it harder to create neutral words (Yaguello). Despite such concerns, linguistics still proposes present constructive advice in creating gender-neutral terms and promoting further

equity (Viennot et al.). Bézert from the Canadian Digital Service summarizes three common ways to neutralize French: by using abbreviated doublets (adding feminine and plural endings to masculine words, such as *les avocat-e-s*), replacing gendered term with neutral terms (such as using individual over man or women in English), or generating non-binary nouns such as *iel* for *il* (he) and *elle* (she) (Bézert). Such approaches, however, may receive more opposition than before, as it may be seen as a further violation of the French culture. The president of France, Emmanuel Macron, has publically stated that the masculine words represent neutral, and there is no need to add any dots within, even showing support to ban gender-inclusive words from official texts at the inauguration of the French language center in Villers-Cotterêts (*REPLAY: Macron Inaugurates French Language Museum in Far-Right Heartland • FRANCE 24 English*).

9. Summary

The focus of this paper is to track the history of the issue concerning the feminization of French professional titles and peek through the future development of this problem along with the development of personal identity and rights. French has historically favored masculine forms, particularly in professional titles, reflecting and reinforcing male dominance in the industry as a language steeped in gender inequalities. Despite opposition from conservative French authorities, calls for the invention and use of feminine forms of professional titles have grown with the rise of the feminist movement in the 20th century. These initiatives eventually resulted in feminized titles being officially recognized and gaining wider social acceptance, which came to represent the acceptance of women's participation in the workplace. Despite the acceptance from the Académie Française, there are also great ongoing challenges towards this problem, such as difficulties in standardizing new feminine forms, mixed public reactions, and the continued underrepresentation of women in higher-ranking professional titles. Furthermore, the discussion about gender-neutral language in French may still be at the starting point, and more efforts should be put to propel any changes.

Works Cited

- Académie française. *La Féminisation des Noms de Métiers et de Fonctions*. Académie française, p. 20.
- Baluta, Camelia. *Linguistic Analysis of the Feminization of Professional Titles in French: Evidence from French, Canadian, and Belgian Press*. 2005. The University of Florida, Report, <https://ufdcimages.uflib.ufl.edu/UF/E0/01/33/68/00001/UFE0013368.pdf>.
- Baron, Dennis. “French Academy Sees Linguistic Diversity Endangering National Identity”. 20 June 2008, <https://blogs.illinois.edu/view/25/4530>.
- Becquer, Annie, et al., editors. *Femme, j’écris Ton Nom: Guide d’aide à La Féminisation Des Noms Des Métiers, Titres, Grades et Fonctions*. Documentation française, 1999.
- Bézert, Marie-Sophie. “Writing without Excluding: Inclusivity in the French Language.” <https://digital.canada.ca/2023/03/20/writing-without-excluding-inclusivity-in-the-french-language/>.
- Burnett, Heather, and Olivier Bonami. “Linguistic Prescription, Ideological Structure, and the Actuation of Linguistic Changes: Grammatical Gender in French Parliamentary Debates.” *Language in Society*, vol. 48, no. 1, Feb. 2019, pp. 65–93. DOI.org (Crossref), <https://doi.org/10.1017/S0047404518001161>.
- Fitzsimmons, Michael P. *The Place of Words: The Académie Française and Its Dictionary during an Age of Revolution*. Oxford University Press, 2017. DOI.org (Crossref), <https://doi.org/10.1093/oso/9780190644536.003.0008>.
- Polinsky, Maria, and Ezra Van Everbroeck. “Development of Gender Classifications: Modeling the Historical Change from Latin to French.” *Language*, vol. 79, no. 2, June 2003, pp. 356–90. DOI.org (Crossref), <https://doi.org/10.1353/lan.2003.0131>.
- “REPLAY: Macron Inaugurates French Language Museum in Far-Right Heartland”. FRANCE 24 English. 2023, <https://www.youtube.com/watch?v=6vXYy98F4p4>.
- Rhodes-Robinson, Skye. *The Feminization of French Profession Words*. 2007. Bryn Mawr College, https://www.swarthmore.edu/sites/default/files/assets/documents/linguistics/2008_rhodes_robinsonskye.pdf.
- Van Compernelle, Rémi A. “What Do Women Want?: Linguistic Equality and the Feminization of Job Titles in Contemporary France.” *Gender and Language*, vol. 3, no. 1, Aug. 2009, pp. 33–52. DOI.org (Crossref), <https://doi.org/10.1558/genl.v3i1.33>.
- Viennot, Éliane, et al. *Le Langage Inclusif: Pourquoi, Comment*. Éditions iXe, 2018.
- Viennot, Eliane. *Non, Le Masculin Ne l’emporte Pas Sur Le Féminin!* Éditions Ixe, 2014.
- Warren, Meredith, and Cathy Jellenik. “L’Académie Française and Gender Disparity: The Need for Female Immortals.” *CLA Journal*, vol. 7, 2019, p. 10.
- Yaguello, Maria. *Les Mots et Les Femmes: Essai d’Approche Sociolinguistique de La Condition Féminine*. Éditions Payot & Rivages, DL.

Understanding Global Population to Understand International Relations

By Baron Jiang

Introduction

Our planet has seen unprecedented change in the contemporary era, with countries around the world experiencing political and economic development at unparalleled scales. The many modern innovations of humankind have created a world where cultures transcend geographical boundaries—where ideas and goods seemingly spread instantaneously from one place to another. With this period of prosperity comes an increasing awareness of rising contemporary issues; climate change looms as a heavy burden for current and future generations to shoulder, advocacy for human rights has been bolstered by political turmoil in states across the world, and incidences of war seem to be on the rise. At the heart of these issues are the collective people themselves. Thus, defining how the global population changes over time is key to solving the world's problems in a span of technological revolution. To inquire upon this growing issue, the crisis must be understood through nuanced analysis of the factors that affect population growth and decay, most notably in demographics and environmental sustainability. We may never reach a definitive answer for what a sustainable carrying capacity looks like on our planet; as proven in our past, the future is unpredictable and unreliable. However, only through careful examination of these aforementioned factors can we truly begin to understand the subtleties of the global population's effects on the international sphere.

Demographics

The global population has been rising at an exponential rate in the past couple of centuries, unprecedented by any other time in history. In fact, the world population did not hit one billion until the early 1800s, followed by two billion around a century ago—and since then, our world's population has quadrupled to eight billion (Ritchie et al. 2023). By extrapolating past trends with geographical models, it is predicted that our population will peak at around 2086, with an expected figure of 10.4 billion (Wade et al. 2024). However, demographers typically only make predictions up until the turn of the century, and it is difficult to know how quickly the process of depopulation is to occur afterwards (Spears 2023). Following this peak, our population faces an imminent decline at an unpredictable rate. Toward this end, the demographic transition model can be used to help predict future growth or decline.

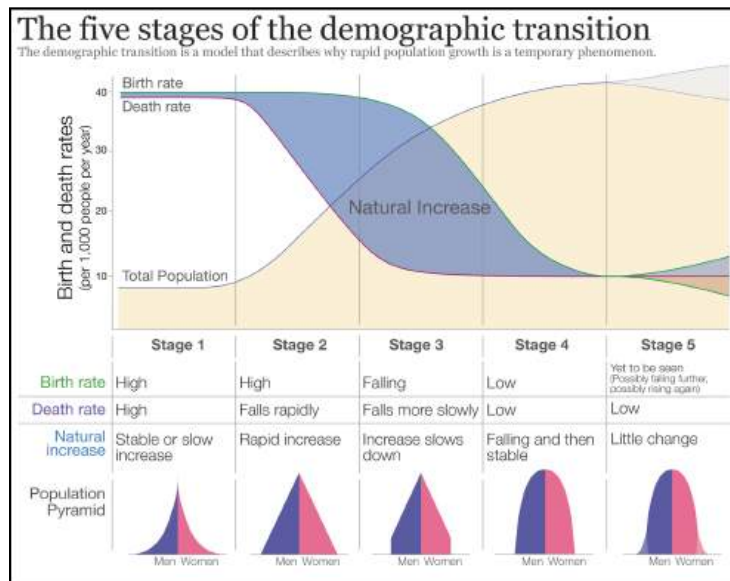


Fig. 1: Demographic Transition Model.

Taken from ourworldindata.org/demographic-transition

As depicted in Fig. 1, the natural increase rate—essentially the difference between the crude birth rate and the crude death rate—increases and then decreases as a country’s population proceeds through the stages of the model. Populations start at the high stationary stage, where very little population growth or decline occurs and the populace has a low life expectancy. No country remains at this stage, save for extremely isolated indigenous populations in the Amazon or in Sub-Saharan Africa (Scanlan 2024). Afterwards, populations enter the early expanding stage of rapid population growth, characterized by industrialisation and a rising life expectancy. Following this, in the late expanding stage the population continues to see growth, albeit at a lesser pace, as industrialisation occurs fast and life expectancy rapidly increases. The fourth stage is where growth seems to significantly slow down; there is very little population growth as countries become fully industrialized and are making a shift towards the services sector in their economic industries. During this period, the population is balanced between young and old, but is beginning to age.

Ultimately, the fifth stage is arguably where some of the most developed countries in the world currently lie. The population slowly declines as death rates are low but steadily rising due to an aging population, further illustrating that the world is currently headed into a period of population decline. The dependency ratio would thus be higher in more developed countries (MDCs) that are experiencing the later stages of the demographic transition model, and inversely so for less developed countries (LDCs) that have undergone population booms in recent decades. In MDCs, aging populations can be a burden on the workforce as countries may feel obligated to

extend the age of retirement to reduce the dependency ratio and enact new policies to accommodate the older population; for instance, take Japan's current population conundrum. Japan is the prime example of an aging population; it has a median age of 48, with over a third of its citizens being 60 or older (Ausubel 2020). In the majority of Japan's political institutions, young people are essentially marginalized as older politicians are the predominant assemblymen and policymakers. Because of this, there are worries that "the decisions made by mostly older politicians will lead to welfare policies that favor the elderly at the expense of younger families" (McClellan 2020). Younger generations will be heavily impacted by the decisions of politicians that may prioritize the needs of older generations over other long-term issues. Furthermore, in contrast to LDCs, fertility rates in MDCs are much lower. This can be attributed to improved healthcare and a change in social values, particularly towards women. MDCs provide women with better opportunities and access to education, decreasing birth rates as more women begin to prioritize pursuing their careers rather than starting a family (OECD 2011). Additionally, access to contraceptives will logically result in less children being birthed. As of 2021, the United States has a fertility rate of 1.66 births per woman. If the entire world had this same exact fertility rate, after peaking at 10 billion, the global population would drastically fall to under two billion people in three centuries' time (Spears 2023). As countries make improvements on their government, infrastructure, and economy, they slowly shift from a developing country to a developed one. It has already been established that MDCs experience a trend of population decline as they move into the fifth stage of the demographic transition model; therefore, the world as a whole seems to already be heading in a direction of depopulation.

Moreover, demographics can be heavily affected by population policies that governments enact to either increase or decrease their state's population—with varying effects. The greatest modern example of this occurred in China. With about 1.41 billion people, China is the second most populous country in the world (World Bank 2022). However, the Chinese population is aging at a high pace; China's population growth turned negative a couple years ago. The state's working age population is also shrinking, which can be directly attributed to the damaging effects of its anti-natalist policies from the past. China's one child policy was designed in the '70s and '80s as an attempt to reduce the state's uncontrollably increasing population growth rate, limiting families to just one child. Propaganda at the time promoted financial benefits or better jobs for those with one child, while fining families with more children. The policy also resulted in sex-selective abortions because of a preference for boys in Chinese culture (Mullen 2021). Due to this, there are millions more men than women in their childbearing years today. China's negative natural increase rate is also driving the country into a population crisis, with the looming threat of depopulation ever-present as a worry in the coming years. The policy was eventually scrapped in 2015, but perhaps much too late to reverse the damaging aftereffects of a government pushing unethical antinatalist strategies: "China could have completely scrapped population controls in 1980 and growth would have moderated naturally" (Kuo & Wang 2019). Overall, population shifts in countries are difficult to deal with, and political policies that force a

populace into taking certain measures in an attempt to control the population size may not be the best solution, as shown by China's failures to preempt and foresee their current crisis.

Agriculture and Environmental Sustainability

British philosopher Thomas Malthus first contemplated this population growth issue in the late 18th century when he formed his own Malthusian Theory: while population increases exponentially, agricultural capacity only grows linearly. He assumes that agricultural capacity would not be able to keep up with the rapidly expanding global natural increase rate, and that this eventually leads to depopulation, famine, and social unrest. Additionally, limits in agricultural carry capacity will effectively set a barrier in which the human population will fail to exceed. However, the contemporary problem with his theory is that it fails to account for modern innovations in the agricultural sector. The Second Agricultural Revolution and Green Revolution both greatly increased agricultural capacity through introduction of modern machinery, chemical fertilizers and GMOs (Thompson 1968; Wu et al. 2004). Yet, it would not be entirely correct to say that agricultural factors do not pose a threat to the global population in the modern era. The global arable land per capita has seen a significant decline as countries further develop their urban cores; between 1961 and 2021, arable land per person decreased by over 50%, from 0.36 hectares to 0.18 hectares (World Bank 2021). A decrease in arable land also means an increase in physiological population density, resulting in a higher number of people who are dependent on less space to grow their food. Because of this, agricultural productivity would need to grow, leading to agricultural methods that are reflective of modern intensive farming practices, which harm the ecosystem by degrading the soil and using a lot of chemical pesticides.

Furthermore, the planet's resources are finite. Freshwater is limited in areas across the globe due to the geographical positions of countries, especially in arid regions like the Middle East. Kuwait almost exclusively relies on the process of desalination to produce clean drinking water for its population, as it is extremely freshwater-scarce; the country has virtually no renewable internal freshwater resources per capita (World Bank 2020). However, desalination is an extremely expensive process and is not considered a long-term sustainable solution. Fossil fuels will also eventually run out as the world is heavily dependent on coal, oil, and natural gas to produce energy. If current consumption rates persist, natural gas will deplete the fastest at 49 years, followed by oil and coal at 56 and 139 years respectively (Energy Institute 2020). Overall, the Earth's dwindling resources support the theory of global depopulation. A region with more people produces more waste and pollution while consuming more resources—in fact, the 20 most polluted cities in the world are in Asia (Vanzo 2024). On the other hand, less people results in balanced consumption rates of food and natural resources, reducing the prominence of these worries and allowing for people to instead focus on achieving further innovation that provides long term solutions towards maintaining population size.

Conclusion

After analyzing the nuances in the demographics, agricultural, and environmental sustainability of different countries, it seems that the world is headed towards a decline in the global population. Although, this poses another important question: is this depopulation we face objectively good for the planet? Seemingly, the problem is that we've never experienced such a mass population decline occurring everywhere at the same time, so no one knows exactly what to expect. It is hard to say what states can do to predict the outcomes of their changing population sizes, but it is clear that different regions of the world will be distinct in how they experience this issue. Countries like Japan and China must pay more heed to their dwindling populations, particularly their working age. Comparatively, countries in Sub-Saharan Africa are still experiencing rapid population growth, but as they evolve into more developed countries, they will likely experience the same issues that MDCs undergo today, albeit decades into the future. Ultimately, countries and their governments must pay more attention to the issue of population growth and decline in the coming years, otherwise humanity as a whole will be unprepared for what could come in the next few decades.

Works Cited

- Ausubel, Jacob. "Populations Skew Older in Some of the Countries Hit Hard by COVID-19." Pew Research Center. 22 Apr. 2020. <https://www.pewresearch.org/short-reads/2020/04/22/populations-skew-older-in-some-of-the-countries-hit-hard-by-covid-19/>. Accessed 7 Sept 2024
- Energy Institute. "Years of Fossil Fuel Reserves Left, 2020." Graph. Our World in Data. <https://ourworldindata.org/fossil-fuels>. Accessed 7 Sept. 2024.
- Hickson, William Edward. *Malthus: An Essay on the Principle of Population in Refutation of the Theory of the Rev. T.R. Malthus*. Taylor, Walton, and Maberly, 1849. <https://jstor.org/stable/60249894>. Accessed 7 Sept. 2024.
- Kuo, Lily and Xueying Wang. "Can China Recover from its Disastrous One-Child Policy?" *The Guardian*, 2 Mar 2019. <https://www.theguardian.com/world/2019/mar/02/china-population-control-two-child-policy>. Accessed 7 Sept. 2024.
- McClellan, Charles. *Silver Democracy: Youth Representation in an Aging Japan*. 2020. U of California San Diego, PhD dissertation. <https://escholarship.org/uc/item/54d3f8bz>. Accessed 7 Sept 2024.
- Mullen, Andrew. "China's One-Child Policy: What Was it and What Impact Did it Have?" *South China Morning Post*, 1 June 2021. <https://www.scmp.com/economy/china-economy/article/3135510/chinas-one-child-policy-what-was-it-and-what-impact-did-it>. Accessed 7 Sept. 2024.
- OECD. *The Future of Families to 2030*. OECD Publishing, Paris, 2011. <https://doi.org/10.1787/9789264168367-en>. Accessed 7 Sept. 2024.
- Pimentel, David et al. "Overpopulation and Sustainability." *Frontiers in Ecology and the Environment*, vol. 4, no. 3, 2006, pp. 155-61. <http://www.jstor.org/stable/3868686>. Accessed 7 Sept. 2024.
- Ritchie, Hannah et al. "Population Growth." *Our World In Data*. 2023. <https://ourworldindata.org/population-growth>. Accessed 7 Sept. 2024.
- Scanlan, Edmund. (2024). "2.5 The Demographic Transition Model." *Fiveable*. 18 June 2024. <https://library.fiveable.me/ap-hug/unit-2/demographic-transition-model/study-guide/qsRzryeP7QdCmPtXs4Wd>. Accessed 7 Sept. 2024.
- Spears, Dean. "The World's Population May Peak in Your Lifetime. What Happens Next?" *New York Times*. 18 Sept. 2023. Op-ed. <https://www.nytimes.com/interactive/2023/09/18/opinion/human-population-global-growth.html> Accessed 7 Sept. 2024.
- Thompson, F.M.L. "The Second Agricultural Revolution, 1815-1880." *The Economic History Review*, vol. 21, no. 1, 1968, pp. 62-77. <https://doi.org/10.2307/2592204>. Accessed 7 Sept. 2024.
- Wade, Matt & Angus Holland. "The Planet's Population Will Get to 10.4 Billion – Then Drop. Here's When We Reach Peak Human." *The Sydney Morning Herald*. 17 Dec 2023. <https://www.smh.com.au/national/the-planet-s-population-will-get-to-10-4-billion-then-drop-here-s-when-we-reach-peak-human-20231213-p5er8g.html>. Accessed 7 Sept. 2024.

- World Bank. "Arable Land (Hectares Per Person)." *World Development Indicators*. 2021.
<https://data.worldbank.org/indicator/AG.LND.ARBL.HA.PC>. Accessed 7 Sept. 2024.
- - -. "Renewable Internal Freshwater Resources Per Capita (Cubic Meters). *World Development Indicators*. 2020. <https://data.worldbank.org/indicator/ER.H2O.INTR.PC>. Accessed 7 Sept. 2024.
- Wright, Quincy. "Population and International Relations." *The Annals of the American Academy of Political and Social Science*, vol. 188, 1936, pp. 318-28.
<http://www.jstor.org/stable/1020384>. Accessed 7 Sept. 2024.
- Wu, Felicia & William P. Butz. *The Future of Genetically Modified Crops: Lessons from the Green Revolution*. RAND Corporation, 2004.
<http://www.jstor.org/stable/10.7249/mg161rc.11>. Accessed 7 Sept. 2024.

Getting Green: Roadmap to Reducing Costs for III-V Multijunction Solar Cells

By Andrew Carson

Abstract

The concern over the climate crisis and push to use renewable energy has led researchers to find ways to make existing energy technology more efficient. While feasible to make and use, the current cost of high efficiency, multi-junction solar cells leads them to be excluded from the market. This literature review aims to look at all costs in manufacturing multi-junction solar cells and show how long it will take them to make up the cost in energy, as well as suggest a path to reduce costs in the future while maintaining their high efficiency. Using articles that provide cost estimates for different manufacturing processes, this paper discusses how much money per unit of energy multi-junction cells provide and the factors that go into this statistic. This paper will also address predicted costs of unimplemented manufacturing methods to determine if any are viable for production of multi-junction cells and could make their price competitive with monocrystalline silicon solar cells. The results show that through increased MOCVD efficiency, semiconductor replacement and price reduction, material recycling, and further research, it is possible to reduce the price of multijunction cells significantly. However, it is worth considering that with the incredible rate monocrystalline solar cells have improved at, implementation of multijunction cells may take too long to be worth the increased efficiency.

Introduction

Human interaction with the environment, namely through the use of fossil fuels, has been shown to be a likely reason for the historically recent increase in global warming and large amounts of climate change. To protect the environment, many countries have put forth resources into researching and using renewable forms of energy, such as solar power.

How Solar Power Works

Solar power is generated from a solar cell, which contains a negatively charged n-type semiconductor layer and a positively charged p-type semiconductor layer. The n-type layer is heavily concentrated with electrons, which gives it its negative charge. The p-type layer is missing electrons which creates electron holes, giving the layer its positive charge. While generally locked in their molecule's electron shell, the electrons will not easily move to create energy. However, the photoelectric effect causes valence electrons to be emitted when the semiconductor layers are hit by light that covers a certain bandgap, such as the frequencies of light that comes from the sun. Now that they are more mobile, the electrons are drawn to fill electron holes and reach a state of equilibrium, and they will take the quickest path to reach an empty hole. The quickest path when the solar panel is immediately functioning is where the two layers make contact; however, the layers have an initial exchange of electrons and holes that creates a depletion zone, which emits a strong force that pushes electrons in the n-type layer to the edge of the panel and holes on the p-type layer to the edge of the panel. The extra electrons

from the negatively charged layer are then allowed to move freely and are drawn to the holes from a lack of electrons on the positively charged layer. Then by giving the two layers a conductive path for the electrons to travel through, it creates an electric current.

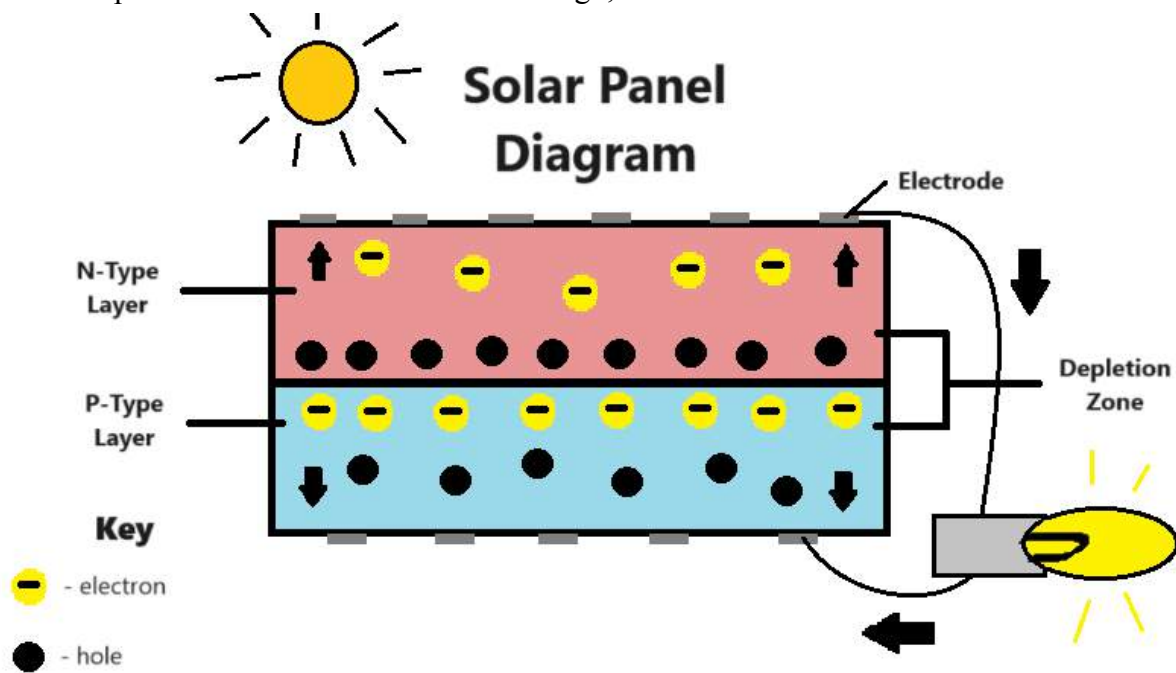


Fig 1: Diagram of a solar panel powering a lightbulb

Solar Technology Discussed

There are many different types of solar panels that have different uses, but this paper mainly considers the most common monocrystalline silicon solar cells compared to 3 junction III-V multijunction solar cells. III-V multijunction solar cells are made of multiple different semiconductor materials layered on top of each other while monocrystalline cells are composed of one semiconductor material. Each semiconductor used for multijunction solar cells has a different bandgap, which means different wavelengths of light excite electrons to generate power on each layer. This allows multijunction cells to cover more of the light spectrum so they're able to generate more electricity. Furthermore, they don't lose as much energy to heat from mismatched bandgaps, as all extra energy beyond the bandgap is converted to heat. By covering more of the light spectrum, III-V multijunction solar cells can have higher efficiencies compared to monocrystalline silicon solar cells. As seen in figure 2, before 2022 the highest efficiency reached by a single junction crystalline solar cell was 26.1%, while the highest efficiency for a multijunction solar cell was 39.46%. However, silicon based cells are significantly cheaper per unit of energy than multijunction cells. This gives multijunction cells niche use cases where they perform better because of their efficiency and size advantage, such as concentrator solar farms and space.

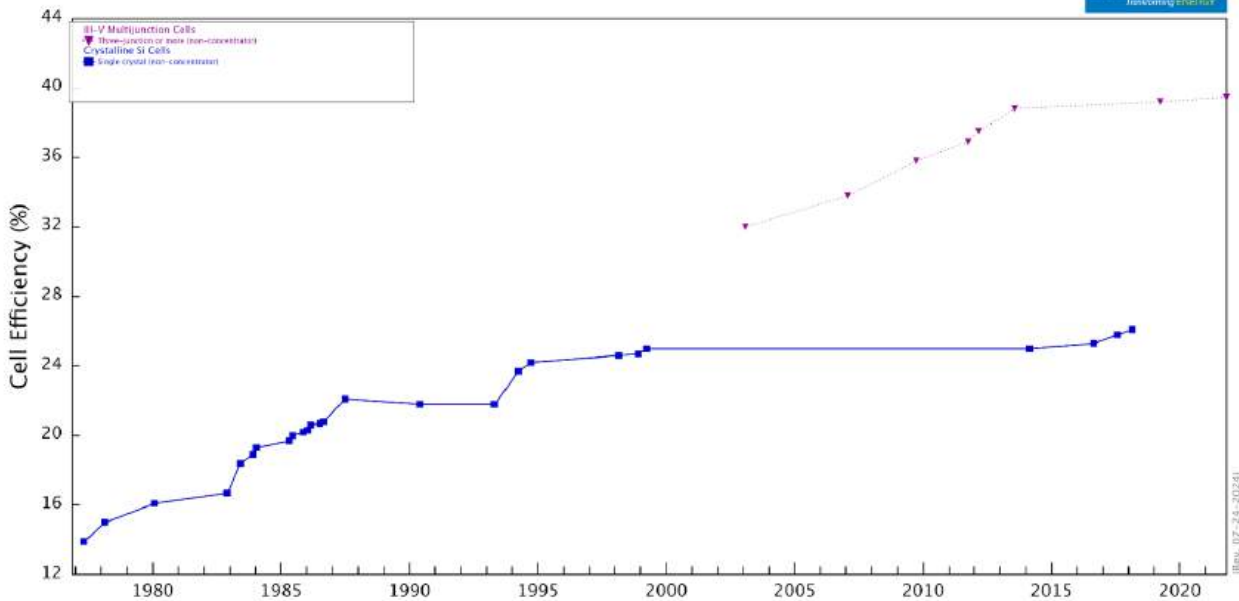


Fig 2: This plot is courtesy of the National Renewable Energy Laboratory, Golden, CO. This graph shows a comparison between the efficiencies of multijunction cells and monocrystalline silicon cells over time.

Manufacturing and Material Choice:

Manufacturing costs are one of the major factors driving up the cost per watt of III-V multijunction solar cells. Compared to the relatively simple construction of silicon based cells, multijunction cells are complicated; they require layers for each semiconductor material, conduction material run through to collect each layer's energy, and often need buffer layers to preserve the purity of the crystals when connecting mismatched lattices. Thus, the tools used to build multijunction cells are more expensive. Another important factor to consider when discussing the price gap between solar cells is material price. Silicon is incredibly cheap compared to semiconductors used in multijunction cells such as Gallium Indium Arsenide and Germanium. Although the price of semiconductor materials does fluctuate and depend on where you're buying them, the large price gap between the cells' materials is a stable and global phenomenon.

One of the most important factors in the price and efficiency of III-V multijunction cells is the semiconductor materials used. There are a few important physical qualities that need to be considered for each semiconductor, being bandgap and lattice constants. As discussed before, to cover as much of the light spectrum as possible, each semiconductor must have a different bandgap. Depending on the manufacturing method, the materials may have to be lattice matched as well, which means that the crystal structure of the semiconductors is aligned. These physical properties put some restraints on what materials can be used in tandem. The current most effective and most used combination of materials for a triple junction solar cell is GaInP/GaInAs/Ge (Germanium Indium Phosphide, Germanium Indium Arsenide, Germanium).

There are multiple ways to construct III-V multijunction solar cells, namely Metal Organic Chemical Vapor Distillation (MOCVD/MOVPE), Molecular Beam Epitaxy (MBE), and Wafer Bonding. MOCVD has applications in mass production of electronics like multijunction solar cells and can fabricate cells with high quality relatively quickly. It is important to consider that although MOCVD is currently the fastest way to mass produce III-V multijunction cells, its lengthy production process and small batch size leave much room for improvement. MBE is more often used in research where getting an extremely pure and efficient cell is most important, as MBE is slower and requires more extreme conditions than MOCVD. Wafer Bonding has recently been explored as a possibility in combining material substrates into a multijunction cell after manufacturing, and it has been proven to be able to work. However, the drop in efficiency and lack of implementation outside of laboratory testing shows wafer bonding requires more research before it can be considered a viable option for making cheap, high efficiency multijunction cells.

Review Goals

Compiling modern research about multijunction solar cells in an effort to reduce manufacturing costs is the main focus of this review, because if we are able to reduce the cost per unit of energy of III-V multijunction solar cells, there could be a large increase in solar efficiency that would accelerate the transition from fossil fuels to renewable energy. This review can also offer a plan to implement the infrastructure required to speed up the production of modern solar technology, which could lead to a long term price reduction in the cost of future solar manufacturing and further encourage renewable energy usage. The objective of this review is to determine if there is a possible alternative to monocrystalline silicon solar cells that are more efficient in terms of cost, efficiency, lifetime, and other metrics. If so, this paper aims to provide a clear roadmap for what it would take to switch to a more efficient technology.

The scope of the study is limited to the comparison of multijunction solar cells to monocrystalline silicon solar cells and not other solar cell types, as there was either not enough research and technology available to see implementation before serious climate consequences or going into them in the detail they deserve would not be possible with the size restraints of the paper. The data collected is viewed from a framework of attempting to increase solar efficiency by switching to new technology, so whether or not that is the absolute best course of action is not as heavily considered in this review. The methodology used to gather and analyze data includes using reliable databases and sources for finding scientific papers, cross checking research in data, and reviewing the sources for papers to make sure that research used was based on reliable facts.

Results (Strategies to Reduce Cost)

When looking at the cost in USD compared to watts generated of monocrystalline cells and multijunction cells, the problem becomes apparent. The price per watt of a 60 cell monocrystalline silicon PV module in the US as of 2018 was \$0.49 per watt (W). In 2017, the

price per watt of multijunction solar cells was \$8.24 per W. While this is a major price difference, there are many ways to save money when making multijunction cells.

Increasing MOCVD Efficiency

Before implementation is considered, it's important that the infrastructure for the production of III-V multijunction cells is in place. Despite being the best way to mass produce multijunction cells, MOCVD requires further research to meet the high demands of the photovoltaics market. However, there are serious advancements being made that can speed up the process and eventually increase efficiency to bring attention to multijunction cells as an alternative to monocrystalline silicon cells. By making the machines more efficient, manufacturing costs will quickly decrease because spending on machines will be significantly lower, especially with each machine going for between \$1-2.5 million.

Lowering Material Costs

Another way to reduce the cost per watt of multijunction cells is to increase the availability of the semiconductor materials that make them up. These materials are often extremely expensive to buy and relatively hard to price, as there are no large scale production multijunction cell manufacturing facilities. There are a few options to handle these high prices, one method being to find cheaper semiconductor materials. Research supports the possibility of using silicon on the bottom layer of multijunction cells for an efficient cost and power, so this may be an option. Another important idea to consider is that as multijunction cells become more thoroughly researched and well known, the materials used to produce them will increase in availability as more are produced. This would hopefully reduce the price of the expensive semiconductors.

Recycling Materials

A cost reduction method that is viable after the implementation of III-V multijunction cells is recycling the semiconductors after the cells reach the end of their lifetime. As previously mentioned, multijunction solar cell materials can be extremely expensive. However, recycling for some expensive parts of the cells is possible with few drawbacks. Due to the dangerous properties of some materials used, the recycling costs for multijunction cells can be higher than other solar technologies, but it can serve as a cost reduction in the long run.

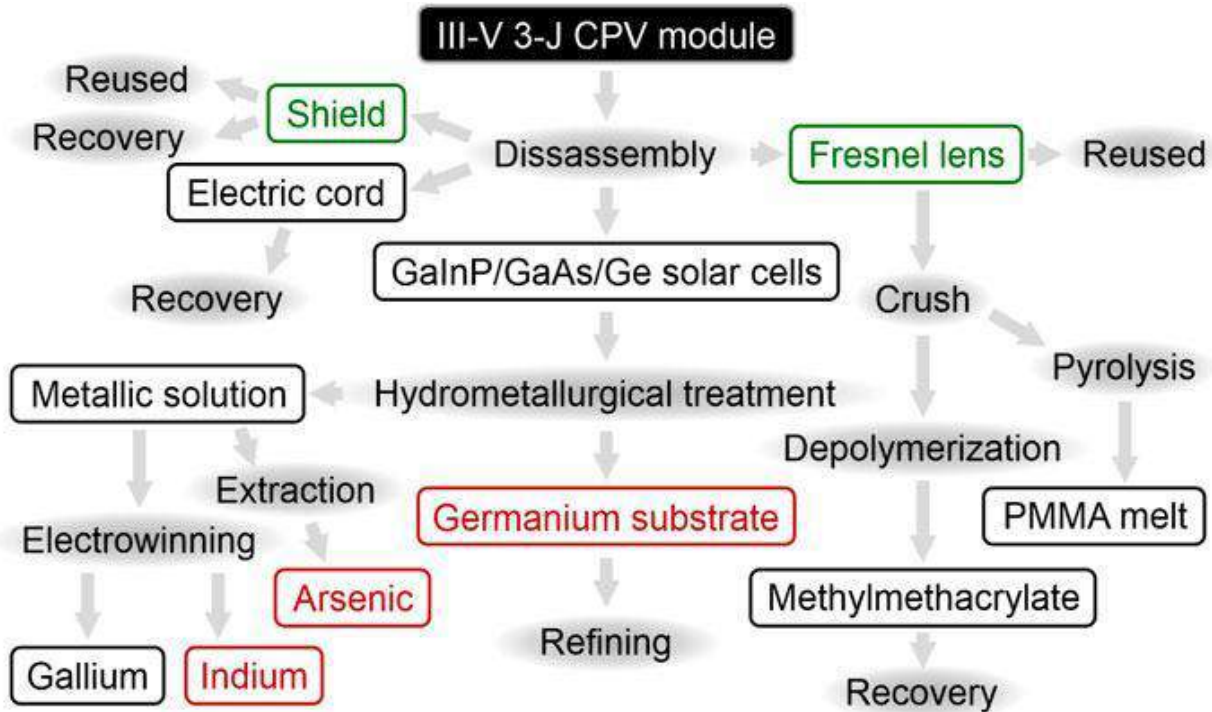


Fig. 3: Recycling process for 3 junction multijunction solar cell in detail. Image from “Status and challenges of multi-junction cell technology” on frontiers (Baiju A and Yarema M).

Future Research

As multijunction cells continue to be researched, a breakthrough in any part of their structure, manufacturing, or materials could push them over the edge to be the most efficient solar cell. Continuing the funding and efforts of research into III-V multijunction cells is the only way to improve upon what is currently possible with this technology. Our investments into research today can cut costs in the future, so for a long term cost goal, research is as crucial as the rest of the methods.

Predictions For Improvement

Although it's impossible to know when and where research will improve technology, we can take a look at how solar technology has improved over time while taking into account the Shockley-Queisser limit on solar efficiency to predict if III-V multijunction cells will improve at a fast enough rate to overtake monocrystalline solar cells. The Shockley-Queisser limit is a theoretical efficiency ceiling for solar panels based on the maximum energy able to be extracted per photon, and it allows us to know what our maximum solar efficiency can be. It is important to consider the maximum efficiency of both monocrystalline solar cells and multijunction solar cells in order to see if investing time into improving them will lead to high efficiency gains. In figure 4 below, you can see the efficiency gain of monocrystalline silicon solar cells used in regular solar farms compared to III-V multijunction cells used in concentrator solar farms modeled over time.

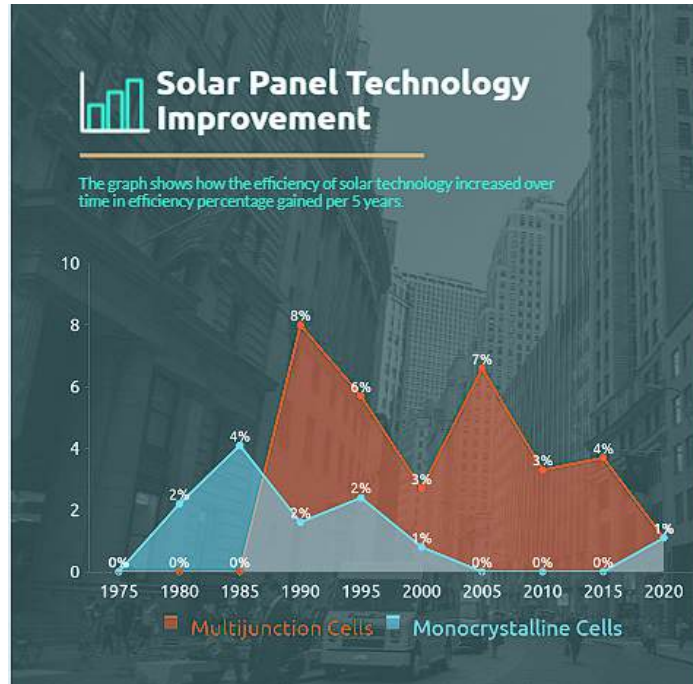


Figure 4. Graph shows solar panel efficiency increase in percentage over 5 year gaps.

Figure 4 shows us that III-V multijunction cells have improved efficiency more over time, while the monocrystalline solar cells have remained relatively stagnant over the past 20 years. Combining this with the fact that the Shockley-Queisser limit, or theoretical highest efficiency of monocrystalline solar cells, is 31.6%, there isn't much room left for these solar panels to grow in efficiency. Comparing this to multijunction solar panels which have grown rapidly in recent years, there is a high possibility that the efficiency keeps increasing because of a significantly higher ceiling to their efficiency. With that being said, it is unclear whether multijunction cells will be able to reach high enough efficiencies before the climate crisis reaches irreversible levels of damage. Combined with a lack of infrastructure for quick production or implementation, there is very little preparation for if multijunction cells are subject to a breakthrough in research. This leads to the conclusion that although multijunction cells are promising, it would take a significant movement to organize construction of necessary infrastructure to implement them before it's too late, and therefore should not be the focus of renewable energy as a solution for the climate crisis unless there is a major increase in affordability or efficiency. However, in the case that there is a breakthrough, it is important to have a plan for implementation, which leads directly into the roadmap.

Roadmap

The recommended steps to take for implementing multijunction solar cells would be to first improve the speed and batch size of the MOCVD manufacturing method. MOCVD is not only used for semiconductor manufacturing; its application in LED production is crucial as well. By improving upon this process, rare semiconductor production would increase, leading to faster

implementation of multijunction solar cells, but it would also have major impacts on global electronics production. Increasing the implementation of multijunction cell technology would likely positively impact other aspects of their price, including recycling, as right now it is extremely expensive and lacks the necessary infrastructure. By bringing attention to multijunction cells through advancements in manufacturing, it would likely lead to an influx of recycling plants and techniques that would make the process easier. Next, the focus should be on improving the materials used to make multijunction cells, and seeing if alternatives are found to the expensive ones used currently. Although there is a large focus on 3 junction multijunction solar cells with Germanium Indium Phosphide, Germanium Indium Arsenide, Germanium, it is entirely possible that there are more efficient or cheaper materials that could be utilized in this process. While all of these steps are being taken, it's necessary to have well constructed, heavily funded research to guide along the process, as that will be the deciding factor to whether multijunction cells will become cost efficient enough for implementation.

Conclusion

III-V multijunction solar cells offer a high efficiency alternative to monocrystalline silicon solar cells whenever efficiency and space management are the primary concerns. There are many ways to reduce the cost of multijunction cells, but at the rate both cells are improving, it is likely that outside of specific use cases, monocrystalline silicon solar cells will be more useful due to their lower price compared to efficiency. However, it is important to consider multijunction cells as important in terms of research surrounding clean energy and electronics production because of their impressive efficiencies compared to other next generation solar technologies.

Methods

When searching for relevant literature about solar energy, there were many factors considered like the date published, research formatting, and author reliability. Papers were found using google scholar including keywords such as “solar power”, “multijunction cell”, and “manufacturing solar cells”. Data was extracted from articles after carefully reading them, organizing by subject matter, cross checking key data points with other scientific literature, and only then allowing usage in the paper. For the purpose of accuracy and true data, papers from sources that were not sufficiently supported by educational or scientific institutions were ignored.

Works Cited

- IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.
- Mozumder, Mohammad Sayem, et al. "Recent developments in multifunctional coatings for solar panel applications: A review." *Solar Energy Materials and Solar Cells* 189 (2019): 75-102.
- Masafumi Yamaguchi, Frank Dimroth, John F. Geisz, Nicholas J. Ekins-Daukes; Multi-junction solar cells paving the way for super high-efficiency. *J. Appl. Phys.* 28 June 2021; 129 (24): 240901. <https://doi.org/10.1063/5.0048653>
- King, R. R., et al. "40% efficient metamorphic GaInP/GaInAs/Ge multijunction solar cells." *Applied physics letters* 90.18 (2007).
- Battersby, Stephen. "The solar cell of the future." *Proceedings of the National Academy of Sciences* 116.1 (2019): 7-10.
- D.J. Friedman, Progress and challenges for next-generation high-efficiency multijunction solar cells, *Current Opinion in Solid State and Materials Science*, Volume 14, Issue 6, 2010, Pages 131-138, ISSN 1359-0286, <https://doi.org/10.1016/j.cossms.2010.07.001>.
- Simon P. Philipps, Frank Dimroth, Andreas W. Bett, Chapter I-4-B - High-Efficiency III–V Multijunction Solar Cells, Editor(s): Soteris A. Kalogirou, McEvoy's Handbook of Photovoltaics (Third Edition), Academic Press, 2018, Pages 439-472, ISBN 9780128099216, <https://doi.org/10.1016/B978-0-12-809921-6.00012-4>.
- Horowitz, Kelsey A. W., Timothy Remo, Brittany Smith, and Aaron Ptak. 2018. *Techno-Economic Analysis and Cost Reduction Roadmap for III-V Solar Cells*. Golden, CO: National Renewable Energy Laboratory. NREL/TP-6A20-72103. <https://www.nrel.gov/docs/fy19osti/72103.pdf>.
- Philipps, Simon P., and Andreas W. Bett. "III–V multi-junction solar cells." (2014).
- Yamaguchi, M., et al. "Super-high-efficiency multi-junction solar cells." *Progress in photovoltaics: Research and applications* 13.2 (2005): 125-132.
- Sugaya, Takeyoshi, et al. "Investigation of InGaP/(In) AlGaAs/GaAs triple-junction top cells for smart stacked multijunction solar cells grown using molecular beam epitaxy." *Japanese Journal of Applied Physics* 54.8S1 (2015): 08KE02.
- Arimochi, Masayuki, et al. "III–V compound semiconductor multi-junction solar cells fabricated by room-Temperature wafer-bonding technique." *Japanese Journal of Applied Physics* 54.5 (2015): 056601.
- Baiju A and Yarema M (2022), Status and challenges of multi-junction solar cell technology. *Front. Energy Res.* 10:971918. doi: 10.3389/fenrg.2022.971918
- Woodhouse, Michael A., Smith, Brittany, Ramdas, Ashwin, and Margolis, Robert M. *Crystalline Silicon Photovoltaic Module Manufacturing Costs and Sustainable Pricing: 1H 2018*

- Benchmark and Cost Reduction Road Map. United States: N. p., 2019. Web.
doi:10.2172/1495719.
- Miccoli, Ilio, et al. "Enabling MOCVD production on next generation 150 mm Indium Phosphide wafer size." *Journal of Crystal Growth* (2024): 127793.
- Ian Mathews, Donagh O'Mahony, Brian Corbett, and Alan P. Morrison, "Theoretical performance of multi-junction solar cells combining III-V and Si materials," *Opt. Express* 20, A754-A764 (2012)
- Connolly, J.P., Mencaraglia, D., Renard, C. and Bouchier, D. (2014), Designing III–V multijunction solar cells on silicon. *Prog. Photovolt: Res. Appl.*, 22: 810-820.
<https://doi.org/10.1002/pip.2463>
- Ehrler, Bruno, et al. "Photovoltaics reaching for the Shockley–Queisser limit." (2020): 3029-3033.
- “The Point of No Return: How Close Is the World to Irreversible Climate Change?” Scientists for Global Responsibility,
www.sgr.org.uk/resources/point-no-return-how-close-world-irreversible-climate-change#:~:text=The%20global%20average%20temperature%20rise%20is%20predicted%20to%20climb%20permanently,very%20soon%20after%20%5B6%5D. Accessed 7 Sept. 2024.

Sound Mind, Sound Memory: The Music-Stress Connection. A Mini Review Exploring the Comparative Effect of Stress and Music on Learning, Memory, and Brain Waves

By Sienna Phillips

Abstract

Stress is any stimulus (intrinsic or extrinsic) that causes a biological response, typically categorized into acute (short-term) and chronic stress (long-term). Chronic stress leads to damage in the hippocampus as well as selected neuronal populations, leading to impairment of learning and memory function. When the brain is stressed, beta waves are predominantly present, a sign of a mentally active and focused brain. In contrast, a relaxed brain is filled with alpha waves. Studies show that one method to achieving this relaxed state is through music, primarily the genre of classical music. Listening and being trained in music results in a more relaxed brain, dominated by alpha waves. Music also has been shown to have positive effects on memory and positive yet contrasting effects on learning. In this review, I focus on the differing effects of stress and music on learning, memory, and brain waves.

Keywords: Stress, Music, Learning, Memory, Brain Waves

Introduction

We have all experienced moments of stress or anxiety. In theory, this is beneficial, allowing the body to defend itself from possible threats or stressors. Stress functions to preserve the homeostasis of the body (Goldstein and Kopin, 2007), giving a jolt of energy in order to survive stressful situations, triggering “fight or flight” (Chu et al., 2024). However, in today's world, people are increasingly stressed over things that have little effect on them. Think about it, being stressed over the grade on an upcoming test or melting sea ice thousands of miles away will not necessarily help the problem.

Although the term is often used loosely, stress is any stimulus (intrinsic or extrinsic) that causes a biological response. Additionally, the response to stress is not always the same due to differences in type, timing, duration, and severity of the stimulus (Yaribeygi et al., 2017).

The main difference in stress focused on in this paper will be duration of the stimulus, broken into acute (short-term) and chronic stress (long-term). Acute stress is a situation in which the stress response disappears quickly after the perceived danger. In the case of chronic stress, the stress remains for weeks or months at a time (Chronic Stress and How to Manage It, 2020), and the stress response does not instantly go away.

More often than not, acute stress is beneficial in memory retention and consolidation with improved memory of stressful events (Goldfarb, 2023). However, chronic stress does not yield the same effect. Unlike acute stress, chronic stress can be detrimental to the brain's ability to learn and remember events. This is due in part to brain atrophy and other damaging factors (Bremner, 2006).

However, music may be a potential tool to mitigate stress. Studies suggest that music such as Mozart or Indian classical music can actually help relieve stress (Verrusio et al., 2015)(Kumar et al., 2017). This in turn would lead to improved memory and learning within the brain (Verrusio et al., 2015)(Chan et al., 1998). Based on this philosophy, treatments such as music therapies and medicine for mental health problems exist, opening the door to affordable and effective stress treatments (Witte et al., 2022).

Stress and the HPA axis

Hypothalamic-Pituitary Axis

The hypothalamic-pituitary (HPA) axis is the organ system that plays a key role in the body's stress management. It consists of the hypothalamus, pituitary gland, and adrenal glands (Hypothalamic-Pituitary-Adrenal (HPA) Axis, 2024). Once a stressor is detected, the information is sent to the hypothalamus through the autonomic nervous system, releasing corticotropin-releasing hormone (CRH). CRH then signals the anterior pituitary to release adrenocorticotropic hormone (ACTH). Next, the ACTH signals the body's adrenal glands to produce cortisol (Hypothalamic-Pituitary-Adrenal (HPA) Axis, 2024). Cortisol is a glucocorticoid that keeps the body alert by increasing the amount of glucose in the blood as well as helping the body repair tissues and slow down nonessential functions during stress (Sandhya et al., 2023).

Fight or Flight

As the HPA axis responds to stress, the hypothalamus uses the autonomic nervous system to alert the rest of the body to the stressor. Within the autonomic nervous system is the sympathetic nervous system which puts the body on high alert so it can defend itself from any potential threat. Signals are then sent through the sympathetic nervous system to the adrenal glands, pumping epinephrine into the blood (LeWine, 2024). Epinephrine, commonly known as adrenaline, releases glucose and fats into the body, providing energy to activate "fight or flight" mode. This involves an increased heart rate, blood pressure, and breathing (Mariotti, 2015). All these things are the body's initial response to acute stress.

Acute versus Chronic Stress

Stress can be broadly classified into two categories: acute stress and chronic stress. In the case of acute stress, epinephrine and cortisol return back to their normal levels once the potential or real threat is gone, making sure the body does not continue to respond to the stressor (LeWine, 2024).

However, in the case of chronic stress, glucocorticoid-dependent negative feedback in the hippocampus is not able to control the stress response and glucocorticoid receptor resistance develops (Mariotti, 2015). This means that glucocorticoid continues to be produced and cortisol does not go back to normal levels so the body remains on high alert (Mariotti, 2015). When stress persists over a long period of time, chronic stress has occurred.

Studies have shown that continuous stress is detrimental to learning and memory. In this mini-review paper I delve deeper into the effect of acute and chronic stress on learning and memory.

Effect of Stress on Memory

Memory is one of the most important functions of the Central Nervous System. It can be categorized into sensory, short-term, and long-term memory (Yaribeygi et al., 2017). The frontal and parietal lobes of the brain play the most important roles in short-term memory whereas long-term memory is dependent on the hippocampus and amygdala regions of the brain (Yaribeygi et al., 2017). The hippocampus region serves as the center for memory to be converted from short-term to long-term (Yaribeygi et al., 2017). Additionally, the hippocampus region is highly susceptible to stress (Yaribeygi et al., 2017).

Stress affects memory in a variety of different ways.

In the case of acute stress, it is found that memory can be improved during stressful situations (Goldfarb, 2023). The purpose of improving memory during acute stress is that when a challenge or dangerous situation is met, the brain better remembers how to adequately respond in the future. Acute stress affects memory in this way as glucocorticoids have been found to enhance memories (Goldfarb, 2023). Studies in rats have shown that acute stress results in improved yet slightly fragmented memories (Diamond et al., 2006). Additionally, human studies involving oral cortisol doses have exhibited enhanced declarative memories (memories involving specific facts or events)(Goldfarb, 2023).

Adversely, chronic and extreme stress has the opposite effect on memory. Drawn-out and intense stress is known to lead to long-term structural changes in the hippocampus (Yaribeygi et al., 2017). The hippocampus is located within the temporal lobe of the brain and is subdivided into regions such as the dentate gyrus, hippocampus, CA1 (Cornu Ammonis 1), CA3 (Cornu Ammonis 3), subiculum, presubiculum, parasubiculum, and entorhinal cortex (Anderson, 2007). Over the past fifty years, the hippocampus has been increasingly studied due to its wide variety of functions, such as its engagement in the processes of learning and memory (Anderson, 2007). Within the hippocampus, memories are converted from short to long-term (Anderson, 2007). The hippocampus' vulnerability to stress is particularly alarming due to its vital role in learning and memory. Chronic stress can lead to significant structural changes in the hippocampus, including brain atrophy, which involves the loss of brain tissue, neurons, and neural connections. These changes can be observed on a larger scale, such as a decrease in the volume of the hippocampus (Bremner, 2006) and adrenal gland (Sarahian et al., 2014). For example, one study found that individuals with chronic occupational stress exhibited a reduced volume of gray matter in the brain (Blix et al., 2013). Studies done in vervet monkeys have shown that one-year exposure to cortisol in the hippocampus led to neuronal shrinkage and dendritic atrophy in the CA3 and CA2 areas of the hippocampus (Sapolsky et al., 1990). An indicator of neuron degeneration is dendritic branch length reduction (Woolley et al., 1990). The dendrites of neurons are responsible for collecting incoming information from other neurons. Studies have demonstrated

that corticosterone-induced stress in rodent models leads to a reduction in dendritic length within the CA3 region of the hippocampus. This dendritic atrophy is a hallmark of neuronal degeneration (Woolley et al., 1990).

The morphological changes induced by chronic stress in the hippocampus significantly impair the brain's ability to acquire and retain memories. Hippocampal atrophy specifically limits the brain's capacity to consolidate short-term memories into long-term storage. A recent study revealed that rats exposed to chronic stress exhibited deficits in acquiring memory for an eight-arm radial maze task (Luine et al., 1994). In a similar study where rats were exposed to psychosocial stress by placing them near a cat for 5 weeks resulted in impaired learning and memory (Park et al., 2001). It is also reported that high concentrations of stress hormones can lead to memory disorders such as Alzheimer's (Lupien and Lepage, 2001). Thus, chronic and extreme stress has been shown to impair memory acquisition, retention, and consolidation.

Effect of Stress on Learning

Each new experience that a person goes through allows the brain to learn. Learning is a complicated process within the hippocampus, primarily generated by connections between neurons through their synapses (Cunnington, 2019). These connections are constantly changing all the way through adulthood, although most learning occurs during childhood (Cunnington, 2019). Learning occurs in two main ways. The first is when new connections are formed between neurons, called synaptogenesis (Qi et al., 2022). Another way learning can take place is long-term potentiation or LTP, where neuronal connections become stronger (Purves et al., 2001).


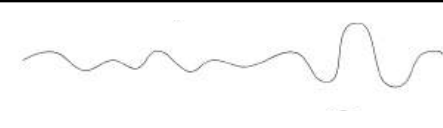

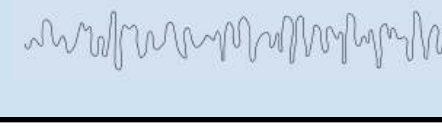
We have established that high levels of stress can damage memory, but does learning experience similar changes as a result of stress? Unlike with memory, the answer is not so straightforward. The Yerkes-Dodson Law states that as psychological or mental arousal increases, so does the performance of a difficult task, up until a certain point. After this point, the opposite is true, with performance decreasing as arousal continues to increase (Diamond et al., 2006). A difficult task can be thought of as challenging learning situations, including divided attention, multitasking, and working memory tasks (Diamond et al., 2006). Performance in the context of this law is nearly synonymous with learning. Arousal is the level of activity within the autonomic nervous system, the system responsible for regulating the body's response to stress. More activity is directly correlated to higher levels of stress. The idea originally laid out by the Yerkes-Dodson Law was later supported by Broadhurst's rat model in which rats were trained to escape from drowning in water (Broadhurst, 1957). When rats were put under intermediate levels of stress in a difficult version of this task they learned rapidly, resulting in the best performance (Broadhurst, 1957). Thus, this study as well as the Yerkes-Dodson Law supports the idea that the best performance of difficult tasks is found at intermediate levels of stress and after this point, impaired learning and performance are observed. If stress levels continue to increase after peak performance is reached, the stress will become extreme or chronic, resulting in less of an ability to learn (Diamond et al., 2006).

However, this is only a portion of the Yerkes-Dodson Law surrounding stress's relationship with learning performance. There is also a relationship between stress levels and the performance of simple tasks (Diamond et al., 2006). For example, in an easy version of Broadhurst's drowning rat model, when rats were put under extreme stress they learned very rapidly (Broadhurst, 1957). For simple tasks, as stress increases, so does learning, leveling off when stress is slightly above intermediate stress levels (Diamond et al., 2006). However, it is important to keep in mind that similarly to stress's effect on memory, high amounts of stress impair complex learning (Diamond et al., 2006).

Earlier in this review we established that extreme and chronic stress hurts the brain's hippocampus through brain atrophy (Yaribeygi et al., 2017). This in turn impairs the brain's memory functioning. A similar phenomenon happens when extreme/chronic stress plagues the brain while learning is taking place. Once again, due to extreme stress, brain atrophy occurs. This involves a reduction in the volume of the hippocampus (Bremner, 2006) and gray matter within the brain (Blix et al., 2013), as well as a decrease in the number of neurons (Sapolsky et al., 1990) and dendritic branch length (Woolley et al., 1990). Because learning occurs by connections between neurons, when neurons are lost as a result of chronic stress, the brain's capacity to learn is impaired. Additionally, the hippocampus is the primary brain structure responsible for learning, so reducing the size of the hippocampus is associated with a reduction in the brain's learning capacity. Ultimately, much like memory, chronic and extreme stress have detrimental effects on complex learning.

Since chronic stress affects the hippocampus and has negative effects on learning and memory, I wanted to explore the effect of stress on specific brain waves.

Brain Waves

Type of Brain Wave	Frequency	Visual Representation	Mental State
Delta	<4 Hz		Deep sleep (Hammond, 2011)
Theta	4-8 Hz		Daydreaming Mental inefficiency (Hammond, 2011)
Alpha	8-12 Hz		Relaxed (Hammond, 2011)
Beta	13-30 Hz		Focused Mentally active (Hammond, 2011)


Gamma	>30 Hz		Intensely focussed (28)
-------	--------	--	-------------------------

Figure 1. Types of neuronal oscillations, their frequencies, visual representations, and associated mental state. (Brain waves have been hand drawn)

The brain is always processing information through the firing of billions of neurons. When this neuronal activity between synapses is rhythmic it produces electrical signals known as neural oscillations, commonly referred to as brain waves. The brain emits brain waves all the time, including during the processes of memory and learning (Başar, 2013). However, not all brain waves look the same. Neural oscillations are mainly categorized by their frequency, with delta waves (<4 Hz) having the lowest frequency, then theta (4-8 Hz), alpha (8-12 Hz), beta (13-30 Hz), and finally gamma (>30 Hz) (Hammond, 2011). Although each category of brain wave is associated with a different mental state, all brain waves are always present within the brain, just to different degrees (Hammond, 2011). In my mini-review, I wanted to explore the effect of stress on alpha and beta waves. In order to fully understand how stress affects the brain it is crucial that we understand the brain's neuronal oscillations.

Alpha waves are slower and larger than beta waves, predominant when the brain is in a relaxed state (Hammond, 2011). When alpha waves are pulsing through the brain, people often have their eyes closed, are relaxed, disengaged, or waiting for something to occur (Hammond, 2011). Compared to alpha waves, beta waves are small and fast (Hammond, 2011). Therefore, these brain waves can be found when the brain is mentally and intellectually active or concentrated on an external factor (Hammond, 2011). However, lower frequency beta waves (closer to 13 Hz) are associated with a more relaxed attentiveness (Hammond, 2011).

Firstly, in situations of stress when the brain is on high alert, beta waves are predominant (Hammond, 2011). One study exposed their subjects to a virtual reality stimulus, with the participants' alpha/beta ratios being recorded before and after they were exposed to the stressor. The study found that when stressed, their subjects had a decreased alpha/beta ratio, indicating increased anxiety and fear (Wen and Aris, 2020). This study supports the theory that beta waves increase as a result of stress while alpha waves decrease.

Secondly, unlike beta waves, alpha waves are found to decrease as a result of stress (Hammond, 2011). The feeling of anxiousness and tenseness, while associated with high beta waves, can also be a result of an alpha wave deficit in the front of the brain, the area in charge of emotional control (Hammond, 2011). One study even found that decreased serum cortisol levels were correlated negatively with alpha wave presence in the brain (Kamei et al., 2000). As cortisol levels increase during the stress response, this finding indicates that alpha waves decrease.

Together, an increase in beta wave and decrease in alpha wave activity in the brain are associated with a state of stress. As most high-frequency beta waves are present when the brain is active and concentrated on something in the external environment, it is logical that when the

brain is on high alert (i.e. stressed) beta waves are predominant. Additionally, alpha waves represent a relaxed, disengaged state. So, when the brain is stressed, alpha waves are less present.

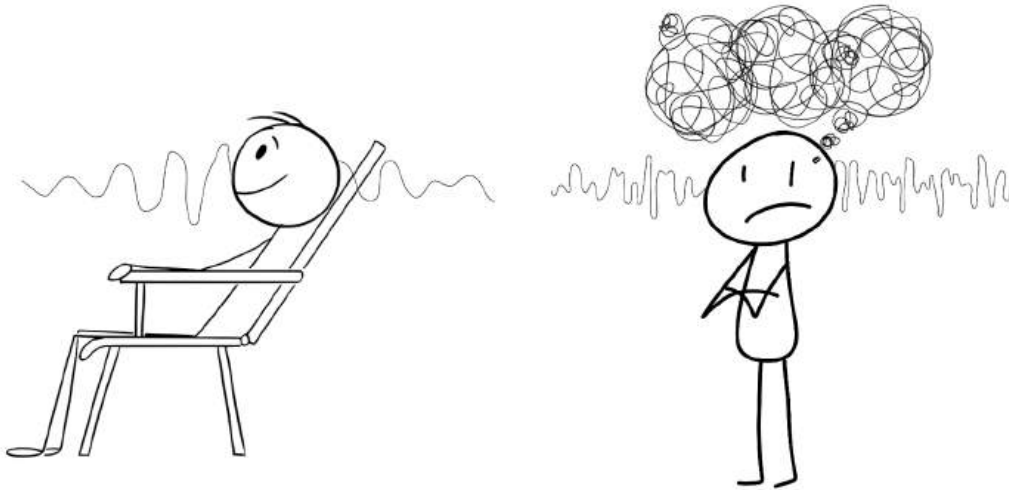


Figure 2. Impact of stress on brain waves. Alpha waves are depicted on the left with an unstressed mind. Beta waves (right) are present when the brain is stressed. (stick figures from Canva)

The effect of certain types of music on the brain is very relaxing. For example, when we go to a spa they play zen music. And many students, including myself, listen to study music while doing school work. Spotify has even created its own genre of music titled “study music”, with almost 4.5 million people following the *Deep Focus* playlist. While using this music, I have felt more relaxed, sparking my interest to delve deeper into the effect of music on stress, learning, memory, and brain waves.

Music

Effect of Music on Stress, Learning, and Memory

Music has always been a tool to evoke emotions, be that love, anger, or a sense of calmness. Well, this idea actually has significant scientific backing, forming a foundation for music therapies and interventions. Music therapy is carried out by music therapists who use specific qualities of music to help address the mental health of their patients (Witte et al., 2022).

More specifically, music is associated with decreased stress, indicated by decreased cortisol levels, heart rate (Witte et al., 2022), and an increase in alpha wave power (Kumar et al., 2017). One study found that participants who listened to relaxing music before being exposed to a stressor resulted in the autonomic nervous system being less affected as well as a lesser psychological stress response (Thoma et al., 2013). Additionally, music therapies have been found to be effective in reducing the stress of patients (Witte et al., 2022).

As a result of a decrease in stress response, memory and learning are seen to improve through listening to and playing music. For example, one study found that listening to Mozart's music resulted in brain wave frequencies associated with memory, cognition, and an open mind to problem-solving (Verrusio et al., 2015). Additionally, another study indicates that adults who receive music training before the age of 12 have better memory of spoken words than those who have not received this training (Chan et al., 1998). It is hypothesized that this improvement is seen due to musicians' better developed cognitive function in the left temporal area, where verbal memory is processed (Chan et al., 1998).

Much like memory, learning is seen to improve when music is listened to. In one study, elementary school students either listened or did not listen to Mozart while reading. Those in the music group were found to have better reading comprehension than those in the silent group, indicating that music is associated with improved learning (Song, 2023). In a related study, students improved their test scores between two similar exams when they received musical training in between (Song, 2023). Because music helps mitigate stress, it is logical to conclude that music can also improve memory and learning functioning within the brain, as supported by these studies.

Although learning and memory are valuable, to fully understand music's effect on stress it is also important to look at how music impacts the brain's neuronal oscillations.

Effect of Music on Alpha and Beta Brain Waves

Music can have the opposite effect on brain waves compared to stress. Most studies on music's effect on the brain utilize Sonata in D Major for Two Pianos K448 by Mozart. The theory that listening to Mozart's music is beneficial for the brain is referred to as the Mozart effect and is due to many factors such as the repetitive nature of the melody and organization of the song. One study found that listening to Mozart resulted in an increase in alpha wave band and median frequency index of background alpha rhythm activity in the brain. Both of these alpha wave changes are associated with memory, cognition, and a growth mindset (Verrusio et al., 2015). A similar study found that when participants listened to Indian classical music, they exhibited increased alpha wave power in the temporal lobe (Kumar et al., 2017). Interestingly, another study found that when professional musicians listened to music they experienced alpha wave activity in the back of the brain (Mikutta et al., 2014). This increase in alpha wave power within the brains of music listeners indicates a relaxed state, harshly contrasting the decrease in alpha waves during stress.

Summary

Throughout this review, it has been established that chronic and extreme stress negatively impact the brain. This is seen through brain atrophy (Bremner, 2006) as well as neuron and dendritic degeneration (Woolley et al., 1990) within the hippocampus. The hippocampus is fundamental to the brain's ability to learn, acquire, and retain memories, thus, undermining both

of these functions (Yaribeygi et al., 2017). Additionally, the brain's stressful state has been associated with an increased beta wave power and presence (Hammond, 2011).

Conversely to stress, music has a positive impact on the brain. Alpha waves, present in a relaxed and disengaged brain are seen to increase when people listen to classical music (Verrusio et al., 2015)(Kumar et al., 2017). Music has also been found to relieve stress (Witte et al., 2022), indicating that as stress decreases, memory and learning functioning improve as a result. Additionally, the brain's capacity for memory and learning are somewhat associated as new learned information must be stored and remembered.

There is significant evidence that music helps relieve stress, resulting in improved learning and memory. Based on this idea, music therapies exist that can help patients suffering from extreme or chronic stress improve their situation. These therapies are noninvasive and promising new approaches to the way stress is treated. Additionally, there are more affordable options such as listening to calming music in the comfort of your own home, be that Mozart, Indian classical, or simply a soothing study playlist.

Although these findings are promising, more research must be done in the areas of music's effect on learning and memory in order to be more certain of music's positive effects on the brain.

Works Cited

- Andersen, Per. *The Hippocampus Book*. Oxford ; New York, Oxford University Press, 2007.
- Başar, Erol. “Brain Oscillations in Neuropsychiatric Disease.” *Dialogues in Clinical Neuroscience*, vol. 15, no. 3, 1 Sept. 2013, pp. 291–300, www.ncbi.nlm.nih.gov/pmc/articles/PMC3811101/.
- Blix, Eva, et al. “Long-Term Occupational Stress Is Associated with Regional Reductions in Brain Tissue Volumes.” *PLoS ONE*, vol. 8, no. 6, 11 June 2013, p. e64065, <https://doi.org/10.1371/journal.pone.0064065>. Accessed 1 Apr. 2020.
- Bremner, J. Douglas. “Stress and Brain Atrophy.” *CNS & Neurological Disorders Drug Targets*, vol. 5, no. 5, 1 Oct. 2006, pp. 503–512, www.ncbi.nlm.nih.gov/pmc/articles/PMC3269810/.
- Broadhurst, P. L. “Emotionality and the Yerkes-Dodson Law.” *Journal of Experimental Psychology*, vol. 54, no. 5, 1957, pp. 345–352, <https://doi.org/10.1037/h0049114>.
- Chan, Agnes S., et al. “Music Training Improves Verbal Memory.” *Nature*, vol. 396, no. 6707, Nov. 1998, pp. 128–128, www.nature.com/articles/24075, <https://doi.org/10.1038/24075>.
- “Chronic Stress and How to Manage It | Pfizer.” pfizer, www.pfizer.com/news/articles/chronic_stress_and_how_to_manage_it#:~:text=Stress%20that%20lasts%20for%20weeks.
- Chu, Brianna, et al. “Physiology, Stress Reaction.” PubMed, StatPearls Publishing, 2024, www.ncbi.nlm.nih.gov/books/NBK541120/.
- Cleveland Clinic. “HPA Axis: The Stress Response System.” Cleveland Clinic, 12 Apr. 2024, my.clevelandclinic.org/health/body/hypothalamic-pituitary-adrenal-hpa-axis.
- Cunnington, Ross. “Neuroplasticity: How the Brain Changes with Learning.” IBE — Science of Learning Portal, 18 Sept. 2019, solportal.ibe-unesco.org/articles/neuroplasticity-how-the-brain-changes-with-learning/.
- Diamond, David M., et al. “The Temporal Dynamics Model of Emotional Memory Processing: A Synthesis on the Neurobiological Basis of Stress-Induced Amnesia, Flashbulb and Traumatic Memories, and the Yerkes-Dodson Law.” *Neural Plasticity*, vol. 2007, 2007, pp. 1–33, <https://doi.org/10.1155/2007/60803>.
- Goldfarb, Elizabeth V. “Enhancing Memory with Stress: Progress, Challenges, and Opportunities.” *Brain and Cognition*, vol. 133, July 2019, pp. 94–105, <https://doi.org/10.1016/j.bandc.2018.11.009>.
- Goldstein, David S., and Irwin J. Kopin. “Evolution of Concepts of Stress.” *Stress*, vol. 10, no. 2, Jan. 2007, pp. 109–120, <https://doi.org/10.1080/10253890701288935>. Accessed 10 Dec. 2019.
- Hammond, D. *Journal of Neurotherapy: Investigations in Neuromodulation, Neurofeedback and*

- Applied Neuroscience What Is Neurofeedback: An Update. 2011, [isnr.org/wp-content/uploads/2019/07/16553-Article-Text-64896-1-10-20160826.pdf](https://doi.org/10.1080/10874208.2011.623090), <https://doi.org/10.1080/10874208.2011.623090>.
- Kamei, Tsutomu, et al. "Decrease in Serum Cortisol during Yoga Exercise Is Correlated with Alpha Wave Activation." *Perceptual and Motor Skills*, vol. 90, no. 3, June 2000, pp. 1027–1032, <https://doi.org/10.2466/pms.2000.90.3.1027>.
- Kumar, Satish, et al. "A Study on Effect of Indian Classical Music on Brain Activity Using EEG Signals." *Journal of Medical Science and Clinical Research*, vol. 05, no. 05, 12 May 2017, pp. 21702–21706, <https://doi.org/10.18535/jmscr/v5i5.76>. Accessed 7 Apr. 2019.
- Howard E. LeWine. "Understanding the Stress Response ." Harvard Health, Harvard Health Publishing, 3 Apr. 2024, www.health.harvard.edu/staying-healthy/understanding-the-stress-response.
- Luine, Victoria, et al. "Repeated Stress Causes Reversible Impairments of Spatial Memory Performance." *Brain Research*, vol. 639, no. 1, Mar. 1994, pp. 167–170, [https://doi.org/10.1016/0006-8993\(94\)91778-7](https://doi.org/10.1016/0006-8993(94)91778-7).
- Lupien, Sonia and Lepage, Martin. "Stress, Memory, and the Hippocampus: Can't Live with It, Can't Live without It." *Behavioural Brain Research*, vol. 127, no. 1-2, Dec. 2001, pp. 137–158, [https://doi.org/10.1016/s0166-4328\(01\)00361-8](https://doi.org/10.1016/s0166-4328(01)00361-8).
- Mariotti, Agnese. "The Effects of Chronic Stress on Health: New Insights into the Molecular Mechanisms of Brain–Body Communication." *Future Science OA*, vol. 1, no. 3, 2015, www.ncbi.nlm.nih.gov/pmc/articles/PMC5137920/, <https://doi.org/10.4155/fso.15.21>.
- Mikutta, C.A., et al. "Professional Musicians Listen Differently to Music." *Neuroscience*, vol. 268, May 2014, pp. 102–111, <https://doi.org/10.1016/j.neuroscience.2014.03.007>. Accessed 29 Nov. 2020.
- Park, Collin R., et al. "Chronic Psychosocial Stress Impairs Learning and Memory and Increases Sensitivity to Yohimbine in Adult Rats." *Biological Psychiatry*, vol. 50, no. 12, Dec. 2001, pp. 994–1004, [https://doi.org/10.1016/s0006-3223\(01\)01255-0](https://doi.org/10.1016/s0006-3223(01)01255-0). Accessed 23 Mar. 2020.
- Purves, Dale, et al. "Long-Term Synaptic Potentiation." *Neuroscience*. 2nd Edition, 2001, www.ncbi.nlm.nih.gov/books/NBK10878/.
- Qi, Cai, et al. "Molecular Mechanisms of Synaptogenesis." *Frontiers in Synaptic Neuroscience*, vol. 14, 13 Sept. 2022, www.ncbi.nlm.nih.gov/pmc/articles/PMC9513053/pdf/fnsyn-14-939793.pdf, <https://doi.org/10.3389/fnsyn.2022.939793>. Accessed 21 Mar. 2023.
- Mayo Clinic. "Chronic Stress Puts Your Health at Risk." Mayo Clinic, 19 Mar. 2019, www.mayoclinic.org/healthy-lifestyle/stress-management/in-depth/stress/art-20046037#:~:text=Cortisol%2C%20the%20primary%20stress%20hormone.
- Sapolsky, Robert, et al. "Hippocampal Damage Associated with Prolonged

- Glucocorticoid Exposure in Primates.” 1990.
- Sarahian, Nahid, et al. “Effect of Memantine Administration within the Nucleus Accumbens on Changes in Weight and Volume of the Brain and Adrenal Gland during Chronic Stress in Female Mice.” *Pathobiology Research*, vol. 17, no. 2, 10 June 2014, pp. 71–82, mjms.modares.ac.ir/browse.php?a_code=A-30-12276-1&slc_lang=en&sid=30.
- Song, Yifei. “Music’s Influence on Children’s Cognitive Development.” *SHS Web of Conferences*, vol. 180, 2023, p. 02011, www.shs-conferences.org/articles/shsconf/pdf/2023/29/shsconf_icepcc2023_02011.pdf, <https://doi.org/10.1051/shsconf/202318002011>.
- Thoma, Myriam V., et al. “The Effect of Music on the Human Stress Response.” *PLoS ONE*, vol. 8, no. 8, 5 Aug. 2013, p. e70156, www.ncbi.nlm.nih.gov/pmc/articles/PMC3734071/, <https://doi.org/10.1371/journal.pone.0070156>.
- Verrusio, Walter, et al. “The Mozart Effect: A Quantitative EEG Study.” *Consciousness and Cognition*, vol. 35, Sept. 2015, pp. 150–155, www.sciencedirect.com/science/article/pii/S1053810015001130, <https://doi.org/10.1016/j.concog.2015.05.005>.
- Wen, T, and Aris, S. “Electroencephalogram (EEG) stress analysis on alpha/beta ratio and theta/beta ratio.” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 17, no. 1, Jan 2019, pp. 175-182. <https://doi.org/10.11591/ijeecs.v17.i1.pp175-182>
- Witte, Martina, et al. “Music Therapy for Stress Reduction: A Systematic Review and Meta-Analysis.” *Health Psychology Review*, vol. 16, no. 1, 27 Nov. 2020, pp. 1–26, www.tandfonline.com/doi/full/10.1080/17437199.2020.1846580, <https://doi.org/10.1080/17437199.2020.1846580>.
- Woolley, Catherine S., et al. “Exposure to Excess Glucocorticoids Alters Dendritic Morphology of Adult Hippocampal Pyramidal Neurons.” *Brain Research*, vol. 531, no. 1-2, Oct. 1990, pp. 225–231, [https://doi.org/10.1016/0006-8993\(90\)90778-a](https://doi.org/10.1016/0006-8993(90)90778-a). Accessed 11 Feb. 2021.
- Yaribeygi, Habib, et al. “The Impact of Stress on Body Function: A Review.” *EXCLI Journal*, vol. 16, no. 1, 21 July 2017, pp. 1057–1072, www.ncbi.nlm.nih.gov/pmc/articles/PMC5579396/, <https://doi.org/10.17179/excli2017-480>.

CAR T-cell Therapy By Sayan Majumder

Abstract

CAR T-cell therapy is currently FDA approved against various blood cancers, but there are some limitations that prevent it from being mainstream cancer treatment. This therapy has over a 60 year history since the time it was first discovered.

CAR T-cells are used in a four step process. The biological structure of Chimeric Antigen receptors(CARs) has great detail and four big components stand out. CAR T-cell therapy fights certain types of the following blood cancers: lymphoma, leukemia, and multiple myeloma. Examples of the lymphomas CAR T-cells can fight are B-cell non-Hodgkin lymphoma, follicular lymphoma, and mantle cell lymphoma. The types of leukemia targeted are acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL). Even amongst the many successes seen with this therapy, there are still heavy side effects associated with it, as well as challenges it has not been able to fight.

Introduction

Chimeric Antigen Receptor(CAR) T-Cell Therapy is a novel cancer treatment that has recently excited scientists. As an immunotherapy, its primary goal is to harness a cancer patient's own immune system to combat the disease. A patient's T cells are a type of white blood cell produced from stem cells in the bone marrow that help fight viruses, infections, and cancers. These T-cells are put through a four step process to target blood cancers, including leukemia, lymphoma, multiple myeloma, and acute lymphoblastic leukemia. First, the patient's blood is collected and put through an apheresis machine to isolate T-cells. Next, a gene coding for the "Chimeric Antigen Receptor", a lab-made receptor that targets surface molecules on tumor cells, is integrated with the T-cells. Once the T-Cells that express "CAR" are produced, they are expanded to produce large quantities of CAR T-cells. Finally, CAR T-cells are infused into the patient. The specialized T-cells target other cells that express proteins corresponding to the epitope of the CAR.

CAR T-Cell therapies have also been used to combat autoimmune diseases, such as lupus, myositis, and sclerosis. These diseases cause conditions that make the immune system mistakenly attack healthy body tissues. The process is similar to that of treating blood cancers, except that the Chimeric Antigen Receptors enable T-cells to target B-cells responsible for the autoimmune response.

Historical Premise of CAR T-Cells

T-cells were discovered in 1961 by the scientist Jacques Miller, but T-cell engineering did not begin until 1992 (*CAR T Cells: Timeline of Progress*, 1960). In the following years, scientists worked to enhance this technology and in 2002, their efforts led to the first CAR T-cells that successfully targeted a prostate antigen. In 2009, the details of CD19 antigen production, used against relapsed leukemia post chemotherapy, were made public. Four years later, a clinical trial

using CD19 CAR T-cells demonstrated success against Acute Lymphoblastic Leukemia(ALL). In 2014, the US FDA called CAR T-cell therapy a “breakthrough” therapy, and in 2015, the first “armored CARs” were produced and tested against ovarian cancer. By 2017, CAR T-cell therapy had advanced to the point where CRISPR could be used to insert CARs in strategic genetic locations on T-cells. It was also the year when CAR T-cells were finally FDA approved with the CD19 antigen for treating Acute Lymphoblastic Leukemia(ALL) in children and young adults (CAR T Cells: Timeline of Progress, 1960).

Structure of CAR T-Cells

At its core, CAR T-cells are composed of CARs, lab-made receptors attached to T-cells. These receptors are based on the biology of T-cells themselves. CARs enable T-cells to fight cancer cells through a cell signaling cascade that instructs the cell to undergo apoptosis or induce killing through perforin and granzyme release. Each Chimeric Antigen Receptor(CAR) consists of four components. The following figure illustrates a CAR T-cell, highlighting the structure of CAR.

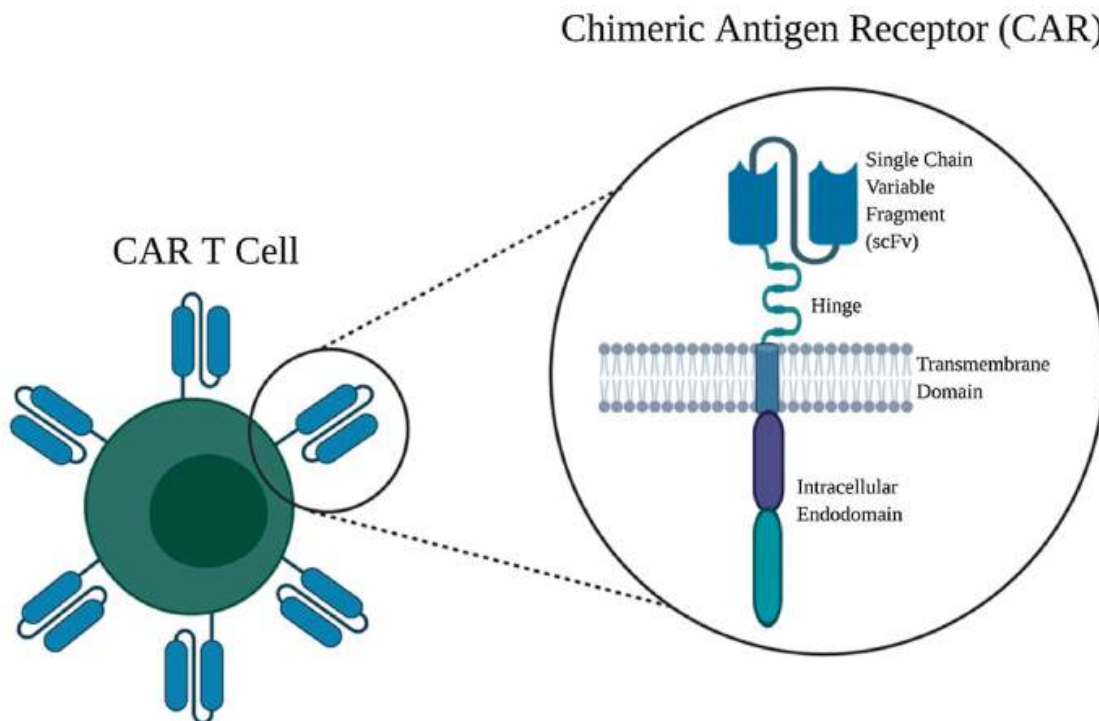


Figure 1 - This figure shows the overall CAR T-cell and a detailed picture of the Chimeric Antigen Receptor(CAR). It shows the four main parts, which are the scFv, hinge region, transmembrane domain, and intracellular endodomain.

Image Credits: Springer Link

The first section is the antigen binding domain, which enables the CARs to recognize and bind to certain antigens. In most cases, this domain is a Single Chain Variable Fragment, though it can differ in some instances (Wang et al., 2022b). The second section is the hinge domain. Also known as the spacer region, this is a flexible spot that allows the overall receptor to change

shape and accommodate different types of antigens. This flexibility helps the CAR bind to proteins on cancer cells. The third part of CAR is the transmembrane domain, which anchors the CAR to the T-cell. The fourth and final region is the intracellular endodomain. This is the internally functional portion of the CAR that allows for signaling to the T-cell nucleus for downstream gene transcription.

How CAR T-cells fights blood cancers

Lymphoma is the overarching term for a cancer that originates in the lymphatic system. The two main types of lymphoma are Hodgkin and non-Hodgkin. Certain FDA approved CAR T-Cell therapies targeting the antigen CD19 have been used to treat B-Cell non-Hodgkin lymphoma (NHL) in adult patients who have not responded to previous treatments or have relapsed (Denlinger et al., 2022). The antigen CD19 is a promising target because it is a glycoprotein seen on B-cells even as they differentiate into plasma or memory cells (Althwaiqeb & Bordoni, 2023; *CD19 CD19 Molecule [Homo Sapiens (Human)] - Gene - NCBI*, n.d.). Other therapies using the same target antigen have been effective against follicular lymphoma and mantle cell lymphoma, two other types of NHL (Denlinger et al., 2022). Leukemia includes a group of blood cancers caused by genetic mutations in the bone marrow, leading to an exponential growth rate of blood cells. The same CAR T-Cell therapies that were used with the target antigen CD19 can fight B-cell acute lymphoblastic leukemia for adolescents and adults (Saleh et al., 2023).

Multiple Myeloma, also called Kahler's disease, is a cancer of B- cells in the blood. Two CAR T-Cell therapies using the antigen BCMA have targeted multiple myeloma in adults who relapsed or did not respond to other treatments (Mishra et al., 2023). MM leads to buildup of myeloma cells in the bone marrow. These cancerous cells displace and deprive healthy blood cells of space and nutrients (Mayo Clinic). Myeloma cells also produce an abnormal antibody called M proteins instead of normal immunoglobulin antibodies produced by healthy plasma cells (*Multiple Myeloma*, n.d.). Unfortunately, M proteins do not fight diseases; instead, they cause damage to various parts of the body, like the kidneys and bones. If M proteins are in the blood, there is a high chance of hyperviscosity syndrome, which thickens the blood and puts stress on the heart to pump blood throughout the body (Multiple Myeloma, 2024). M proteins also disrupt the production of healthy white blood cells, red blood cells, and platelets. The most significant symptoms of multiple myeloma include bone pain, poor immune function, anemia, low white blood cell count, low red blood cell count, and low platelet count (MMRF, 2024).

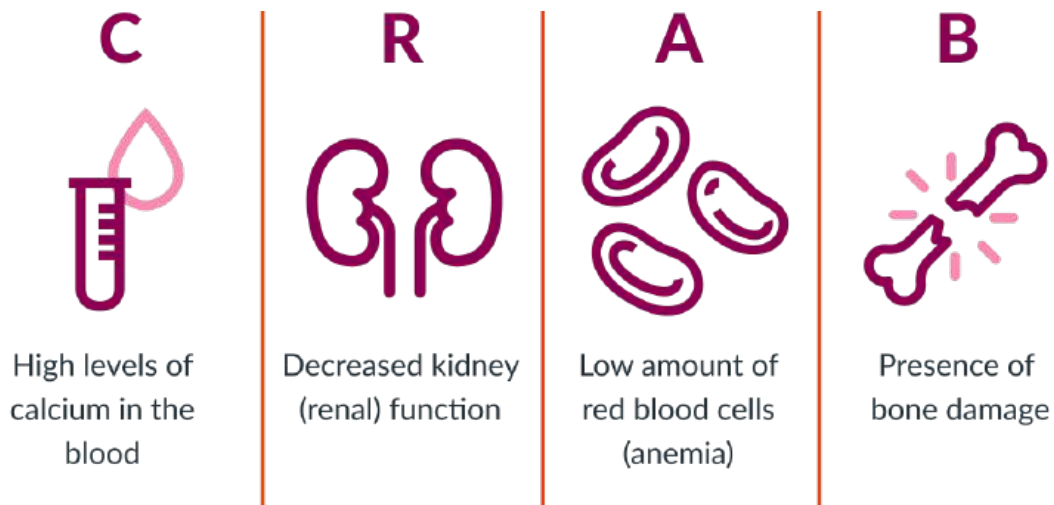


Figure 2: Shown above are four key symptoms of multiple myeloma. They are illustrated with the acronym “CRAB”. The “C” is calcium, the “R” is renal, the “A” is anemia, and the “B” is bone.
Image Credits: Multiple Myeloma Research Foundation

Multiple Myeloma is currently the second most common type of blood cancer in the United States (*Multiple Myeloma* | GSK US, n.d.). Potential target antigens for CAR T-Cell Therapy in treating MM include BCMA, CD19, CD38, and CD70 (Mishra et al., 2023).

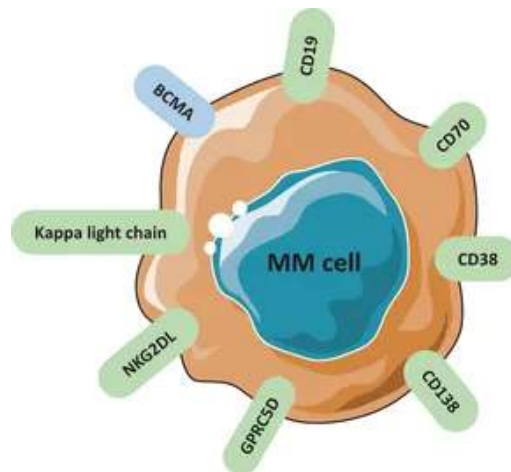


Figure 3: A Multiple Myeloma cell is shown with eight surface antigens attached. These have been studied in the clinic through various trials. BCMA, the antigen in the color blue, is the most commonly targeted surface antigen on MM cells through CAR T-cell therapy.

Image Credits: Frontiers

Another CAR T-cell therapy to fight Multiple Myeloma is called ciltacabtagene autoleucel or Carvykti for short. It is used in adults who have refractory or relapsed multiple myeloma and have already tried many treatments. Carvykti found success in certain clinical trials where the BCMA antigen on myeloma cells was targeted. Based on a clinical trial called CARTITUDE-1 with 97 patients, 98% had reduced amounts of cancer, though for some this was partially. Moreover, 78% of patients had absolutely no signs of cancer in bone marrow and blood

(FDA Approves Carvykti CAR T-Cell Therapy for Multiple Myeloma, 2022). This is called a stringent complete response and it is the highest classification of treatment response for multiple myeloma patients (Myeloma Central, n.d.).

The key components used in the structure of T-cells to fight against MM cells are described in the following figure (Wang et al., 2022).

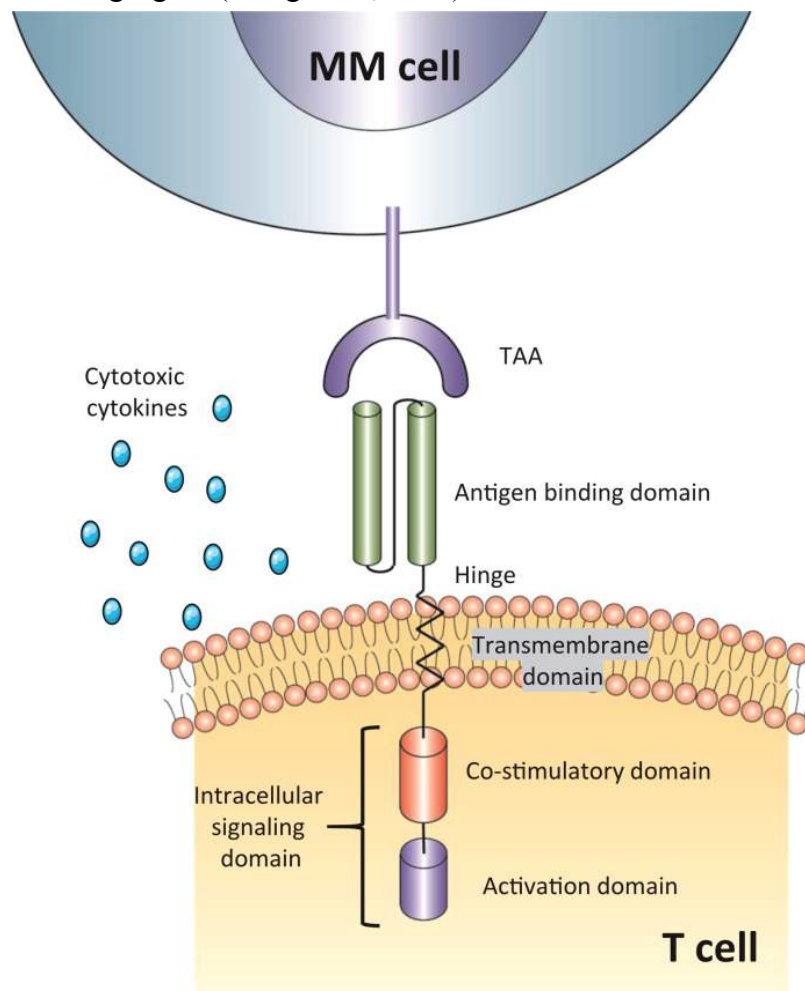


Figure 4 - The figure shows the primary components that make up a T-cell. Ultimately, the antigen binding domain attaches to the Tumor Associated Antigen (TAA) of a Multiple Myeloma (MM) cell. Cytotoxic cytokines are also pictured.

Image Credits: *Frontiers*

CAR T-cell therapy has also been utilized in Acute Myeloid Leukemia (AML). AML happens when blood stem cells, which would normally mature into blood cells over time, are stunted in their developmental period. As a result, the body's bone marrow produces an excess of dysfunctional blood cells, including red blood cells, white blood cells, and platelets. Many AML cases involve rapid growth of premature white blood cells called myeloblasts. This type of blood cancer can spread to various parts of the body, including the central nervous system, skin, and gingiva. Additionally, the overproduction of myeloblasts limits space for healthy blood cells in the blood and bone marrow, which results in a high risk of infection, anemia, and an inability to

form blood clots when bleeding. Risk factors for AML include old age, smoking, radiation exposure, and a family history of blood disorders (National Cancer Institute).

Though clinical trials are still underway and limited data has been released, CAR T-cells are known to respond to relapsed AML. In the case of a certain 41 year old man, CD33-targeting CAR T-cells led to a reduction of AML blasts in the bone marrow. Also, a clinical trial at the City of Hope Medical Center with 24 patients enrolled as of February 2019 revealed that one patient was able to reach a temporary stage free of leukemia. The patient later received an infusion of CAR T-cells 3 months later and had a 77% reduction in AML blasts (Mardiana & Gill, 2020).

Limits to CAR T-Cell Therapy

Despite the successes of many clinical trials, CAR T-cell therapy still has some limitations. One is that solid tumors are hard to fight with this therapy. CAR T-cells struggle to combat solid tumors and even if they can, penetrating the tumor microenvironment(TME) remains a formidable task. Solid tumors present a challenge due to the limited availability of tumor-specific antigens (TSAs) and the presence of heterogeneous tumor-associated antigens (TAAs). This means CAR T-cell targets will only be present on some of the tumor cells, allowing other cells to survive and proliferate. For this reason, CAR T-cells are most often used to treat blood cancers, where there are no solid tumors present.

Also, CAR T-cell therapy is very expensive. This therapy can cost between \$500,000 and \$1,000,000. This is due to the costs of biopsies, an apheresis, lab production of CAR T-cells, hospital stays, imaging studies, and more. Only some insurance companies cover the costs of CAR T-cell therapy, and even these companies are not able to cover everything (*CAR T-Cell Therapy: Managing Costs and Finding Financial Assistance*, n.d.).

The most significant issue with CAR T-cell therapy is its associated toxicities. One toxicity is called cytokine release syndrome(CRS), caused by cytokines released by CAR T-cells transfused into the patient's body. This results in hyper stimulation of the patient's natural immune system and a mass cytokine release (*Cytokine Release Syndrome (CRS)*, 2024). CRS can lead to respiratory symptoms or in severe cases, organ failure. Another toxicity is ICANS, or Immune Effector cell-associated neurotoxicity syndrome. This can cause a range of problems such as encephalopathy, seizures, confusion, and aphasia (Rees, 2022).

Conclusion

CAR T-cell therapy is a recent cancer immunotherapy that has excited both scientists and the medical community as a whole. It has an intricate design that has been re-engineered over the years to improve efficiency. This therapy has found success against various blood cancers. However, there are toxicities and other health hazards associated with it that make it hard to use on patients. Thus, there is still a lot of work to be done to ensure this therapy can be safer, more efficient, and more affordable.

Works Cited

- “Acute Myeloid Leukemia Treatment.” Cancer.gov, 6 Mar. 2024,
www.cancer.gov/types/leukemia/patient/adult-aml-treatment-pdq#:~:text=Adult%20acute%20myeloid%20leukemia%20.
- Althwaiqeb, S. A., & Bordoni, B. (2023, May 29). Histology, B cell lymphocyte. StatPearls - NCBI Bookshelf.
<https://www.ncbi.nlm.nih.gov/books/NBK560905/#:~:text=The%20first%20signal%20is%20C%20cross,ability%20of%20memory%20cell%20production>.
- Andrea, Alain E., et al. “Advances in CAR-T Cell Genetic Engineering Strategies to Overcome Hurdles in Solid Tumors Treatment.” *Frontiers in Immunology*, vol. 13, Feb. 2022,
<https://doi.org/10.3389/fimmu.2022.830292>.
- “Autoimmune Diseases.” National Institute of Environmental Health Sciences,
www.niehs.nih.gov/health/topics/conditions/autoimmune.
- Brudno, Jennifer, and James Kochenderfer. Figure 5. 30 June 2016. *Blood*. Accessed 11 Aug. 2024.
- Byju’s. Hinge Region in Antibodies Is Important Because. 4 July 2022,
byjus.com/question-answer/hinge-region-in-antibodies-is-important-because-it-provides-flexibility-to-the-antibody-molecules-helps-change.
- “CAR T cells: Engineering immune cells to treat cancer.” Cancer.gov, 10 Mar. 2022,
www.cancer.gov/about-cancer/treatment/research/car-t-cells.
- “CAR T Cells: Timeline of Progress.” Memorial Sloan Kettering Cancer Center, 1960,
www.mskcc.org/timeline/car-t-timeline-progress.
- “CAR T-cell Therapy and Its Side Effects.” American Cancer Society, 2022,
www.cancer.org/cancer/managing-cancer/treatment-types/immunotherapy/car-t-cell1.html. Accessed 11 Aug. 2024.
- “CAR T-cell Therapy.” Cancer Research UK,
www.cancerresearchuk.org/about-cancer/treatment/immunotherapy/types/CAR-T-cell-therapy#:~:text=Collecting%20your%20T%20cells&text=One%20tube%20removes%20the%20blood,tube%20in%20your%20other%20arm.
- “CAR T Cell Therapy.” Penn Medicine,
www.pennmedicine.org/cancer/navigating-cancer-care/treatment-types/immunotherapy/what-is-car-t-therapy. Accessed 11 Aug. 2024.
- CD19 CD19 molecule [Homo sapiens (human)] - Gene - NCBI. (n.d.).
<https://www.ncbi.nlm.nih.gov/gene/930#:~:text=CD19%20is%20continuously%20and%20stably,antibody%20targeting%20treatment%20of%20leukemia>.
- Chen, Chen, et al. Figure 3. 2022. *Frontiers*. Accessed 11 Aug 2024.
- Chen, Chen, et al. Figure 4. 22 Dec. 2022. *Frontiers*. Accessed 11 Aug 2024.
- Cytokine release Syndrome (CRS). (2024, May 1). Cleveland Clinic.
<https://my.clevelandclinic.org/health/diseases/22700-cytokine-release-syndrome>

Denlinger, N., Bond, D., & Jaglowski, S. (2022). CAR T-cell therapy for B-cell lymphoma. *Current Problems in Cancer*, 46(1), 100826. <https://doi.org/10.1016/j.currprobcancer.2021.100826>

FDA approves Carvykti CAR T-Cell therapy for multiple myeloma. (2022, March 30). Cancer.gov. <https://www.cancer.gov/news-events/cancer-currents-blog/2022/fda-carvykti-multiple-myeloma#:~:text=The%20cancer%20can%20form%20tumors,called%20a%20stringent%20complete%20response.>

Figure 2. 2024. Multiple Myeloma Research Foundation, Multiple Myeloma Research Foundation Inc. . Accessed 12 Aug. 2024.

Hofmann, Susanne, et al. “Chimeric Antigen Receptor (CAR) T Cell Therapy in Acute Myeloid Leukemia (AML).” *Journal of Clinical Medicine*, vol. 8, no. 2, Feb. 2019, p. 200. <https://doi.org/10.3390/jcm8020200>.

“Leukemia - Symptoms and Causes - Mayo Clinic.” Mayo Clinic, 21 Sept. 2022, [mayoclinic.org/diseases-conditions/leukemia/symptoms-causes/syc-20374373#:~:text=In%20general%2C%20leukemia%20is%20thought,signs%20and%20symptoms%20of%20leukemia.](https://www.mayoclinic.org/diseases-conditions/leukemia/symptoms-causes/syc-20374373#:~:text=In%20general%2C%20leukemia%20is%20thought,signs%20and%20symptoms%20of%20leukemia.)

“Leukemia—Patient Version.” Cancer.gov, www.cancer.gov/types/leukemia.

Mardiana, S., & Gill, S. (2020). CAR T cells for acute myeloid leukemia: State of the art and future directions. *Frontiers in Oncology*, 10. <https://doi.org/10.3389/fonc.2020.00697>

“---.” MD Anderson Cancer Center, www.mdanderson.org/cancer-types/multiple-myeloma.html#:~:text=If%20a%20plasma%20cell%20becomes,do%20not%20fight%20against%20infection.

Mishra, A. K., Gupta, A., Dagar, G., Das, D., Chakraborty, A., Haque, S., Prasad, C. P., Singh, A., Bhat, A. A., Macha, M. A., Benali, M., Saini, K. S., Previs, R. A., Saini, D., Saha, D., Dutta, P., Bhatnagar, A. R., Darwal, M., Shankar, A., & Singh, M. (2023). CAR-T-Cell therapy in multiple myeloma: B-Cell maturation antigen (BCMA) and beyond. *Vaccines*, 11(11), 1721. <https://doi.org/10.3390/vaccines11111721>

MMRF. “What Is Multiple Myeloma? Symptoms, Causes, and Prognosis.” MMRF, 29 May 2024, [themmrf.org/multiple-myeloma](https://www.themmrf.org/multiple-myeloma).

Multiple myeloma | GSK US. (n.d.). <https://us.gsk.com/en-us/company/oncology/multiple-myeloma/#:~:text=Multiple%20myeloma%20is%20the%20second,to%20be%20diagnosed%20this%20year.>

Multiple myeloma. (2024, May 1). Cleveland Clinic. <https://my.clevelandclinic.org/health/diseases/6178-multiple-myeloma>

“Multiple Myeloma.” MD Anderson Cancer Center, www.mdanderson.org/cancer-types/multiple-myeloma.html.

Myeloma central. (n.d.). Myeloma Central. <https://www.myelomacentral.com/livingwithmm/multiple-myeloma-symptoms-and-diagn>

- osis/glossary-multiple-myeloma-definitions#:~:text=Stringent%20complete%20response%20(sCR),measured%20by%20immunohistochemistry%20or%20immunofluorescence.
- “Multiple Myeloma - Symptoms and Causes - Mayo Clinic.” Mayo Clinic, 30 July 2024, www.mayoclinic.org/diseases-conditions/multiple-myeloma/symptoms-causes/syc-20353378#:~:text=Multiple%20myeloma%20is%20a%20cancer,Antibodies%20find%20and%20attack%20germs.
- “NCI Dictionary of Cancer Terms.” Cancer.gov, www.cancer.gov/publications/dictionaries/cancer-terms/def/car-t-cell-therapy.
- Rees, J. H. (2022). Management of Immune Effector Cell-Associated Neurotoxicity Syndrome (ICANS). In Springer eBooks (pp. 141–145). https://doi.org/10.1007/978-3-030-94353-0_27
- Saleh, K., Pasquier, F., Bigenwald, C., De Botton, S., Ribrag, V., & Castilla-Llorente, C. (2023). CAR T-Cells for the treatment of B-Cell acute lymphoblastic leukemia. *Journal of Clinical Medicine*, 12(21), 6883. <https://doi.org/10.3390/jcm12216883>
- Sterner, Robert C., and Rosalie M. Sterner. “CAR-T Cell Therapy: Current Limitations and Potential Strategies.” *Blood Cancer Journal*, vol. 11, no. 4, Apr. 2021, <https://doi.org/10.1038/s41408-021-00459-7>.
- “Toxicities of Chimeric Antigen Receptor T Cells: Recognition and Management.” *Blood*, ashpublications.org/blood/article/127/26/3321/35340/Toxicities-of-chimeric-antigen-receptor-T-cells. Accessed 11 Aug. 2024.
- Verma, Amitesh, and Sarwish Rafiq. “Chimeric Antigen Receptor (CAR) T Cell Therapy for Glioblastoma.” *Cancer Treatment and Research*, Jan. 2022, pp. 161–84. https://doi.org/10.1007/978-3-030-96376-7_5.
- Verma, Amitesh, and Sarwish Rafiq. Chimeric Antigen Receptor Structure. 13 May 2022. Springer Link. Accessed 4 Aug. 2024.
- Wang, Zehua, et al. “Chimeric Antigen Receptor T-cell Therapy for Multiple Myeloma.” *Frontiers in Immunology*, vol. 13, Dec. 2022, <https://doi.org/10.3389/fimmu.2022.1050522>.
- Zhang, Cheng, et al. “Engineering CAR-T Cells.” *Biomarker Research*, vol. 5, no. 1, June 2017, <https://doi.org/10.1186/s40364-017-0102-y>.

What recent advances have been made in C-H methylation for late-stage functionalization?

By Angela Wang

Abstract

The late-stage functionalization of drug molecules plays an invaluable part of the chemical life sciences industry. The addition of methyl groups to drug scaffolds has been shown to greatly enhance the efficacy of the drug for given targets. Indeed, this effect is so substantial, it has been denoted as the ‘magic methyl effect’. Organic chemists have therefore made specific attempts to exact methylation strategies to common drug scaffolds, with perhaps the most valuable being C-H methylation strategies due to the ubiquity of C-H bonds in drug candidates. Therefore, in this review, the most recent advances in methodologies to perform C-H methylations will be discussed. This will involve assessing photochemical, electrochemical and transition metal-catalyzed reactions that now enable selective C-H methylations of drug-like scaffolds.

Introduction

Late-stage functionalization (LSF) can be defined as the performing of a chemical transformation on molecules featuring high structural complexity, requiring an excellent level of chemoselectivity and, preferably, presenting an important site-selectivity, without the need for pre-introduction of additional specific functional groups (Guillemard et al. 545; Börgel et al. 1887). It plays an important part in modern pharmaceutical research where there is a necessity for preparing a large number of variations of compounds to explore structure-activity relationships (SARs). However, LSF has only emerged as a ubiquitous area of research with the onset of a plethora of new reaction methodologies exhibiting the combination of chemoselectivity and a broad functional group tolerance. The ubiquity of C-H bonds in organic molecules makes them by far the most pervasive group to target for LSF strategies and so the chemoselective and regioselective activation of specific C-H bonds remains the challenge for useful LSF methods (Guillemard et al. 545). In addition to this, the most synthetically useful LSF protocols enable the installation of small groups, such as fluoro, chloro, methyl and hydroxy groups with high chemoselectivity onto unprotected lead compounds. This is because the introduction of large functional groups, such as (hetero)aryl rings, often results in too large an alteration of a lead compound’s structure and thus can result in a reduction in its biological efficacy (Guillemard et al. 545). Small substituents, however, can modulate the properties of a lead compound in a beneficial manner with respect to either binding affinity and/or drug metabolism and pharmacokinetic properties.

The installation of methyl groups is of high importance in LSF campaigns due to the remarkable ability for methyl groups to drastically impact the solubility, hydrophilicity and conformation of a lead compound, thus profoundly influencing its biological activity, pharmacokinetic profile, and physical properties (Aynedinova et. Al. 5563). The particular

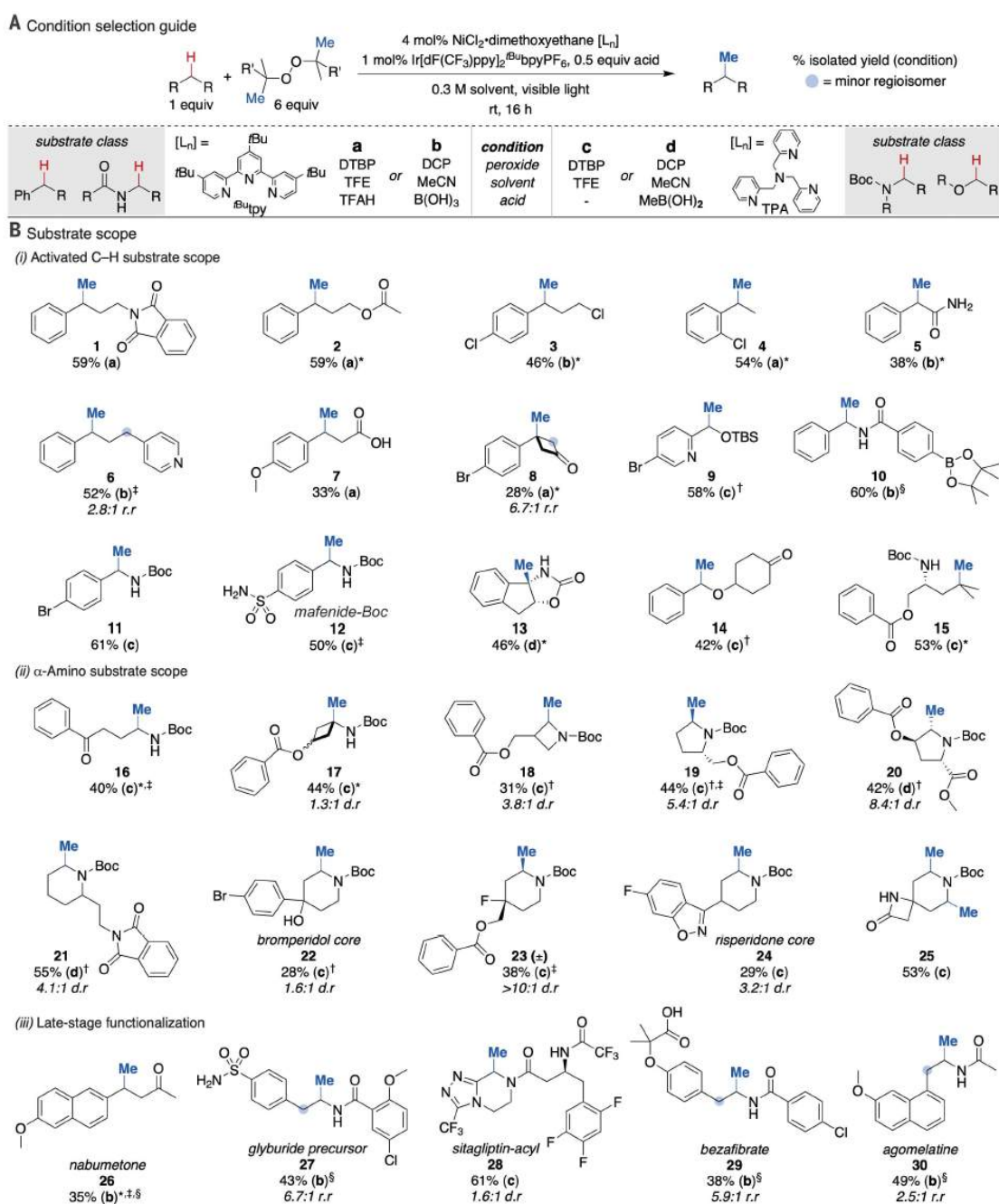
impact of methyl groups to this end has led to the coining of the term ‘magic methyl effect’ within medicinal chemistry circles over the last 10 years. Indeed, a literature survey of >2000 cases revealed that 8% of methyl installations led to a >10-fold potency boost, and >100-fold activity increases in 0.4% of cases (Leung et al. 4500). As a result, methyl analogues of drug candidates are emerging as highly valuable targets in drug discovery campaigns, with 40% of the best-selling drugs in 2019 containing at least one methyl group (Njardson 1). Introducing methyl groups, however, is non-trivial, with most new methyl derivatives of lead compounds requiring lengthy *de novo* syntheses using more conventional chemistries (Börgel et. al. 1887). The development of methods to install methyl groups in place of C-H bonds therefore has a high demand, with the ability to also further drive innovation within the medicinal chemistry field. In recent years, there have been some important developments towards this goal, particularly in the fields of photochemistry, electrochemistry and transition metal catalysis, which now allow chemoselective C-H methylation under relatively mild conditions. This review shall cover some of the key contributions to this area since 2017. However, the methylation of heteroarenes by Minisci-type reactivity, shall not be a focus of this review due to the plethora of methods developed for this subsection of C-H methylations (Zhang et. al. 3626).

Photochemical C-H methylation

Photochemistry has been through a renaissance in the last 15 years with the advancement in LED technologies and the uptake of photocatalysis. Indeed, the development of visible-light-mediated photocatalytic reactions has enabled the engagement of otherwise unreactive molecules selectively under mild conditions using just visible light and a photocatalyst (Belloti et. al. 4352). The highly reactive intermediates that can be generated photochemically have broadly expanded the range of methylation strategies possible in synthesis. Furthermore, with specific developments in hydrogen-atom transfer (HAT) catalysis, remote and seemingly unreactive (by classical standards) C(sp³)-H bonds can be readily activated and substituted with a variety of functional groups (Belloti et. al. 4352). There have therefore been concerted efforts in recent years towards C(sp³)-H methylation strategies using photochemical procedures (Liao et. al. 8203). There have also been a plethora of methods developed towards C(sp²)-H methylations of arenes, however these generally proceed *via* Minisci-type reactivity and thus access specific sites on aromatic heterocycles (Zhang et. al. 3626).

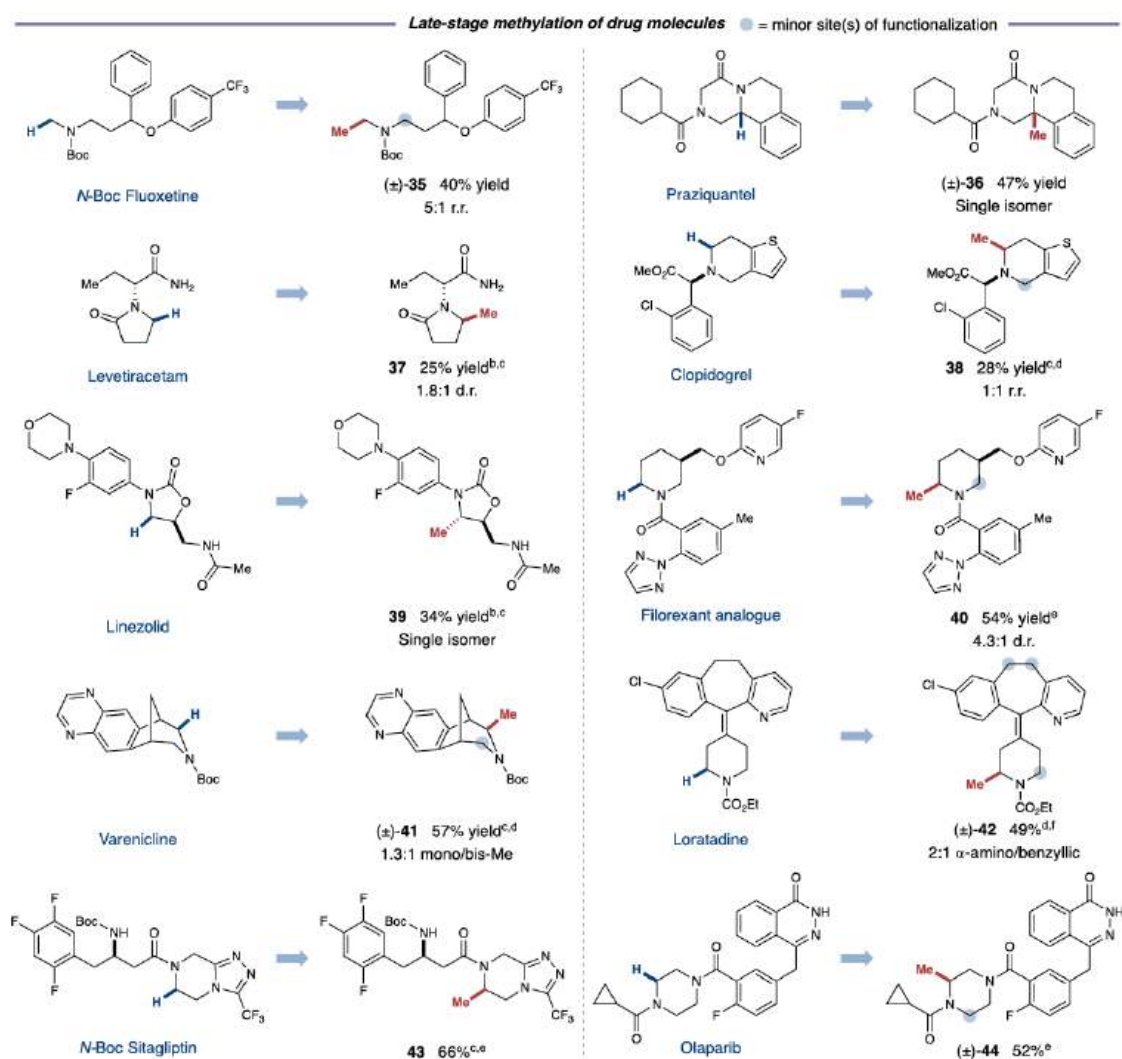
In terms of C(sp³)-H methylation, however, a variety of different photochemical strategies have been employed. In 2021, the Stahl group disclosed a novel photochemical approach of using tertiary alkyl peroxides to perform HAT on activated C(sp³)-H bonds and subsequently methylate the position by the *in situ* generation of methyl radicals, also from the peroxide source (Vasilopoulos et. al. 403). By introducing nickel to this mixture, the methyl radicals produced were cross-coupled effectively with the HAT-generated alkyl radicals of a diverse range of α -amino and/or benzylic substrates, without the need for a multi-step chemical process. The authors also presented an investigation on selectivity challenges observed with substrates bearing both α -amino and benzylic sites and showed that by forming the ammonium

salt of the amine enabled selective activation of just the benzylic position. As a result, a selection of APIs were methylated at either benzylic or α -amino positions, offering a direct opportunity for a late-stage methylation strategy (Scheme 1). In contrast with previous methods, this approach is applicable under mild conditions, which not only ensure substrate compatibility under visible light and nickel catalyst but also improve molecule stability that is beneficial for late stage functionalization of drug molecules.



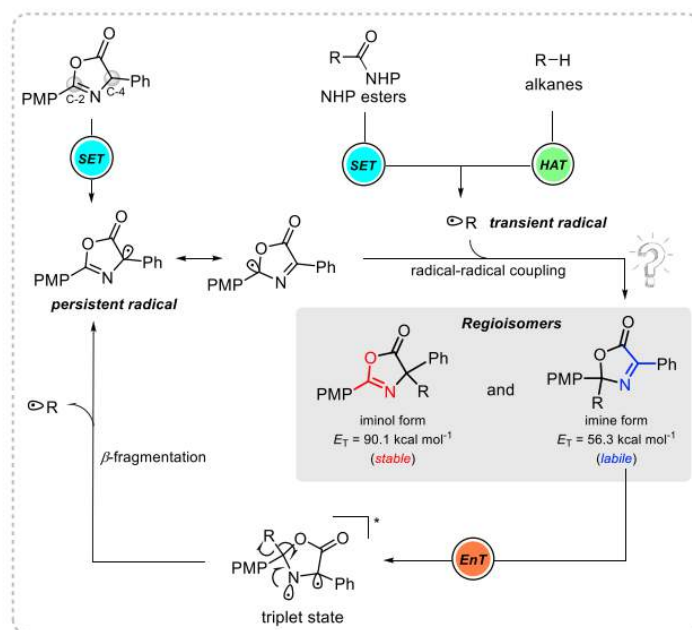
Scheme 1. The use of iridium-nickel dual catalysis and peroxides to perform C-H methylations at benzylic and α -amino positions (Vasilopoulos et. al. 403).

In 2023, Mao and MacMillan disclosed a strategy for α -amino and α -oxy methylation by merging decatungstate HAT catalysis with a nickel co-catalyst (Mao et. al. 2793). In doing so, they were able to activate a range of C(sp³)-H positions on a broad selection of *N*- and *O*-containing aliphatic heterocycles and derivative them to their methylated analogues in generally good yields. Not only this, but they also demonstrated the ability to perform late-stage C(sp³)-H methylations on a wide range of active pharmaceutical ingredients (APIs), with complexity in both scaffold and functional group composition (Scheme 2). In this case, they employed *N*-acetyloxypthalimide as the methyl radical source, which is generated after single electron reduction and subsequent decarboxylation. The decatungstate HAT photocatalyst was able to undergo polarity-matched HAT at the most hydridic C-H bonds in each substrate selectively. The generated methyl radicals and α -amino/ α -oxy radicals were then coupled by nickel, producing the methylated derivatives.

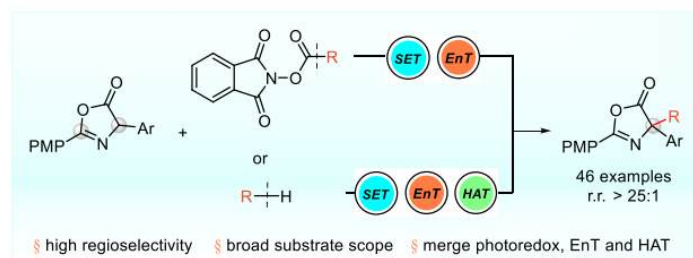


Scheme 2. The use of decatungstate HAT catalysis and a radical source to perform late-stage C-H methylations on the α -amino position of drug compounds (Mao et. al. 2793).

A prevalent motif in many biologically active molecules is the presence of amino acid residues. Also increasingly common are α,α -disubstituted unnatural amino acid fragments, which have received much attention in the last decade for the building of peptidometric libraries. Wu and Lu developed a method to functionalize azlactones at the C-4 position *via* the interplay of single electron transfer (SET), hydrogen atom transfer (HAT) and energy transfer (EnT) steps (Zhu et. al. 4902). In this way, the formal selective alkylation of azlactones was achieved by a series of radical mechanisms. Since the employment of azlactones in radical reactions by direct energy transfer have been limited historically due to the inability to distinguish the C-2 and C-4 positions selectively, Wu, Lu and co-workers employed photocatalysis for the activation of feedstock molecules. This in turn enabled the activation selectively of the C-2-substituted products also bearing a *para*-methoxyphenyl (PMP) substituent *via* an energy transfer mechanism, thus resolving C-2-substituted products to the more stable C-4-substituted products (Scheme 3). With this reaction, they were able to prepare a selection of C-4-alkylated azlactones with a variety of substituents in typically good yields. Amongst these, was a methylated azlactone, produced in 64% yield under the standard reaction conditions, offering the opportunity for the methylation at the alpha position of amino acid residues. The ability to do this as a late-stage protocol remains unreported to date, however.

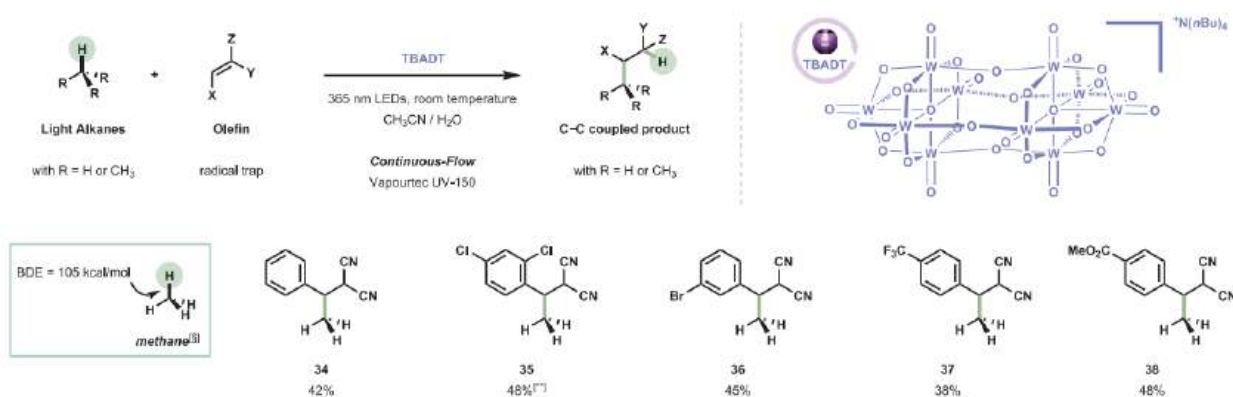


B. This work



Scheme 3. The functionalization of azlactones *via* photocatalytic radical coupling (Zhu et. al. 4902).

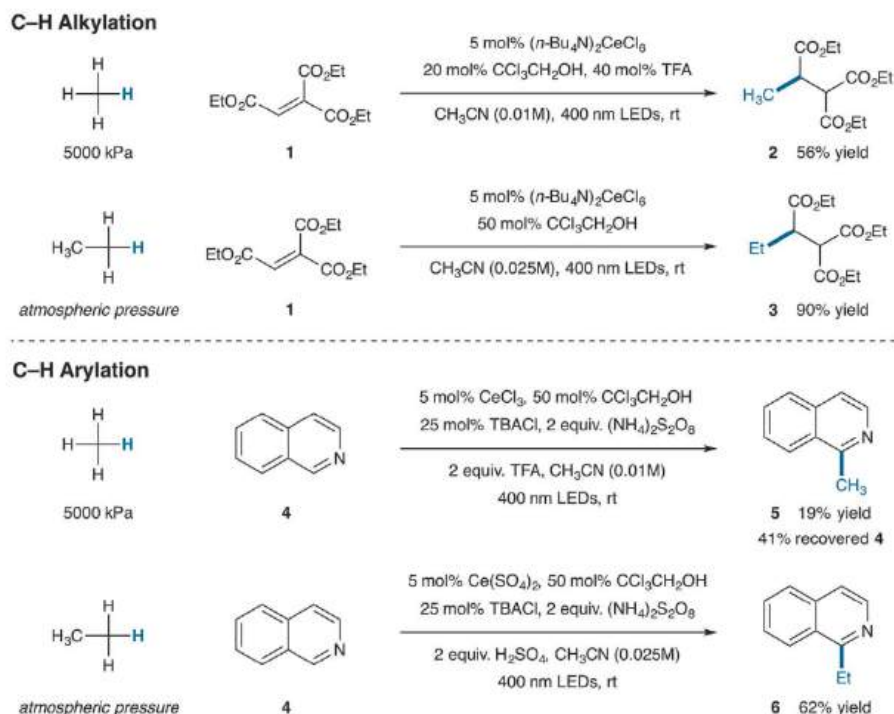
There has been much focus on the various sources of the methyl source for late-stage methylation strategies. In 2020, Noël and co-workers disclosed a strategy to selectively activate gaseous hydrocarbons using decatungstate HAT photocatalysis in flow (Laudadio et. al. 96). By doing so, they were able to target the difficulty in transforming alkanes into high-value chemicals due to its low reactivity with nearby functional groups, thus simplifying synthetic conditions. Although the authors reported the direct decatungstate-mediated HAT activation of alkanes worked optimally for more highly substituted alkanes, impressively they demonstrated that methane could also be activated to generate methyl radicals under the conditions. In the case of each alkanes activation, the alkyl radical produced was trapped with a highly activated and matched olefin acceptor to produce alkylated product. In the case of the methylation reaction, yields ranging from 38-48% were achieved (Scheme 4). Crucially, the flow apparatus enabled the effective irradiation of the reaction mixture whilst under pressure without the additional hazards associated with a batch set-up. This work provides a tangible opportunity for selective methylation strategies from just methane, although late-stage applications are yet to be delivered.



Scheme 4. The functionalization of methane by decatungstate HAT photocatalysis in flow (Laudadio et. al. 96).

In a similar vein, Zuo and co-workers reported in 2018 the development of a photocatalytic C-H amination, alkylation and arylation of methane, ethane and higher alkanes under mild reaction conditions (Hu et. al. 672). Since the intrinsic inertness of C-H bond challenges the catalytic systems as has been discussed before, the authors applied the combination of LMCT and HAT catalysis *via* the use of cerium (III) salts as photocatalysts and simple alcohols as co-catalysts, to generate hard to access light alkyl radicals. These were then trapped with radical acceptors such as activated alkenes, arenes and diazo compounds (Scheme 5). This method was later applied to more complex alkanes, which not only enhance the mixed-phase gas/ solution reactions, but also offers opportunities for the functionalization of feedstock alkanes. Crucially, the use of cheaper metals such as cerium for photocatalysts provides cost advantages to many of the transition metal photocatalysts typically employed for this type of chemistry, such as iridium and ruthenium. Not only this, they demonstrated the application of a photoflow reactor to enable the activation of light alkanes under high pressure

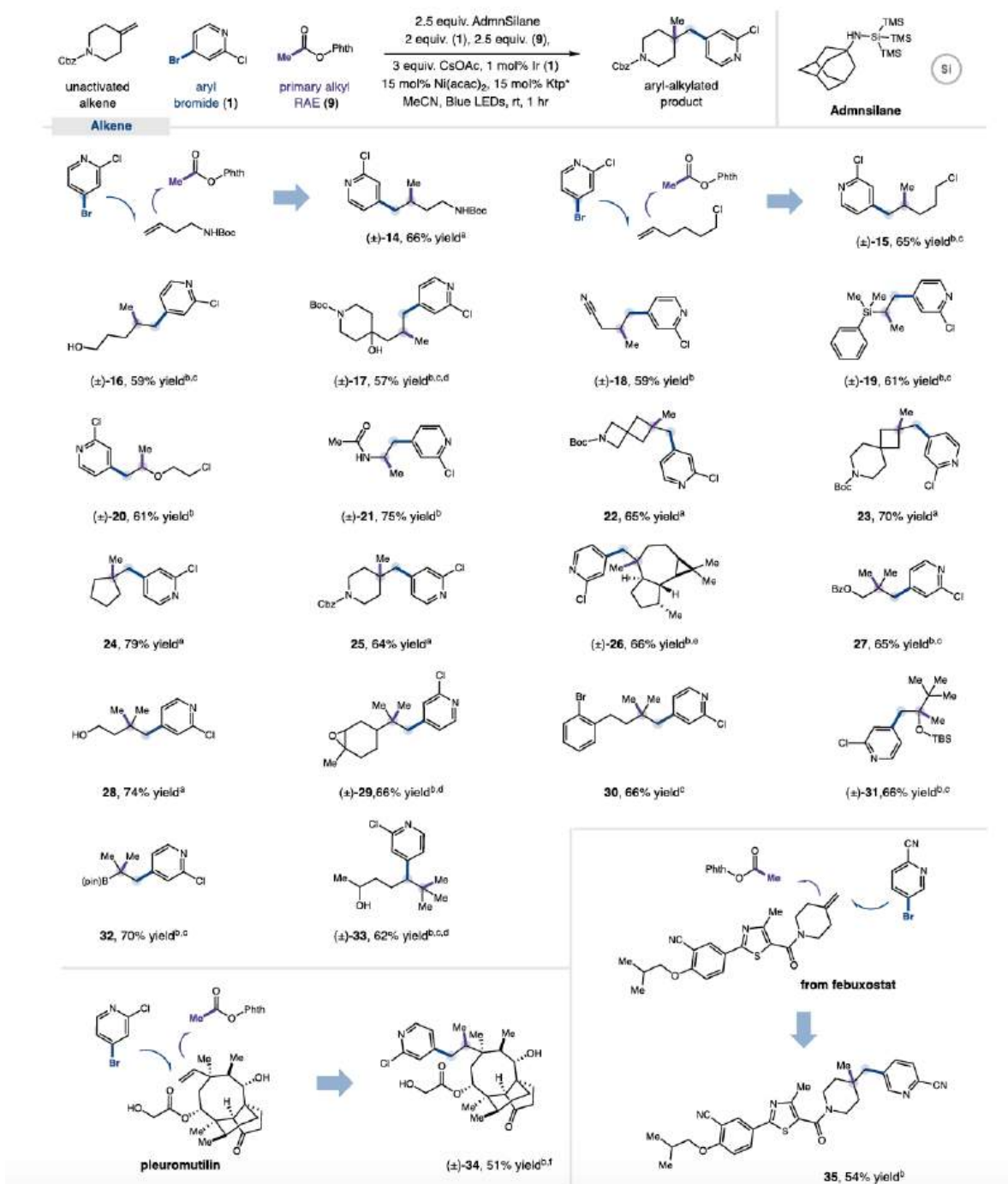
with a reduced hazard to the user. This ensured the solubilization of the alkane starting material, which is challenging to do in a batch setting. However, as with the Noël work, applications towards late-stage functionalization are yet to be fully explored.



Scheme 5. The functionalization of methane and ethane using cerium (III) salts as photocatalysts (Hu et. al. 672).

An alternative approach to these C-H methylation strategies to incorporate new methyl groups into drug compounds was suggested in 2024, when MacMillan and co-workers disclosed an approach towards the arylalkylation of alkenes through a triple radical-sorting mechanism (Wang et. al. 15700). A general method to achieve arylalkylation of unactivated alkenes would involve generating both aryl radicals and alkyl radicals in the presence of an alkene, which could avoid traditional limitations with alternative cross-coupling strategies such as competitive β -hydride elimination. Through a nickel-catalyzed bimolecular homolytic substitution (SH2) mechanism, they were able to use the concept “radical sorting” to avoid other limitations with radical reactions, including uncontrolled radical recombination. Arylalkylation in this way effectively distinguished between aryl radicals, primary alkyl radicals and hindered alkyl radicals generated by aryl radical addition into unactivated alkenes, which could then be coupled together selectively and in the desired order. With this reaction, a range of alkenes were methylated at the most hindered position, whilst also receiving an aryl substituent at the least hindered position. A range of aryl and heteroaryl groups were successfully engaged, as were a range of alkenes, each bearing a variety of functional group substituents (Scheme 6). Interestingly, APIs and natural products containing olefin substituents could be derivatized in reasonable to good yields, providing an opportunity for dual-late-stage functionalization. They also demonstrated as an

extension of the arylalkylation protocol, a so-called “couple close protocol” strategy, in which three bonds were formed in two steps to generate highly substituted fused heterocyclic fragments. In summary, both approaches create novel pathways to challenge traditional limitations and help accelerate access to pharmaceutical C(sp³) rich chemical space.

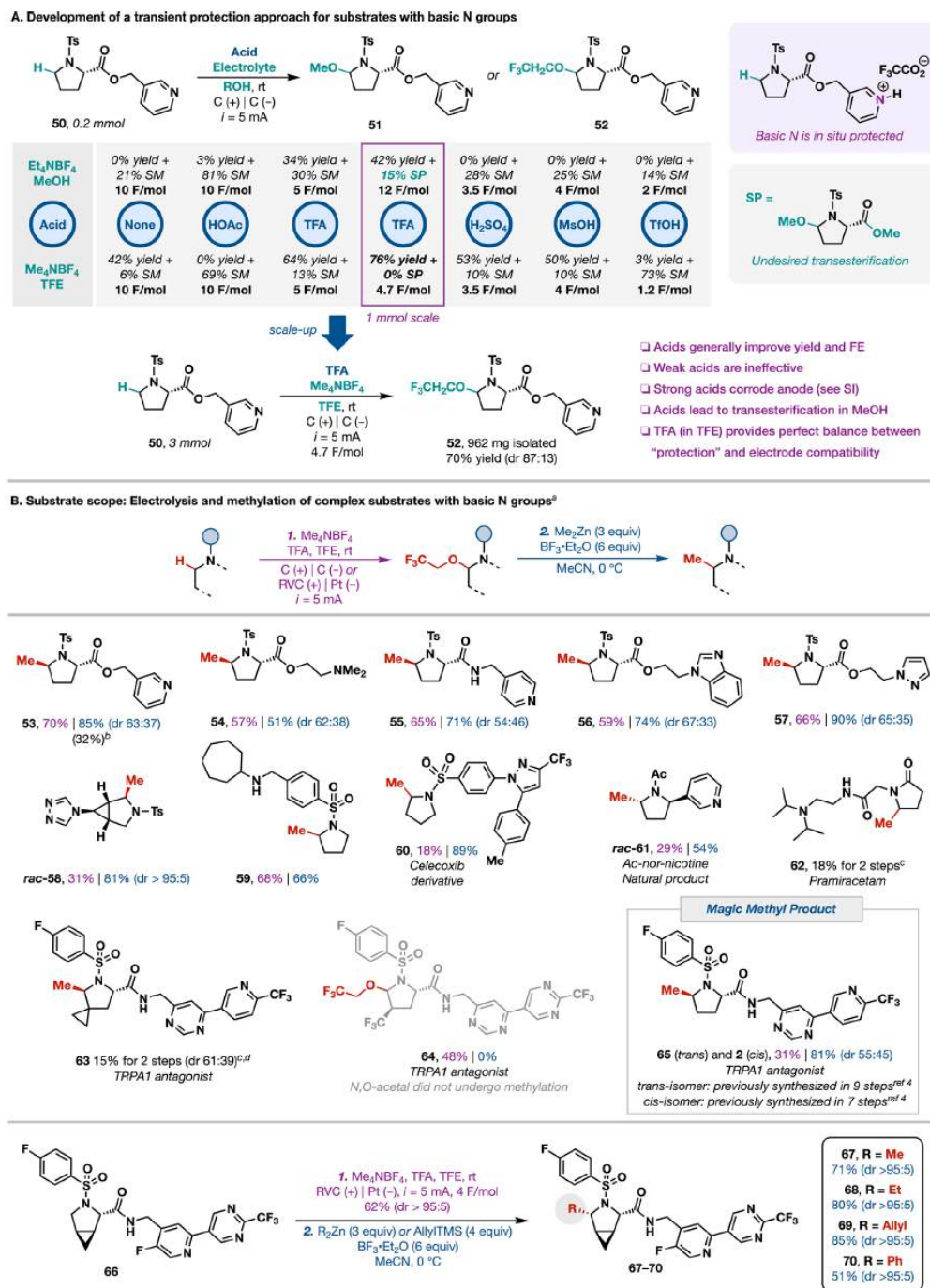


Scheme 6. The photocatalytic arylmethylation of olefins (Wang et. al. 15700).
Electrochemical C-H methylation

Electrosynthesis has a history of nearly 200 years and has been an important field within synthetic chemistry. Indeed, many developments within the field have enabled the development of some of the chemical industry's most important processes, such as the Simons fluorination process and the Monsanto adiponitrile process. Its ability to selectively perform redox chemistry on particular functional groups has also led to the development of important well-established reactions, such as the Shono oxidation (Wang et. al. 11335). Developments in the late 20th century with electrochemical technologies, reactors and chemistries has resulted in the discovery of electroauxiliaries and *cation pool* and indirect electrolysis *via* mediators. These increased the viability of electrosynthesis in bespoke synthesis settings. However, in recent years, these key innovations have resulted in the successful application of electrosynthesis towards the C-H activation, reductive cross-coupling, alkene difunctionalization, radical chemistry, total synthesis and the LSF of natural products and medicinally relevant compounds (Wang et. al. 11335). Crucially, towards the efforts in electrochemical LSF has been the ability to induce chemoselective functionalization under reducing or oxidizing conditions, that was challenging beforehand. Electrocatalysis in particular has enabled several strategies involving site-selective C-H activation for LSF. Amongst these strategies are some notable advances in late-stage C-H methylation reactions (Liao et. al. 8203), which can be broadly split into 2 camps. Firstly, site-selective C(sp³)-H α -methylation of amines and aliphatic azaheterocycles *via* anodic oxidation in the same fashion as Shono oxidations. Secondly, the synergistic merging of electrosynthesis and transition metal catalysis, or so-called electrocatalysis, for site-selective C(sp²)-H methylation of aromatic rings containing a substituent able to act as a directing group.

The most notable example of electrochemical C(sp³)-H methylation in recent years is that of Neurock, Terrett, Lin and co-workers in 2022 (Novaes et. al. 1197)). Inspired by Shono's oxidation reaction, these authors developed a new electrochemical method to overcome traditional methods of late stage methylation in compounds with basic nitrogen containing groups. They began by testing Shono oxidation conditions on six arylsulfonamides with various electronic properties. They found that substrates with electron-withdrawing groups such as bromine and nitro required higher oxidation potentials, which also resulted in competitive solvent oxidation and significant efficiency losses. Therefore, the researchers employed solvents that are more resistant to anodic oxidation for the development of the methylation protocol. This method allowed for the methylation (and alkylation/acylation) of various substrates, which was optimized for yield and selectivity whilst maximizing its compatibility with other functional groups. The reaction proceeded *via* oxidation of the azaheterocycle to the corresponding iminium ion, which was trapped out by the trifluoroethanol solvent as the hemiaminal. Treating this initial product with dimethylzinc formed the desired methylated analogue. The researchers extended their methodology to apply it onto compounds with basic nitrogen-containing groups in pharmaceuticals. By adding Bronsted acid, such as TFA, researchers found that it protected the basic nitrogen *in situ*, which broadened the range of methylated substrates and expanded the prevalence to apply it into medicinal compounds (Scheme 7). In summary, this method efficiently installed methyl and other alkyl/aryl groups in complex amine derivatives using a

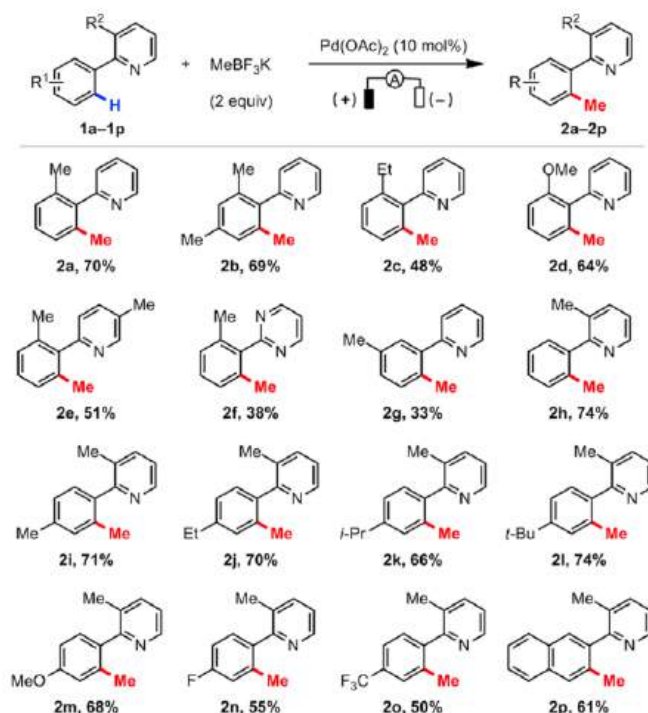
telescoped electrochemical platform.



Scheme 7. Electrochemical telescoped α -amino C-H methylation with the use of TFA to protect basic nitrogen groups (Novaes et. al. 1197).

There have been a few recent applications of electrocatalysis for site-selective C(sp²)-H methylation of aromatic rings containing a substituent able to act as a directing group. In 2017,

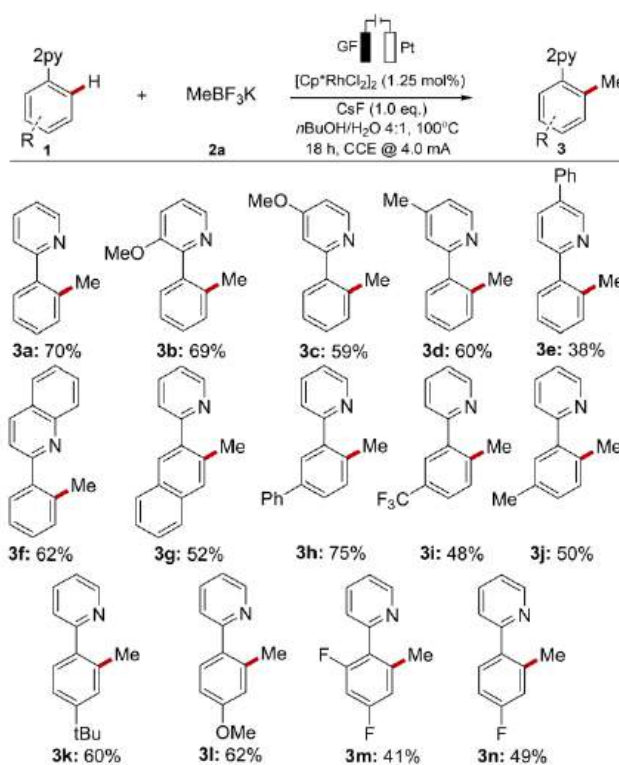
Mei and co-workers reported the use of oxime directing groups with palladium-electrocatalysis to install methyl groups *ortho*- to the oxime (Ma et. al. 12192). This was made possible with the use of potassium methyltetrafluoroborate and enabled the preparation of a selection of methylated aryloximes. In 2019, Li, Wu, Mei *et al.* reported a method that involved a Pd-catalyzed electrochemical C–H alkylation of arenes and alkyl boron reagents using an undivided cell in water (Yang et. al. 1212). The reaction tolerated a range of functional groups, particularly favoring substrates with electron-rich substituents (Scheme 8). In addition, the authors explored other alkyl boron reagents, including potassium trifluoroethylborate and potassium trifluorobutylborate. These proved effective by their moderate yield. However, potassium trifluorophenethylborate resulted in lower yields due to reduced conversion, indicating limitations in the method’s applicability to certain alkyl groups. To probe the reaction mechanism, Li, Wu, Mei and colleagues synthesized a palladacycle intermediate and conducted electrochemical and kinetic studies, which the results suggest that the methylation usually occurs at either Pd(III) or Pd (IV) oxidation states and that C-H activation is the rate determining step. Therefore, the researchers proposed a mechanism where the nitrogen atom coordinates with the palladium catalyst, triggering C–H activation and formation of palladacycle intermediates, which is then oxidized, alkylated and then reductively eliminates to produce the products and regenerated Pd(II).



Scheme 8. Pyridine-directed electrocatalytic C-H methylation with the use of palladium (Yang et. al. 12192).

In 2022, Ackermann and colleagues disclosed a method using Rhodium-electrocatalyzed C–H methylation of 2-aryloximes while using a user-friendly, undivided cell setup without

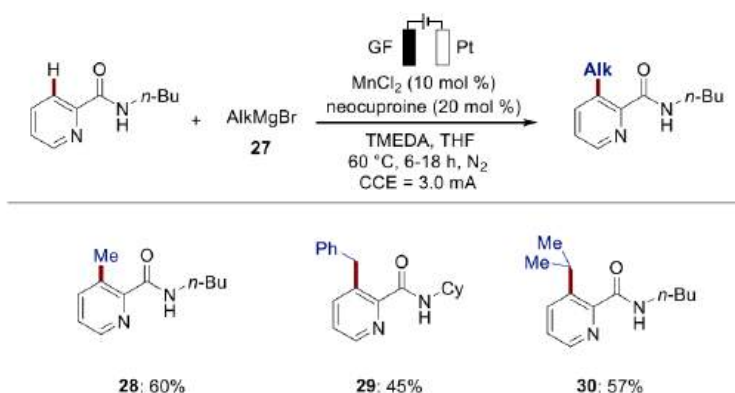
additional electrolytes (Kuciński et. al. 5). Starting by optimizing the reaction conditions, the researchers found that electron-rich and electron-deficient compounds were successfully methylated, with competent substrates such as purines demonstrating high selectivity and moderate yield (Scheme 9). This method was also effective for the mono- and dimethylation of diazepams. Then, the researchers extended their investigation to other alkylations, such as potassium vinyltrifluoroborate with $[\text{Cp}^*\text{RhCl}_2]_2$. To investigate the reaction mechanism, the researchers conducted an experiment using isotopically labeled compounds in the divided cell setup, and found out that a catalytic cycle, involving cyclo-rhodation and high-valent pathways, was likely. This in turn led to C–H alkylation and reduction to form the final product. In summary, the researchers demonstrated a viable late-stage C–H methylation strategy, though the method was limited by the directing group and the requirement for an expensive rhodium catalyst. However, this work was followed up by Yang, Guo and colleagues, who demonstrated that a wider range of azine directing groups and boron-derived methylating agents could enable such a C–H methylation using rhoda-electrocatalysis (Yang et. al. 1212).



Scheme 9. Azine-directed electrocatalytic C–H methylation with the use of rhodium (Kuciński et. al. 5).

Avoiding the use of expensive transition metals would increase the attractiveness of an electrocatalytic LSF procedure to install methyl groups. In 2021, Ackermann and colleagues reported an manganese electrocatalytic C–H acylation and alkylation reaction (Massignan et. al. 11649). The reaction enabled conversion of various amide-substituted azines, however it also

demonstrated a single example of a C-H methylated product, perhaps unlocking a potentially general manganese-catalyzed C-H methylation reaction (Scheme 10). Cyclic voltammetry and SEM-EDS studies provided some explanation for the mechanism of the catalytic cycle, suggesting that it involves the formation of an active manganese(II) complex, followed by C–H activation, transmetalation, and an anodic oxidation event, leading to the final product and regenerating the active catalyst. In summary, this approach highlights the development of the first manganese-catalyzed electrocatalytic C(sp²)-H methylation reaction and its potential to be used for late-stage functionalization in pharmaceutical research.



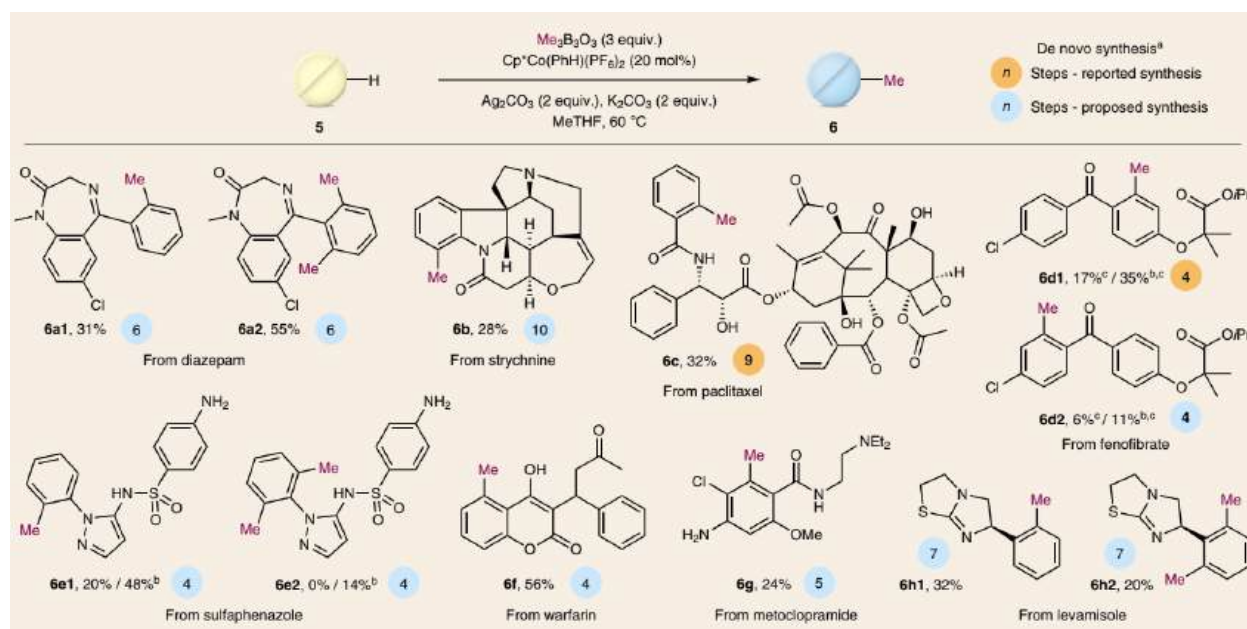
Scheme 10. Amide-directed electrocatalytic C-H methylation with the use of manganese (Massignan et. al. 11649).

Transition Metal-Catalyzed C-H methylation

The use of transition metals to activate otherwise inert C-H bonds has emerged as a useful tool for LSF (Castellino et. al. 8153). Indeed, the use of precious transition metals for this purpose has now been expanded to 3d transition metals, offering potentially cheaper methods to access derivative compounds. The majority of transition metal-catalyzed C-H activation strategies make use of directing groups, either exogenous to the substrate (and so are added on beforehand and removed afterwards) or as an inherent part of its structure, which is generally the more popular method.

In 2020, Friis, Johansson and Ackermann presented a strategy for the cobalt-catalyzed late stage C-H methylation of structurally complex drug molecules (Friis et. al. 519). They used a multiparameter optimization approach in this investigation, which allowed them to examine many possible cross-optimization and combinations. This resulted in the development of a directed aryl C-H methylation strategy using Me₃B₃O₃ as the methylating agent and a dicationic Cp*-coordinated cobalt catalyst under relatively mild conditions. In the methylation catalyzed by earth-abundant cobalt, the authors were able to compare the reactivity between different substrates as well as ranking the diverse selection of additives applied in the compatibility screen and how well they were tolerated. A range of directing groups were viable for this reaction, enabling the preparation of a variety of C-H methylated arenes. By applying these findings to the late-stage functionalization of more structurally complex drug compounds, a variety of methylated drug and natural product derivatives were prepared in modest to good yields (Scheme

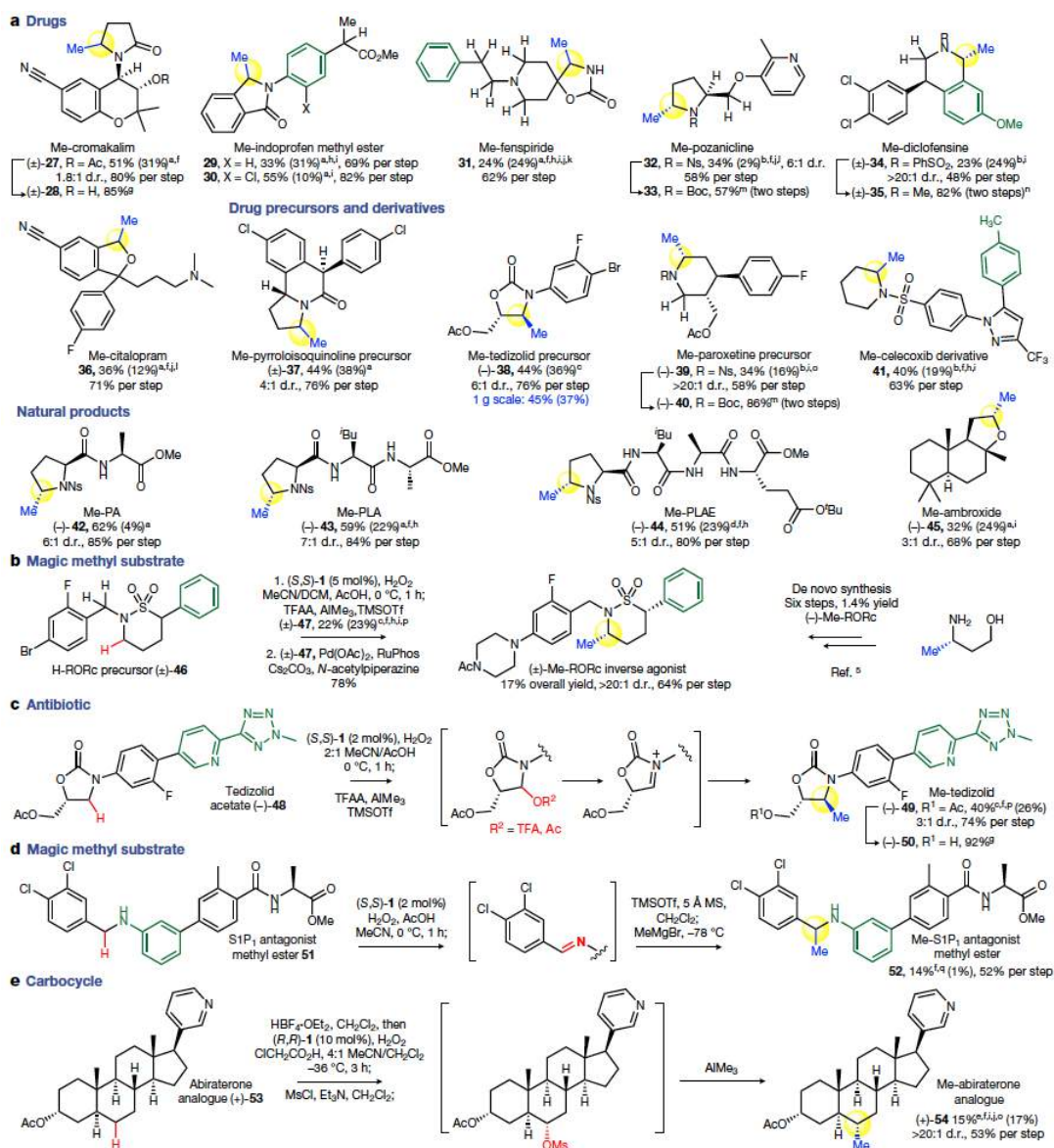
11). Crucially, the authors emphasize the advantages of using an earth-abundant metal catalyst for this reaction, which still retains conversions and selectivities of more expensive catalyst choices. However, under the reaction conditions reported, a stoichiometric excess of Ag_2CO_3 is still required for the reaction to occur, detracting from its reported cost effectiveness. To fully realize a cost-effective procedure, a re-optimization would need to be conducted to remove the necessity for expensive additives.



Scheme 11. Late-stage directed cobalt-catalyzed C-H methylation of APIs (Friis et. al. 519).

In 2019, White and colleagues disclosed a new method for regioselective and chemoselective oxidative $\text{C}(\text{sp}^3)\text{-H}$ methylation of *N*- and *O*-containing aliphatic heterocycles (Feng et. al. 627). Akin to other strategies involving the oxidation of the α -position of aliphatic amines, this method initially forms a hydroxylated intermediate before methylation at the α -position. A manganese catalyst enabled faster C-H hydroxylation whilst achieving high chemoselectivity and yields for the hydroxylated products. In more complicated scenarios, the hydroxylated intermediates were separated before adding ethyl groups, in order to avoid the formation of inseparable mixtures. The researchers found that previous catalysts involving iron produce complex mixtures, but that this particular manganese catalyst lacked the reactivity for the hydroxylation of complex substrates without the addition of other additives. To enable the methylation of more complex structures, White and colleagues developed strategies for different functional groups. For compounds with Lewis basic or electrophilic functional groups, they use fluorine-assisted methylation with reagents like diethylaminosulfur trifluoride (DAST) or boron trifluoride diethyl etherate ($\text{BF}_3 \cdot \text{OEt}_2$) to activate hydroxyl groups. For molecules lacking these functional groups, the BF_3 activation method was used. 22 compounds with various heterocycles were successfully methylated in good yields using the DAST-activation strategy. Additionally,

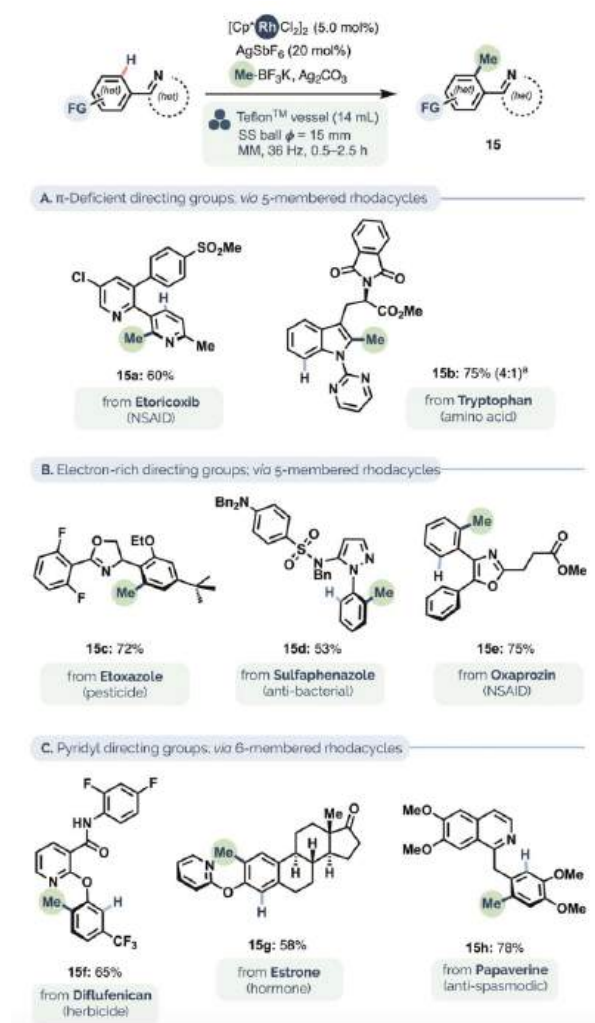
the oxidative methylation method demonstrated chemoselectivity by methylating specific sites on substrates while avoiding promiscuous interception of other reactive groups, which makes this method beneficial for the development of pharmaceutical compounds and providing pathways for late stage functionalization (Scheme 12).



Scheme 12. Manganese-catalyzed C(sp³)-H late-stage methylation of APIs (Feng et. al. 627).

In 2021, Pilarski and colleagues disclosed a mechanochemical method enabling catalyzed C-H methylation under solvent-free conditions (Ni et. al. 6666). They noted that the increasing demand for C-H methylation strategies in drug production provides an opportunity for transition metal catalysis to achieve this goal. However, there are environmental concerns with the conduction of many chemical reactions. This study therefore introduces a mechanochemical method for C-H methylation, which offers a more sustainable alternative to other methods, by

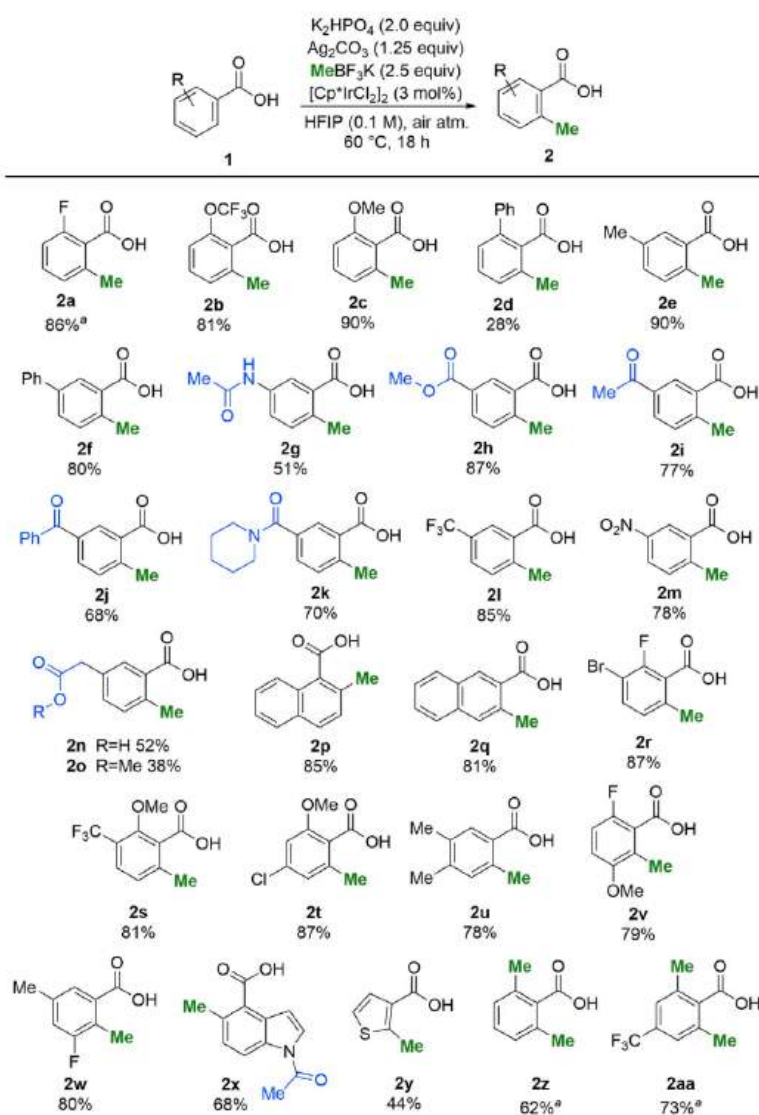
avoiding the use of solvents and carrying out air-sensitive reactions under aerobic conditions. While experimenting, the researchers used phenylpyridine as a model substrate and MeB(OH)_2 as the methyl source to avoid otherwise toxic methylating reagents and found that the catalyst $[\text{Cp}^*\text{RhCl}_2]_2$ was most effective. This method achieved high yields (up to 92%) with significantly shorter reaction times than traditional methods and also demonstrated high regioselectivity. Thus, the researchers then applied this method to the late-stage methylation of several bioactive compounds containing aryl azines as well as arenes bearing alternative directing groups, which demonstrated its high versatility and compatibility with various bioactive compounds. In summary, this method demonstrated how mechanochemistry can be utilized for the LSF of complex bioactive molecules (Scheme 13), which attempts to greatly improve molecule potency and promote environmental sustainability. However, the use of highly expensive transition metals perhaps reduces the attractiveness of such a strategy.



Scheme 13. Rhodium-catalyzed $\text{C}(\text{sp}^2)\text{-H}$ late-stage methylation of APIs (Ni et. al. 6666).

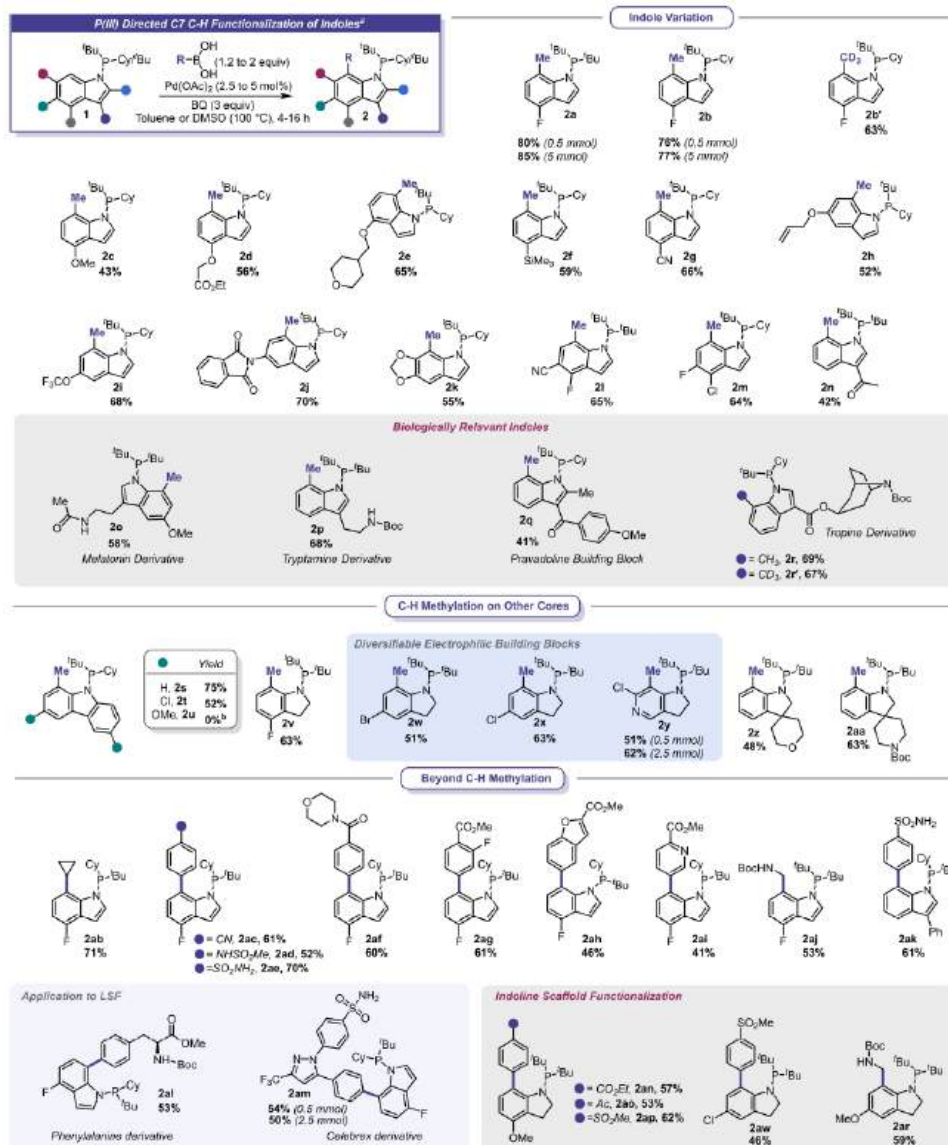
In 2021, Johansson and co-workers suggested a new approach of utilizing iridium

catalyzed carboxyl-directed C-H methylation for the *ortho*-methylation of benzoic acid derivatives (Scheme 14) (Weis et. al. 102467). This approach suggested that by using an iridium catalyst, high selectivity for *ortho*-methylation can be achieved for complex molecules. Indeed, several biologically active compounds were successfully derivatized using this strategy, producing compounds with enhanced pharmacokinetic properties. Though the potential benefits are still being investigated, Johansson and co-workers used different methyl sources for simulation, which found out that it has very similar yields and ensures the metabolic stability of methylated analogues. As with other strategies, this work is limited from being readily adopted for large-scale applications due to the cost of both the iridium catalyst and the stoichiometric AgCO_3 base.



Scheme 14. Iridium-catalyzed C(sp²)-H late-stage methylation of benzoic acids (Weis et. al. 102467).

In 2024, Kelly, Salinas and co-workers were seeking a method for the LSF of tryptophan moieties to produce unnatural amino acids (UAAs) that could be important for modifying peptide drugs (Kelly et. al. 2055). UAAs can be valuable building blocks and improve drug properties, such as potency and stability. Here they developed a reaction for the site-selective C-H methylation of indolyl cores, which could be subsequently used to target tryptophan-containing peptides. They eventually found that using a di-*tert*-butylphosphine substituent on the indole nitrogen in combination with a palladium catalyst and benzoquinone oxidant allowed selective C-H methylation of the indole core and the C7 position. They found that lowering the temperature decreased the yield, and that Pd(OAc)₂ was an effective catalyst to control methylboronic acid levels. The use of a cyclohexyl-*tert*-butyl phosphino directing group improved the stability, robustness and reactivities during late stage functionalization, and enabled the reaction to be conducted on a gram scale on complex molecules (Scheme 15). Directly applying the developed strategy to tryptophan functionalization, they produced the desired C7-modified product, indicating the feasibility of creating unnatural amino acids. The ultimate goal was to incorporate C-H functionalization into solid-phase peptide synthesis (SPPS) for on-peptide editing, which would provide protection and purification of the peptide structure and activities. In summary, this research produced a method to enhance peptide properties and is applicable to understand complex biological systems in pharmaceutical research.



Scheme 15. Palladium-catalyzed C(sp²)-H late-stage methylation of di-*tert*-butylphosphine substituted indoles (Kelly et. al. 2015).

Conclusion

Overall, there has been much interest in the development of C-H methylation strategies for late-stage functionalization. These methods broadly target C-H bonds that can either be activated by a metal with a tethered directing group, or C-H bonds that can be abstracted radically, or oxidized electrochemically/photochemically. The targeting selectively of different C-H bonds within a drug candidate remains therefore elusive unless the structure contains C-H groups distinguishable by the above methods. Developing further methods which would allow the selective methylation of different C-H bonds on the same structure to build libraries of methylated derivatives of the same compounds would be highly advantageous to drug development. Therefore, there remains much to be seen in this area in the future.

Works Cited

- Guillemard, L.; Kaplaneris, N.; Ackermann, L.; Johansson, M. J. *Nat. Rev. Chem.* 2021, 5, 522 - 545.
- Börgel, J.; Ritter, T. *Chem.* 2020, 6, 1877 - 1887.
- Aynetdinova, D.; Callens, M. C.; Hicks, H. B.; Poh, C. Y. X.; Shennan, B. D. A.; Boyd, A. M.; Lim, Z. H.; Leitch, J. A.; Dixon, D. J. *Chem. Soc. Rev.* 2021, 50, 5517 - 5563.
- Leung, C. S.; Leung, S. S. F.; Tirado-Rives, J.; Jorgensen, W. L. *J. Med. Chem.* 2012, 55, 4489 - 4500.
- Njardson group website. Accessed 2024-08-09:
<https://njardarson.lab.arizona.edu/content/top-pharmaceuticals-poster>.
- Zhang, X.; Li, S.; Qiu, F.; Ang, H. T.; Wu, J.; Jia, P. *Green Chem.* 2024, 26, 3595 - 3626.
- Bellotti, P.; Huang, H.-M.; Faber, T.; Glorius, F. *Chem. Rev.* 2023, 123, 4237 - 4352.
- Liao, F.; Wei, Z.; Guan, Y.; Zhuang, Z.; Xu, K.; Tan, J. *Green Chem.* 2024, 26, 8161 - 8203.
- Vasilopoulos, A.; Krska, S. W.; Stahl, S. S. *Science* 2021, 372, 398 - 403.
- Mao, E.; MacMillan, D. W. C. *J. Am. Chem. Soc.* 2023, 145, 2787 - 2793.
- Zhu, K.; Ma, Y.; Wu, Z.; Wu, J.; Lu, Y. *ACS Catal.* 2023, 13, 4894 - 4902.
- Laudadio, G.; Deng, Y.; van der Wal, K.; Ravelli, D.; Nuño, M.; Fagnoni, M.; Guthrie, D.; Sun, Y.; Noël, T. *Science* 2020, 369, 92 - 96.
- Hu, A.; Guo, J.-J.; Pan, H.; Zuo, Z. *Science* 2018, 361, 668 - 672.
- Wang, J. Z.; Mao, E.; Nguyen, J. A.; Lyon, W. L.; MacMillan, D. W. C. *J. Am. Chem. Soc.* 2024, 146, 15693 - 15700.
- Wang, Y.; Dana, S.; Long, H.; Xu, Y.; Li, Y.; Kaplaneris, N.; Ackermann, L. *Chem. Rev.* 2023, 123, 11269 - 11335.
- Novaes, L. F. T.; Ho, J. S. K.; Mao, K.; Liu, K.; Tanwar, M.; Neurock, M.; Villemure, E.; Terrett, J. A.; Lin, S. J. *Am. Chem. Soc.* 2022, 144, 1187 - 1197.
- Ma, C.; Zhao, C.-Q.; Li, Y.-Q.; Zhang, L.-P.; Xu, X.-T.; Zhang, K.; Mei, T.-S. *Chem. Commun.* 2017, 53, 12189 - 12192.
- Yang, Q.-L.; Li, C.-Z.; Zhang, L.-W.; Li, Y.-Y.; Tong, X.; Wu, X.-Y.; Mei, T.-S. *Organometallics* 2019, 38, 1208 - 1212.
- Kuciński, K.; Simon, H.; Ackermann, L. *Chem. Eur. J.* 2022, 28, e202103837.
- Yang, Q.-L.; Liu, Y.; Liang, L.; Li, Z.-H.; Qu, G.-R.; Guo, H.-M. *J. Org. Chem.* 2022, 87, 6161 - 6178.
- Massignan, L.; Zhu, C.; Hou, X.; Oliveira, J. C. A.; Salamé, A.; Ackermann, L. *ACS Catal.* 2021, 11, 11639 - 11649.
- Castellino, N. J.; Montgomery, A. P.; Danon, J. J.; Kassiou, M. *Chem. Rev.* 2023, 123, 8127 - 8153.
- Friis, S. D.; Johansson, M. J.; Ackermann, L. *Nat. Chem.* 2020, 12, 511 - 519.
- Feng, K.; Quevedo, R. E.; Kohrt, J. T.; Oderinde, M. S.; Reilly, U.; White, M. C. *Nature* 2020, 580, 621 - 627.
- Ni, S.; Hribersek, M.; Baddigam, S. K.; Inger, F. J. L.; Orthaber, A.; Gates, P. J.; Pilarski, L. T. *Angew. Chem. Int. Ed.* 2021, 60, 6660 - 6666.
- Weis, E.; Hayes, M. A.; Johansson, M. J.; Martín-Matute, B. *iScience* 2021, 24, 102467.
- Kelly, C. B.; Padilla-Salinas, R.; Chen, W.; Muuronen, M.; Balsells, J. *Adv. Synth. Catal.* 2024, 366, 2044 - 2055.

Applying AI to Enhance Stock Price Prediction: Improving Investment Strategies and Risk-Management Through Data-Driven Insights By Lakshya Gupta and Odysseas Drosis

Abstract

In the finance world, “beating the stock market” by predicting the stock prices of different companies is imperative for a multitude of people. For company leaders, it aids in making informed investment decisions, mitigating risks, and maximizing profits (since large financial institutions rely on these predictions for quick decision-making and risk management, as well as planning strategic moves like mergers or raising capital for a new business venture). For individuals, accurate predictions of future stock prices can be the major difference between earning profits and incurring losses.

As a result of these reasons, our study explores how artificial intelligence (AI) can potentially be used as a transformative method to enhance stock price prediction compared to traditional, more speculative methods. We aimed to identify the most effective AI models, ranging from linear regression to neural networks to decision trees to a sentiment analysis model, and determine the most optimal data sources, including historical prices and tweets about companies. Simultaneously, we aimed to ensure that our AI models can adapt to market anomalies and sudden shocks.

Our methodology involved developing, combining, and testing three AI models on datasets comprising historical stock prices from various S&P 500 companies such as Apple, Google, General Motors, and Microsoft. We obtained these datasets from Yahoo Finance to ensure accurate historical stock prices were factored into our models. By leveraging a multiplicative weight update approach to combine these models, we enhanced prediction accuracy and robustness, particularly in volatile market conditions. Moreover, we obtained a dataset containing 3.3 million tweets about NASDAQ companies which was used in a sentiment analysis model that backed up the predictions from the other three AI models. The results indicate that our overall AI model (on average) achieves a mean squared error of 7.23, which outperforms traditional prediction methods. These findings underscore the transformative potential of AI in financial markets, exhibiting the need for future research and practical implementation of AI-driven investment tools.

Introduction

In recent years, the media and public have fabricated considerable attention because of the prospect of integrating financial markets with artificial intelligence (AI). This is driven by the huge potential that AI can revolutionize the forecasting of stock prices and investment strategies employed by investors. Normally, stock price prediction methods are performed by fundamental analysis and technical analysis, which rely on speculation, some historical data, and economic indicators (Garg et. al. 24). In contrast, AI can offer highly vigorous solutions, by utilizing machine learning algorithms and vast datasets to uncover patterns in data that many statistical methods may miss. This would explain why the subject is highly relevant to the finance field

since accurate predictions can lead to reduced risk and better-informed investment decisions (which, in turn, improves financial outcomes).

Despite the advancement of AI in recent years, the field of stock price prediction still needs to be improved, with challenges created through notable gaps and inconsistencies in models created through existing research. These include integrating only one data source, hardly handling “noisy data” in datasets that haven’t been properly cleaned (Sun et. al. 1738), and the lack of interpretability in these AI models. Moreover, although some studies have depicted promising results, the practical application of these models in real-world trading scenarios is limited due to their complexity and vast computational requirements. We addressed this in our study when we developed and evaluated our own AI models to improve stock price prediction since addressing the limitations of existing models can contribute to the broader understanding of AI applications in finance, paving the way for more robust and reliable predictive systems in the future.

Some of our primary objectives in this study were to: (1) develop AI models that can integrate diverse data sources for stock price prediction (2) refine our models’ performance in terms of their prediction accuracy and practical applicability, and (3) ensure that our AI models can provide useful statistics, which was done at the end of our code by allowing users to enter their initial investment amount, and calculate their Return On Investment (ROI) & profit after 5 years. On top of this, it’s also important to mention the wide scope of our study, which includes the development (and testing) of AI models using real-time data and historical data from the last 5 years using various large companies’ stock prices from the S&P 500 index. However, the training and testing of the AI models have been constrained by limited computational resources and a lack of preprocessed data for local stock markets outside the USA.

We developed a technique to leverage supervised learning models, including our neural networks and decision trees, to create one huge classifier model. We extracted data from Yahoo Finance, news archives, and social media sites such as X (formerly known as Twitter). Data collected were cleaned and normalized using libraries in Python such as NumPy and Pandas. The models were trained and evaluated with standard metrics, such as mean-squared error, but with a special focus on understanding their implications for investment strategies and risk management.

Methods

This section outlines the AI models employed in this study to perform stock price predictions, each offering unique approaches to capturing relationships within the financial data presented to it. We will focus on the methodologies used in our four models (in our case: a linear regression model, a neural network, a decision tree, and a sentiment analysis model). We also touch on our datasets and their sources

Linear Regression

For starters, let’s discuss our linear regression model. It’s a statistical method that models the relationships between an independent variable (which was the historical values of a single

stock's value of a company, spanning 5 years) and a dependent variable (current stock price). This model assumes a linear relationship, which the following equation can express:

$$Y = P_0 + P_1X_1 + P_2X_2 + \dots + P_nX_n + \epsilon$$

Where Y represents the predicted stock price; X_1, X_2, \dots, X_n are the independent variables; P_0 is the y-intercept; P_1, P_2, \dots, P_n are the coefficients, and ϵ is the error term. The coefficients are estimated using a Mean Squared Error equation (shown below). This equation measures error in statistical models by using the average squared difference between observed and predicted values. Initially, the coefficient values are randomly generated, and then the function fine-tunes them to reach a global minimum that yields the lowest mean squared error. In turn, we can then create the least square regression line (LSRL), also known as a line of best fit, which gives us our linear regression model.

$$\frac{1}{2n} \times \sum (y_i - p_i)^2$$

Where y_i is the i^{th} observed value, p_i is the corresponding predicted value for y_i , and n is the number of data points. The Σ indicates that a summation is performed over all values of i until it parses through all the data points, n .

Once trained on historical data, these coefficients are applied to forecast future stock prices. Although its primary limitation lies in its assumption of linearity in the correlation between our dependent and independent variables — which may not consistently hold true in complex financial markets — linear regression is particularly valued for its simplicity and ease of interpretation, making it a useful baseline against which more complex models can be compared, like our neural network and decision tree models (Hastie et. al. 44).

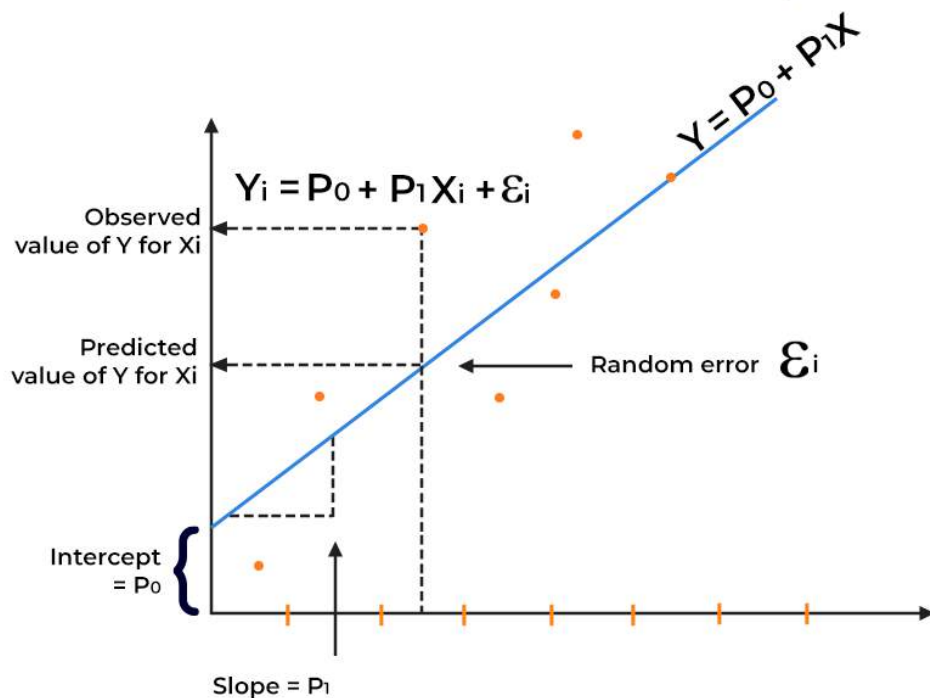


Fig 1: Visualization of Linear Regression Model (source: www.spiceworks.com/tech/artificial-intelligence/articles/what-is-linear-regression/)

Neural Network

We used a multi-layer perceptron (MLP) — a type of neural network. They were used to capture nonlinear relationships in the data. A feedforward neural network, consisting of multiple layers of neurons (like an input layer, multiple hidden layers, and an output layer) is trained using the backpropagation algorithm, which adjusts the weights of connections between neurons to minimize prediction errors based on the output of the mean squared error function (Singh 1). Activation functions like ReLU (Rectified Linear Unit) introduce nonlinearity, allowing the model to learn complex patterns. Hyperparameters such as the number of layers, neurons per layer, and learning rate are optimized through cross-validation. Neural networks are valuable in stock price prediction because they can capture complex, nonlinear relationships in financial data, which are often missed by simpler models like linear regression. Their ability to learn from diverse inputs (like historical prices and sentiment analysis) and adapt to changing market conditions (as more data becomes available) makes them particularly effective in modeling the intricate patterns that drive stock prices in the dynamic environment of financial markets.

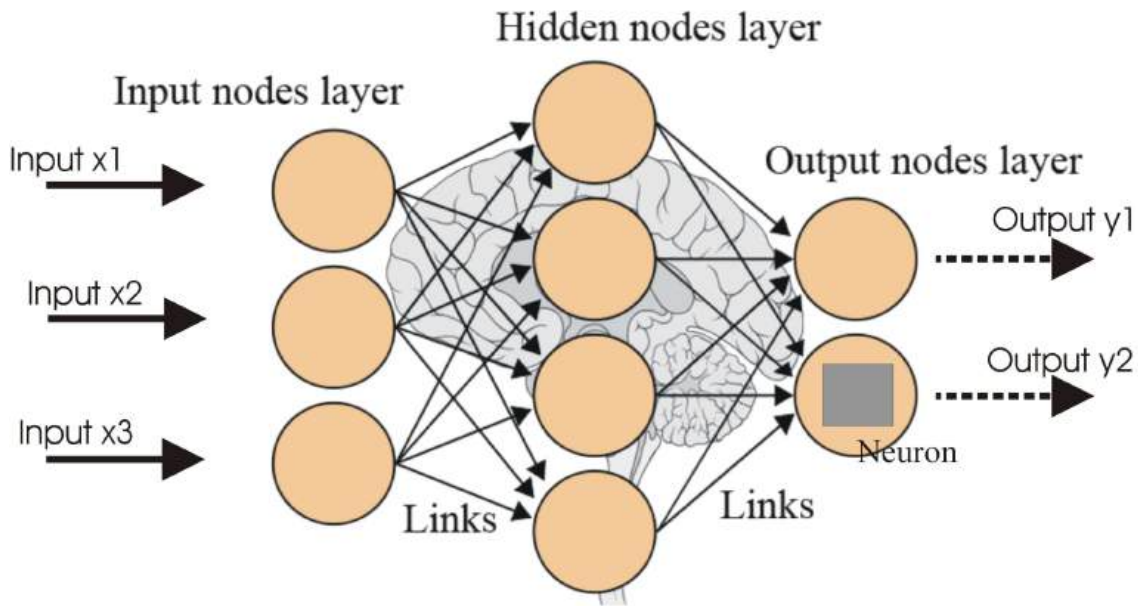


Fig 2: Visualization of MLP Model (source: medium.com/@avicebooks/part-18-ml-deep-learning-and-reinforcement-learning-468f7340aaf1)

Decision Tree

A decision tree is a supervised machine learning algorithm that can be utilized to solve regression problems like our stock price prediction problem. The recursive division of the data space according to the most informative feature is achieved by the decision tree algorithm; the tree-like structure grows as each internal node of the tree represents a decision taken regarding which feature is most informative, each branch represents an outcome of that decision, and each leaf represents the final prediction made based on the path taken to arrive at it (IBM 1). Splitting the data is necessary to reduce any prediction error to improve the accuracy, so a clear path can be created from the root node to the leaf node, which ultimately determines the output for any given input. Even if they tend to overfit, they can still be applied effectively to find non-linear relationships in our code without requiring extensive data augmentation or data preprocessing.

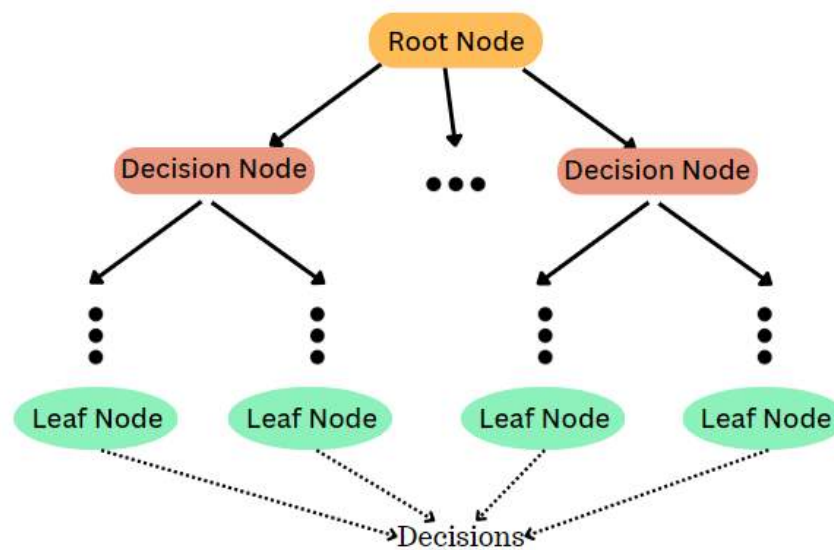


Fig 3: Visualization of Decision Tree (source: medium.com/@nidhigahlawat/decision-trees-a-powerful-tool-in-machine-learning-dd0724dad4b6)

Multiplicative Weight Update

After all three models had been developed and rigorously tested, the multiplicative weight update technique accomplished the task of combining the predictions forecasted by the three models into a single output by assigning weights to each model based on their performance and iteratively adjusting those weights.

Initially, we made all models have equal weights. For each prediction of the next day's stock price, a weighted average was computed by combining the outputs from the linear regression, neural network, and decision tree models, with the weights reflecting their respective contributions. After each prediction, the weights were updated based on the errors of the models. Specifically, the weight for each model was adjusted inversely proportional to its prediction error, with the aim of reducing the influence of models that performed poorly and increasing the influence of those which performed better. The weights were then normalized to ensure they were summed to one, maintaining their relative importance in the weighted average, making it easier to compare the influence of the three models. The final weight updates for the models are shown in the results section.

To summarize, this iterative adjustment process allows the combined model to leverage the strengths of each model while mitigating their weaknesses, ultimately providing a more robust and accurate forecast of stock prices.

Yfinance

In this study, the financial dataset was sourced using the "Yfinance" library on Python, which is an open-source Python library that provides access to historical financial data (and fetches real-time stock data) from Yahoo Finance. For the purpose of this research, only the

historical price data was retained, with all other columns dropped, allowing the models to focus exclusively on predicting stock prices based on this and sentiment analysis.

Tweets Dataset

The tweets dataset was obtained from Kaggle, which is an online community platform that allows users to find datasets they want to use in building AI models, publish datasets, work with other data scientists and machine learning engineers, and enter competitions to solve data science challenges. This particular dataset contains roughly 3.3 million unique tweets which contain information about the text body of the tweet, the post date, the author, the number of likes and retweets, and which company they are referring to. These companies include Amazon, Apple, Google, Microsoft, and General Motors. We solely extracted the body of the text and the associated company (using Pandas) so that our sentiment analysis model could be trained using this information.

Sentimental Analysis

Sentiment Analysis is a Natural Language Processing (NLP) technique that identifies the emotion conveyed by a given text. In our case, we wanted the tweets to be labeled positive, neutral, or negative. Rather than having to read through all the tweets and decide which ones are positive, neutral, or negative, sentiment analysis models can automatically understand the perspectives of people on S&P 500 companies, which we can gain insights from, that can help back up our AI models' predictions. To analyze the sentiments of these tweets, we wielded a pre-trained sentiment analysis model from the transformers library in PyTorch to analyze all the tweets and give a label, as well as the confidence score. These scores were correlated with historical stock prices to determine whether they could be a reliable indicator for prediction purposes. This approach was integral to bridging the gap between the market sentiment with traditional financial data in our predictive combined model.

Results

We split our dataset into three sections: a training set, a validation set, and a test set. As shown below in Table 1, we have displayed the mean squared error for each model based on the validation set and the average mean squared error (for each model) for four familiar S&P 500 Companies, which we used in order to prevent the models from overfitting. Overfitting is an undesired, yet common problem in machine learning which happens when a model gives accurate predictions on the training set, but cannot generalize to unknown cases presented to it in datasets such as the validation or test sets, hindering the model's accuracy. Usually, this takes place when the machine learning algorithm is too complex, and it fits noise or irrelevant patterns that are found from the model's training inputs.

Table 1: Mean Squared Error (MSE) Evaluation of the AI Models on Financial Datasets from Four S&P 500 Companies

S&P Company	MSE (linear)	MSE (neural network)	MSE (decision tree)
Apple	8.09	8.36	10.66
Google	3.53	4.85	6.69
General Motors	0.85	0.87	2.36
Microsoft	21.02	28.01	26.78

To mitigate overfitting, we experimented with hyperparameters for the neural network, such as “max_iter” (max iterations — which is the total number of iterations of all the training data in a cycle for training the neural network — which controls the complexity of the neural network), and also for the decision tree, such as “max_depth” (which controls the overall complexity of the decision tree). Through hyperparameter tuning, we tested various values and found that a max_iter of 450 optimized our neural network on the validation set, and a max_depth of 7 provided the best results for our decision tree model on the validation set, which can be seen in Tables 2 and 3 below:

Table 2: Max Iterations vs. Mean Squared Error (MSE) Evaluation for regulating overfitting in the Neural Network model

Max Iterations	100	300	450	500	600
MSE (train set)	9.69	9.18	8.99	9.75	9.66
MSE (validation set)	10.78	10.25	10.09	9.24	9.84
MSE (test set)	8.26	7.80	7.81	9.73	9.53

Table 3: Max Depth vs. Mean Squared Error (MSE) Evaluation for regulating overfitting in the Decision Tree model

Max_Depth	3	5	7	10
MSE (train set)	37.28	6.25	2.39	0.56
MSE (validation set)	35.53	9.19	10.66	8.81
MSE (test set)	46.71	10.40	8.42	9.62

To boost the accuracy of the model further while minimizing overfitting, we adopted a multiplicative weight update approach. By modifying the weights allotted to the different models based on their performance, this approach produces a prediction framework that is more equitable and resilient. The final weight updates for the models were as follows:

- Linear Model: 0.33538863877723524
- Neural Network: 0.6637687520812994
- Decision Tree: 0.0008426091414653441

This improvement reveals that when the MLP neural network model was optimized using multiplicative weight updates, it produced the best results in our ensemble technique. This highlights even more how important sophisticated model integration techniques are for managing complex financial data and improving prediction accuracy.

Discussion

While our study makes sincere efforts in applying AI to stock price prediction, several open questions remain, which can guide future research and enhance the models developed. One significant area of exploration is the integration of political factors and company financial reports into the predictive models, because policy changes due to elections, and optimistic (or pessimistic) quarterly earnings statements can have notable impacts on a company's stock prices. By incorporating these elements, AI-driven predictions can be more reliable and accurate, especially during volatile market conditions. Another potential avenue for future research is extending these AI models to companies outside the US stock market. Our study focused on the US stock market due to data availability and preexisting financial infrastructure however applying these models to international markets could lead to new insights on companies that don't operate primarily in the USA — albeit with some unique challenges and circumstances like different economic conditions and different regulatory environments, which would require procuring large datasets with financial information of different corporations outside the USA, and possible modifications to the models. It's worth mentioning that many firms on Wall Street have developed sophisticated proprietary mathematical models for stock price prediction, so there is value in investigating how AI models can be integrated with those already in use by Wall Street professionals. Combining these with AI-based approaches could leverage the strengths of both, likely leading to more robust and accurate predictions. These open questions, executed correctly, can bridge the gap between academic research and practical application in the finance industry, encouraging collaboration between researchers and industry professionals, and yield significant advancements in the use of AI-driven models for predicting stock prices.

Works Cited

- Garg, Ashish, and Vikram Mohan. "Comparative Analysis between the Fundamental and Technical Analysis of Stocks." *International Journal of Research in Finance and Marketing*, vol. 6, no. 4, 2016, pp. 15-24. ResearchGate
- Sun, Huidong, et al. "Identifying Big Data's Opportunities, Challenges, and Implications in Finance." *Mathematics*, vol. 8, no. 10, 2020, pp. 1738. MDPI
- Hastie, Trevor, et al. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, 2009.
- Singh, Sonal. "Understanding Multilayer Perceptron (MLP) Neural Networks." 2023, www.shiksha.com/online-courses/articles/understanding-multilayer-perceptron-mlp-neural-networks/.
- IBM. "Decision Trees." IBM, 2024, www.ibm.com/topics/decision-trees.

The Effect of Augmented Reality and Haptics Technology on Student Engagement and Learning in STEM Education By Zachary Jones

Abstract

As our society is quickly evolving to become reliant on technological systems which are evolving just as rapidly, it is apparent that the need for a workforce that is highly skilled in the fields of Science, Technology, Engineering, and Math (STEM) is at an all-time high. However, if our educational systems and pedagogical models, which aim to produce the next generation of STEM workers, are not evolving at the same rate as our society's need for advanced STEM knowledge, then our future workforce may be unprepared to face the challenges of a world, which will be more globalized and reliant on digital technology than ever before. This paper aims to identify and address some shortcomings of current STEM education while investigating the effects of augmented reality and haptic technology on student engagement and learning. After exploring the results of research conducted in various subcategories of STEM, this paper concludes that augmented reality and haptics technology are effective at improving student engagement and learning while empowering valuable skills for high-level STEM work, including experimentation, collaboration, creativity, and critical thinking. Although there remain lingering obstacles with how such technology could be implemented in the classroom, such as the costs and challenges with traditional education models, this paper maintains an optimistic view on the potential to improve STEM education through augmented reality and haptics technology.

Introduction

1. The Societal Shift Towards STEM

Continuous technological innovation has ensured that behind every industry – from sports and entertainment to travel and commerce – there are countless technicians, engineers, and mathematicians (Barone, R., 2023). Additionally, as the world becomes increasingly interconnected through globalization, digital technologies are playing a huge role in empowering connections all across the globe (Youssef et al., 2019). The more that we rely on technological systems as a society and as a part of our economy, the more valuable and in-demand workers who have received a high-quality STEM education will be. According to the Bureau of Labor, the amount of STEM jobs is projected to increase by 10.8% by 2029, which is significantly larger than the average growth for a “stable” job, at 4.9% (Barone, R., 2023, para. 6). In contrast, according to the National Association of Manufacturing, the United States is on pace to have produced only 1.5 million “highly skilled candidates” by 2025, not even half of the projected 3.5 million STEM job openings (Barone, R., 2023, para. 1). Furthermore, a startling trend that is beginning to emerge is that of automation and technological takeover in the workforce; as technological capabilities, machine learning, and artificial intelligence systems continue to evolve, there becomes more and more pressure on the human workforce to demonstrate that their skills are invaluable and irreplaceable (Youssef et al., 2019). Based on the projection that

anywhere from 400 to 800 million jobs are at risk of becoming automated over the next 10 years (Youssef et al., 2019, para. 3), it is clear that to stand out in the workforce of tomorrow, prospects will need to possess valuable skills that are best acquired through STEM education, including cooperation, critical thinking, and creativity. Just as our society and its economy shift towards technology to run more efficiently, perhaps so too can our education systems – it is clear that we are in need of a large wave of high-quality STEM workers, large enough to fill the exponentially expanding technology market and skilled enough not to be replaced. To keep up with the constantly evolving need for technological experience, we must use the existing available technology to empower the highest quality STEM education possible.

2. Current Problems

2.1. Regressions in STEM interest and literacy

American society used to be the center for technological innovation and high-quality STEM education. Back in the 1950s, following the end of World War II and the beginning of the Cold War, the USA undertook enormous education programs, such as the National Defense Education Act (NDEA), to compete with the Soviet Union's technological progress and visions of being the first country to reach outer space (U.S. Senate, 2023). The program was a massive success and resulted in the flourishing of STEM education, along with the creation of NASA (U.S. Senate, 2023). STEM literacy and proficiency steadily grew until the early 2000s (Burke, A., 2022), but since then unfortunate trends in STEM literacy and interest have ensued. Beginning in the 2000s there has been a flatline in student performance in STEM (Burke, A., 2022). For example, the US has steadily fallen down the rankings of mathematics literacy in teens among Organization for Economic Co-operation and Development (OECD) nations, where it now sits at 25/37 - directly in front of Lithuania, and behind Spain (Burke, A., 2022). Furthermore, according to the MIT Technology Review, only about 28% of American adults are considered scientifically literate, meaning approximately 216 million U.S. adults are considered scientifically illiterate (Duncan, D. E. 2020). This is not to say that our programs are suddenly getting worse, but rather that they aren't expanding and improving at the rate we should be hoping for. Stagnation is not the worst-case scenario, but to keep up with society's increasing reliance on technology we're going to need to see significant progress that potentially strays away from traditional methods of STEM education.

2.2 The issues with traditional STEM education

STEM education in the majority of today's classrooms has not evolved very much over the past few decades (Petrov, P. D., 2020). Despite many aspects of STEM being highly complex, dynamic, and 3-dimensional, many classrooms still present these topics in a limited, two-dimensional view. For example, schools often opt to utilize the standard of lectures and textbook readings which have been around for ages instead of daring to test the waters with newer technology (Kesim et al., 2012). Presenting complicated topics in a manner that lacks any

immersion and practicality creates two primary issues. Firstly, there is simply less interest in whatever subject is being taught, as students can't engage with the content in a meaningful way (Sanfilippo, F., 2022). When complex STEM subjects are taught in a purely theoretical context, it can demotivate students from applying themselves, as they fail to realize the broader significance of the subject (Sanfilippo, F., 2022). For example, despite being a necessity in fields such as engineering and economics, calculus is often perceived by students as a rather useless thought experiment instead of a life skill that they should pay attention to (Wu et al, 2018). Secondly, even if students can recognize the importance of STEM, they may not be able to understand the material and its applications as thoroughly as their professors would have envisioned (Poonja et al., 2023). The implementation of skills that the student only learned about from a lecture or textbook reading is going to be extremely difficult, if not impossible, simply because the student was not taught in a way that prepares them for the real world. If a student took an engineering course that covered many of the necessary principles via lectures, readings, and slideshows, they still might find themselves lost when tasked with a more practical assignment that utilizes specialized technological tools – tools that they will have had zero previous experience with (Youssef & Zomor, 2019). Thus, it is crucial that our pedagogies shift to prioritize student interest and proficiency in STEM and ensure that we can keep up with the continual demand for STEM in the workforce.

3. Summary of Augmented Reality and Haptic Technology

This paper will focus on two forms of technology: augmented reality (AR) and haptics. AR is similar to virtual reality (VR) in that it imposes virtual information into the real world, projecting imagery for the user to see that isn't there (Kesim et al., 2012). However, while VR serves to create an entirely new immersive virtual world, AR only overlays projected imagery onto the surfaces or spaces of the real world (Kesim et al., 2012). Haptic technology is technology that serves to simulate the sensation of physical touch when there isn't any real object or force that would produce that sensation in real life (Poonja et al., 2023). For that reason, haptic technology is often paired with AR and VR to increase the degree of immersion, typically simulating features such as texture, vibration, temperature, and even physical force – like pressure (Youssef & Zomor, 2019). Both AR and haptics have a variety of associated devices that are used in various contexts, and here I will cover some of the most popular variations that are used in specifically an educational context. The most basic form of AR is simple apps that can be installed on handheld devices such as a smartphone or iPad (Pila et al., 2022). The more commonly used form of AR is the headset, or head-mounted display, in which a helmet-like device is worn to enable one or two optic displays to project virtual imagery directly into the subject's line of sight as if they were wearing magic glasses (Kesim et al., 2012).

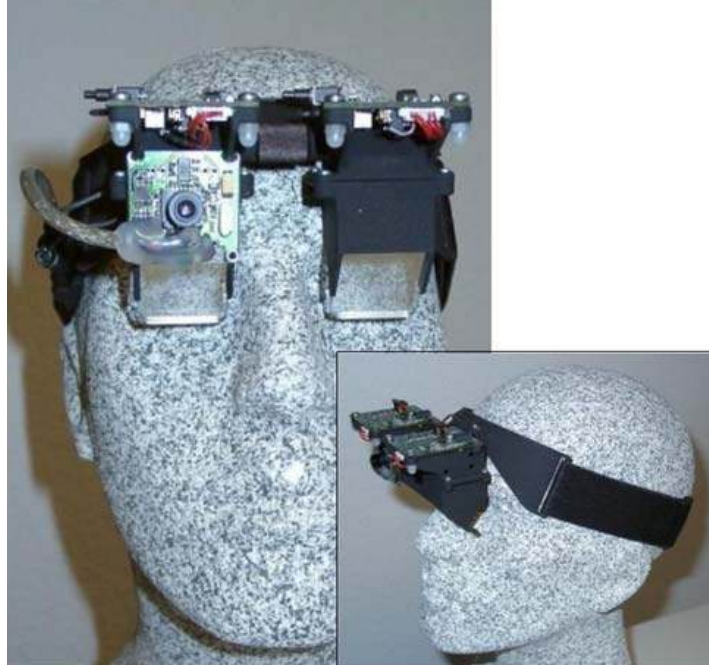


Figure 1. One example of a Head-Mounted Display Prototype from Trivisio Engineering

Source: Jorge, J. (2017). *Trivisio optical see-through HMD* [photograph]. ResearchGate. https://www.researchgate.net/figure/Trivisio-optical-see-through-HMD_fig3_221096725

The most basic example of haptic technology, similar to handheld AR devices, is just vibrations on a touch screen (Pila et al., 2022). An example of slightly more advanced haptic technology is the pinch glove, which is worn by the user like normal gloves, and detects gestures made by the hands to replicate those actions in the virtual landscape (Kesim et al., 2012).



Figure 2. Haptic Glove, with Tactile Pads on the Fingertips

Source: Hoang, T. (2013). The New Passive Deformable Haptic Glove [photograph]. ResearchGate.

https://www.researchgate.net/figure/The-new-Passive-Deformable-Haptic-Glove_fig1_26003466

For example, the gloves could detect if the user was making a grabbing motion, and then trigger the virtual object to become under the user's control until they released their hand (Kesim et al., 2012). Pinch gloves are not inherently haptic technology, as motion detection can be used on its own, but it is common to pair the gloves with some kind of sensor on the finger pads that can simulate tactile sensation (Hoang, T., 2013). Another typical example of haptic technology is a relatively simple controller, fitted with buttons and perhaps a joystick. The user will typically hold the controller while simultaneously using an AR headset to control their actions or manipulate an object in the simulation (Pila et al., 2019). These devices are customizable to the desired experience and thus come in many shapes and sizes, with various forms of haptic feedback (Sanfilippo, F., 2022). The most common forms of haptic feedback in these cases come in the form of vibrations, and pads on the controller which can simulate texture (Radu & Schneider, 2019). Overall, various forms of AR and haptic technology can have a range of designs, capabilities, structures, and prices to properly suit the needs of the user (Sanfilippo, F., 2022).

4. Purpose Statement

This paper aims to investigate the effect of AR and haptic technology on STEM education. It will focus on three separate capabilities of the technology, in an attempt to isolate which aspects of AR and haptic technology are more or less effective in increasing student engagement and student interest, respectively. Each aspect of the technology is viewed through the lens of two primary research questions:

1. Does AR and haptics technology improve student engagement and interest in STEM education?
2. Does AR reality and haptics technology enhance student learning and the development of skills in STEM education?

Literature Review

1. Imagery Projection

Student Engagement and Interest

Imagery projection is the most fundamental capability of AR technology, as I am defining it as the projection of virtual imagery into the real landscape, without necessarily being able to interact with the projected image. In terms of student engagement and interest, imagery projection seems to serve mostly as supplementary material which can increase students' overall excitement about the class, thus encouraging increased participation and meaningful

engagement. Additionally, if the results demonstrate an increase in student collaboration it can indicate improved engagement, as collaborating with other students often makes the material more enjoyable. It is worth noting that many metrics of student engagement and interest will be based on researcher observations and student surveys, as it can be difficult to define precise criteria upon which to evaluate engagement. One study conducted by Iulian Radu and Bertrand Schneider of Harvard University investigated the effect of augmented reality as an educational medium for teaching abstract physics concepts, such as magnetic fields and electromagnetic

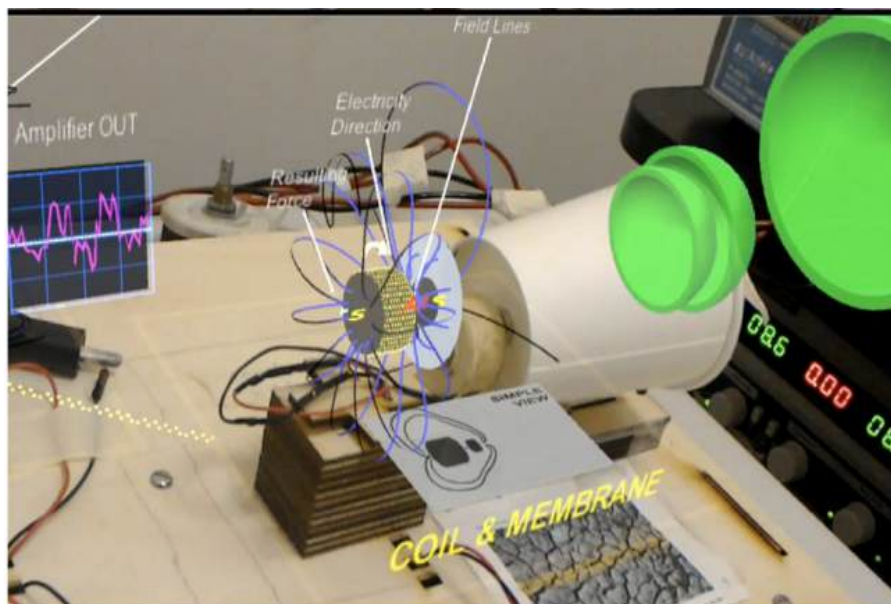


Figure 3. An AR Projection of an Electric Coil Producing a Magnetic Field

Source: Radu and Schneider. (2019). The magnetic fields around the coil and the magnet that are generating the sound waves. [photograph]. ACL.

<https://dl.acm.org/doi/pdf/10.1145/3290605.3300774>

waves. Students of all ages in the experimental group were given an augmented reality headset, allowing them to watch the magnetic field around an energy coil shift as the coil turned. After the demonstration, the researchers conducted an interview and test with both the students who had used AR, as well as the control group who had the topic explained normally by a teacher. It was reported that students in the experimental group “enjoyed the experience” more than those in the control group, and “felt it was more worthwhile” (Radu & Schneider, 2019, pg 7). Both the experimental and control group were reported to have worked collaboratively, although it is noted that the experimental group “had statistically higher ratings of time management, and worked more efficiently,” (Radu & Schneider, 2019, pg. 9). A separate study conducted by Omar Youssef and Mohammed Zomor of Florida International University involved undergraduate engineering students, who utilized a headset to visually represent sound production in various

rooms. Through an AR headset linked with microphones and decibel testers, students were able to observe noise intensity and direction, which they then used to try to design rooms that could better insulate sounds against urban environments.



Figure 4. Imagery Projections of Noise, Converted to an Infrared Scale

Source: Youssef & Zomor. (2019). This module provides a thorough understanding of methods of investigating human audial comfort indoors and outdoors using data acquisition systems. [photograph]. ASEE.

<https://peer.asee.org/coupling-haptic-learning-with-technology-to-advance-informal-stem-pedagogies>

Students were substantially immersed throughout the entire procedure and reported that it was “very unique and engaging,” (Youssef & Zomor, 2019, pg 10) although no such lesson was conducted before the use of AR to serve as a baseline for engagement. Overall interest in the class also increased, as despite two students dropping from the class in the semester prior to AR integration, only one student dropped in the following four semesters when AR was implemented. It is also noted that students were active in encouraging their peers to register for this “haptic learning experience” (Youssef & Zomor, 2019, pg. 10), demonstrating an increased willingness to engage with engineering concepts if they could also use AR technology as part of the procedures.

Student Learning and Skill Development

Imagery projection typically has a relatively minimal effect on skill development, because the concept of imagery projection as I am defining it in this paper implies being able to observe virtual images, without necessarily being able to practice any specialized skills besides critical thinking and making connections. By ‘skill development’ I am referring to specific skills, such as surgical procedures or technical drawings, that are practiced through AR and haptics to enhance student skills in real life. However, imagery projection still has a positive impact on

student learning and comprehension. In the aforementioned undergraduate sound engineering study, it was found that students who were able to engage with AR received significantly better grades in the course than those who previously did not. According to the study, “the letter grade ‘A’ increased from 47% of students to 85%, while the Letter grade ‘B’ decreased from 35% to 15% and the letter grade ‘C’ dropped from 17% to 0%” (Youssef & Zomor, 2019, pg 10). It is worth noting that the topics covered by the new curriculum (with AR implemented) had not been covered in prior semesters, which could have also been a factor in the increase in grades. In the physics study conducted by Radu and Schneider, the researchers had all students take a 21-question test after their respective learning experiences. The test covered seven topics related to electromagnetism and magnetic fields, and although it was reported that there was no significant difference in student performance between the two groups in five out of seven categories, the experimental group (which used AR) was reported to score more highly on questions regarding the shapes of magnetic fields and about the relationship between electricity and the magnetic field. The experimental group had a 64% better performance than the control group on questions about the shapes of magnetic fields, and 23% better than the control group on questions about the relationship between electricity and the magnetic field (Radu & Schneider, 2019, pg 8). For every other category, the difference in average test score was <10% (Radu & Schneider, 2019, pg 8).

2. Skill Simulation

Student Engagement and Interest

I am defining skill simulation as a procedure in which the students are directly practicing a real skill in a virtual environment that they will need to utilize in real life. For example, one study conducted by Umaiyalini Uruthiralingam and Paul M. Rea investigated the potential for AR to train medical students as part of their residency with procedures such as ventriculostomy, in which the surgeon must puncture the patient’s lumbar spine with a needle to drain cerebrospinal fluid. These simulations served to “provide a 3D image of the needle in the relative position” that would be necessary during a real surgery (Uruthiralingam & Rea, 2020, pg 7). Although simulations of these were primarily meant to serve as training for the procedure itself, it is worth noting that “student’s motivation and engagement was high” (Uruthiralingam & Rea, 2020, pg 7). The fact that intimidating and sometimes risky surgical procedures can be performed in a safe and controlled environment where failure and growth are allowed could make the entire field seem much more approachable, and such simulations could serve as great tools for improving student confidence and drawing interest towards the program. Another study conducted by Omar Huerta and others at the University of Huddersfield investigated the effect of AR simulations on student’s ability to execute technical drawings (TDs) - an extremely valuable skill possessed by any engineer - which allows designers to effectively communicate the technical product specifications to manufacturers. In this study college students used AR to view 3D animated walkthroughs demonstrating how to properly annotate and dimension various

models of parts of a car, in order to convey the most important features of a TD in a manner that is both effective and engaging. Although this study did not outline any specific engagement metrics or provide statistics, it is worth noting that the researchers reported AR simulations helping “the student with imagination and visualization whilst increasing their motivation and engagement,” (Huerta et al, 2019, pg 3).

Student Learning and Skill Development

In the study previously mentioned, in which students performed virtual ventriculostomy, it was found that the students demonstrated great improvement and proficiency in the given task itself, and also improved at some of the skills associated with the task. For example, it was reported that students who used AR to “provide a 3D image of the needle in the relative position performed quicker in insertion and procedure time,” (Uruthiralingam & Rea, 2020, pg 8). Additionally, many of the students experienced improvements in hand-eye coordination throughout the course of the study, and reported that by the end of the study they were all able to perform the procedure “calmly, efficiently, and correctly,” (Uruthiralingam & Rea, 2020, pg 8). One major beneficial factor is the fact that they were able to perform the same procedure as many times necessary for them to get the hang of it, and with progressive levels of difficulty to ease them into it. In the aforementioned TD study by Omar Huerta and others, the degree of success for the AR simulation was determined by the observations of the researchers and did not include any specific data points. Although it was stated that the “process of creating AR/VR content was time-consuming” (Huerta et al., 2019, pg 7), it was also observed that “Interactivity is essential” when it comes to creating simulations that translate into real skills (Huerta et al., 2019, pg 7). Overall, the researchers seemed to agree that the students who had the chance to learn from the interactive AR simulation demonstrated deeper understanding than those who were taught traditionally, and were able to more quickly utilize the demonstrated skills.

3. Tactile Feedback

Student Engagement and Interest

Although the full capability of haptics is technically broader than just tactile feedback, tactile feedback is the most key element in that the primary focus of haptic technology is to accompany some kind of AR/VR technology by adding in the sense of touch. So what exactly does that sense of touch add to the student’s learning experience; How big of a difference does it make? A study conducted by Civelek and others at Yildiz Technical University sought to investigate to what degree a haptic augmented simulation could potentially boost student attitudes towards physics, as well as achievement in physics. Throughout the study, 109 11th-grade students were taught about the gravitational forces between the earth and the sun by an AR application which displayed the earth rotating around the sun, accompanied by a haptic force feedback device that simulated the gravitational forces between the Earth and the sun.



Figure 5. Students in the Experimental Group Interacting with the Haptic Force Feedback Device

Source: Civelek, T. (2014). Students in the experimental group [photograph]. *Eurasia Journal of Mathematics, Science & Technology Education*.

<https://www.ejmste.com/article/effects-of-a-haptic-augmented-simulation-on-k-12-students-achievement-and-their-attitudes-towards-4340>

After the students had a chance to experiment with the haptic device, the researchers conducted a survey to gauge factors such as motivation, collaboration, and autonomy. The responses were given scores 1-5 depending on how strongly they indicated success in any respective category. On average, it was found that students who used the haptic device were 44% more motivated, 34% more collaborative, and 33% more autonomous (Civelek et al, 2014, pg 6). The way that haptics were able to assist in breaking down seemingly difficult material and instead make it seem simple serves to break stigmas and instill confidence in students to pursue challenging STEM subjects in the future. In another study conducted by Stephen Hegedus of the University of Massachusetts Dartmouth, it was investigated whether or not haptic technology could increase interest and proficiency in advanced mathematics, specifically geared towards younger children aged 9-10 years old. The study focused on geometric concepts that could be easily projected, and “developed an exploration activity using solids and a plane to explore how these objects interact” (Hegedus, 2013, pg 13). In addition, the geometric simulation was paired with the Omni haptic device, which allowed students to manipulate the geometric figures with the stylus.



Figure 6. The Omni Haptic Stylus Device

Source: Hegedus, S. (2013). “Phantom Omni” - 6 DOF master device. [photograph]. Delft Haptics Lab. <https://delfthapticslab.nl/device/phantom-omni>

This study did not outline any specific metrics for recording student engagement and interest, but the author did note that the students who used the device appeared to be “encouraged to want to continue to learn,” (Hegedus, 2013, pg 4). Furthermore, having never been exposed to some complex 3D shapes including various prisms and spheres, the students “engagement in discovering various types of intersection was immediate and endured for almost an hour,” (Hegedus, 2013, pg 14).

Student Learning and Skill Development

In the aforementioned physics study conducted by Civelek and others, students were asked to take an achievement test after completion of the haptic module. The test consisted of five short response questions regarding the fundamental laws at play behind the forces that they had experienced earlier. Student responses were scored on a scale from 1-10, and it was found that students in the experimental group scored on average 3.6 points higher than students in the control group. It was concluded that the additional level of immersion which haptics is an extremely helpful factor in creating a “fruitful learning environment for students... particularly in terms of improving student’s learning,” (Civelek et al., 2014, pg 9). Additionally, the mathematics study conducted by Martin Hegedus which investigated the effects of haptic technology on young students’ performance in geometry found that the younger students seemed to have a much easier time comprehending topics such as planar intersections, volume, density, and 3D angles when they were able to use the haptic stylus to interact with the projected

imagery. Hegedus notes on page 10 that the experimental group was able to comprehend the subjects “much deeper and very quickly.” As soon as the subject matter becomes less theoretical and more interactable, students are not only able to digest content much more efficiently, but are also exposed to experimental skills which are crucial to the pursuit of knowledge in any STEM-related subject.

Findings

Overall, it is reasonable to conclude that AR in combination with haptics has a positive effect on both student engagement as well as student performance and achievement. I suspect that at the core of the effectiveness of the technology for empowering continued interest and achievement in STEM are skills which aren't specific to any field or job title, but instead are keys to finding success throughout all of STEM, at any level. Firstly, by creating an environment in which content becomes more visibly digestible, students are motivated to collaborate in order to solve problems and reach a collective understanding. Being unafraid to ask for help and willing to combine skills in order to overcome a problem which seems initially daunting is a trait that is invaluable across all fields, and it is thus a trait which must be emphasized in STEM education. Secondly, being able to simulate real scenarios and complex theoretical concepts allows students to exercise their critical thinking skills. Simulations can push them to make connections and realizations that could have otherwise been left unturned, if all they had access to was a textbook and lecture notes. Lastly, with visible imagery being such a powerful tool, students can empower their creativity by creating models, analyzing anatomy, and visualizing geometry right in front of their eyes without any external resources. This means that their minds are prompted to expand on what they see, and push their creative limits to form unique theories and designs that they may have never reached without AR. These key skills are largely recognized and strived for throughout traditional STEM education, but many students end up falling short without the aspects of interactivity and visual simulation provided by AR and haptics (Reich, 2020). One of the likely driving factors behind the effectiveness of AR and haptics, in addition to the factors already mentioned, is the effect on student's cognitive load. According to Shayan Doroudi, cognitive load theory states that any learning material should aim to minimize cognitive load, the amount of information that working memory can hold at one time, in order to maximize the student's ability to use their working memory to comprehend the information being presented (Doroudi, 2021). By presenting the information in a format which takes minimal effort to comprehend, AR is able to effectively minimize cognitive load to empower student understanding. This is similarly part of the issue with traditional STEM education - boring lectures and textbook readings require more mental energy to pay attention, and are thus suboptimal in comparison to AR and haptics. At the beginning of this paper I mentioned the threat of automation in the job industry, and how valuable STEM education is going to be at producing irreplaceable workers; the skills described above which AR and haptics encourage are exactly the reason why. Machines are great at routine tasks, but traits like critical

thinking, collaboration and creativity are all uniquely human qualities – qualities that could serve as valuable job security some time in the near future.

Implications

Though I think that AR and haptics have undeniable potential to make STEM education a more enriching experience for all students, it is important to consider that the technology is not a panacea, and there are a number of potential obstacles that stand in the way of properly implementing this technology into classrooms.

1. Cost Efficiency

One unfortunate aspect of all the advanced technology being used is that like most advanced technology, it generally is not cheap (Kesim et al., 2012). Almost all of the studies which I reviewed featured a small amount of AR/haptic devices which students took turns with, not providing unique headsets to each student. If the hope is to have these technologies blended into classrooms, school districts may have to make a decision between quality and quantity. Schools aren't necessarily forced to invest in potentially expensive headsets, as mere apps on iPads can be serviceable (Pila et al., 2022). However, as can be expected, the effect on student engagement and achievement is less pronounced, and the difference between students with and without the technology is nearly indistinguishable (Pila et al., 2022). It is worth noting that many of the studies reportedly used lower-cost options that should be theoretically accessible to school PTAs and boards, just to demonstrate that high-quality results are still possible on a budget. Unfortunately, there are still additional costs beyond just the purchasing of the hardware itself. Oftentimes specialized experts are required in order to properly install and set up these technologies, and although this could potentially be taken care of by a school employee, some may not feel comfortable with the risk of damaging any of the technology (Kesim et al., 2012). Furthermore, many of the applications that are used are associated with licensing fees and software subscriptions that the school would need to be willing to cover (Kesim et al., 2012).

2. Cybersickness

Cybersickness is a phenomenon similar to motion sickness that is associated with VR/AR and can result in the user feeling symptoms of nausea, discomfort, sweating, and dizziness (Kemeny, A., Chardonnet, J.-R., & Colombet, F., 2020). Latency and lag within the simulation, unnatural or sudden movements, and abrupt changes in perspective are all potential triggers (Kemeny, A., Chardonnet, J.-R., & Colombet, F., 2020). For any users who are susceptible to motion sickness, cybersickness could potentially stand in the way of extended session time, severely limiting the duration of virtual experiences. Fortunately, this is more of a concern for VR than AR, because having the real world as a backdrop in AR can help the brain stabilize itself in a way, as opposed to the total immersion of VR (Kemeny, A., Chardonnet, J.-R., & Colombet, F., 2020). Additionally, some newer models of VR and AR technology are implementing audial simulations to accompany the visual effects, which has been found to help

reduce the severity of cybersickness in some users (Kemeny, A., Chardonnet, J.-R., & Colombet, F, 2020). Ideally, as the technology continues to advance, the less cybersickness will be an issue.

3. Challenges with Traditional Pedagogy

Although we'd like to imagine that all teachers just want what's best for their students, the fact remains that there are a number of potentially limiting factors that could limit or prevent the proper implementation of AR and haptics into the classroom. The first issue is familiarity, as most teachers will not have dealt with these forms of technology before (Reich, 2020). Especially with older teachers who have methods of teaching that have remained separate from technology, many may find themselves unsure of how to operate some of the devices, or confused at how to include them in their existing lesson plans (Reich, 2020). This factor alone could potentially warrant a school-wide technological training program, which would cost money, tying back to the issue of cost efficiency. Additionally, despite the research that suggests the effectiveness of AR and haptics on a number of factors relating to student engagement and achievement, it is not unreasonable to assume that many teachers are going to be skeptical (Reich, 2020). If they have been teaching for a long time, they likely have some methodology which they view as tried and true and may be concerned with the pedagogical value that the technology brings to the table – “If they don't trust the tech itself, why would they want to go through all the effort of implementing it?” (Reich, 2020, pg. 135).

Conclusion

Although all of the issues that I have acknowledged are very real and potentially very disruptive, I truly believe that the potential of AR and haptic technology to renew interest in STEM and produce a generation of students with strong STEM proficiency warrants at least an attempt to implement it into the school system. In order to mitigate some of the risks, it will likely have to unfold slowly, beginning with a handful of volunteer schools that are willing to experiment with the technology for a year or two. Though it may start slowly, I am confident that the effect it will have on the minds of young STEM scholars will make it worth it in the end, no matter how long it takes.

Works Cited

- Kesim, M., & Ozarslan, Y. (2012, August 18). Augmented reality in education: Current Technologies and the potential for education. *Procedia - Social and Behavioral Sciences*. https://www.sciencedirect.com/science/article/pii/S1877042812023907?ref=pdf_download&fr=RR-2&rr=7f07b745cc4e4315
- Uruthiralingam, U., & Rea, P. M. (2020). Augmented and Virtual Reality in Anatomical Education – A Systematic Review. *www.sciencedirect.com*. https://link.springer.com/content/pdf/10.1007/978-3-030-37639-0_5.pdf
- Poonja, H. A., Khan, M. J., & Javed, K. (2023). Engagement detection and enhancement for STEM education through computer vision, augmented reality, and haptics. *Science Direct*. <https://www.sciencedirect.com/science/article/pii/S026288562300094X>
- Petrov, P. D., & Atanasova, T. V. (2020). The Effect of Augmented Reality on Students' Learning Performance in Stem Education. *information*. https://mdpi-res.com/information/information-11-00209/article_deploy/information-11-00209.pdf?version=1586939814
- Sanfilippo, F. (2022). A perspective review on integrating VR/Ar with haptics into STEM Education for Multi-Sensory Learning. *robotics*. https://mdpi-res.com/d_attachment/robotics/robotics-11-00041/article_deploy/robotics-11-00041.pdf?version=1648737764
- Pila, S., Lauricella, A. R., Piper, A. M., & Wartella, E. (2022, July 19). Preschool teachers' perspectives on (haptic) technology in the classroom. *Frontiers*. <https://www.frontiersin.org/articles/10.3389/feduc.2022.981935/full>
- Youssef, O., & El Zomor, M. (2019). Coupling Haptic Learning with Technology To Advance Informal STEM Pedagogies. *ASEE*. <https://peer.asee.org/coupling-haptic-learning-with-technology-to-advance-informal-stem-pedagogies>
- Radu, I., & Schneider, B. (2019). What can we learn from augmented reality (AR)? Benefits and Drawbacks of AR for Inquiry-based Learning of Physics. *ACL*. <https://dl.acm.org/doi/pdf/10.1145/3290605.3300774>
- McDonald, C. V. (2016). STEM Education: A review of the contribution of the disciplines of science, technology, engineering and mathematics. *ERIC*. <https://files.eric.ed.gov/fulltext/EJ1131146.pdf>
- National Science Foundation. (2022). The state of U.S. Science and Engineering 2022. National Science Board. <https://nces.nsf.gov/pubs/nsb20221/u-s-and-global-stem-education-and-labor-force>
- Alison Gillespie, Communications Specialist. (2021, September 9). What do the data say about the current state of K-12 STEM education in the US?. *NSF*. <https://new.nsf.gov/science-matters/what-do-data-say-about-current-state-k-12-stem>
- Burke, A. (2022). Science & Engineering Indicators. National Science Foundation. <https://nces.nsf.gov/pubs/nsb20211>

- Hackling, M. W. (2015). Preparing today's children for the workplaces of tomorrow: The critical role of STEM education. *International Journal of Innovation in Science and Mathematics Education*. file:///Users/jjdesign23/Downloads/10334-85-26820-3-10-20160225.pdf
- Kemeny, A., Chardonnet, J.-R., & Colombet, F. (2020). Getting Rid of Cybersickness In Virtual Reality, Augmented Reality, and Simulators. <https://link-springer-com.ezproxy.oberlin.edu/content/pdf/10.1007/978-3-030-59342-1>
- Fang, H., Guo, J., & Wu, H. (2022). Wearable triboelectric devices for haptic perception and VR/AR applications . *Science Direct*. <https://www-sciencedirect-com.ezproxy.oberlin.edu/science/article/pii/S2211285522001938?via%3Dihub>
- Blažauskas, T., Gionata, S., Ramos, I., Vert, S., Radianti, J., Majchrzak, T. A., & Oliveira, D. (2022, April 24). Integrating VR/ar with haptics into STEM Education. *Encyclopedia*. <https://encyclopedia.pub/entry/22189>
- Fouad, M., Mansour, T., & Nabil, T. (2017). Use of haptic devices in education: A review . *SCEEE*. https://sceee.journals.ekb.eg/article_279487_afcd3241a051c4ddd3e30f2ba3f38d8d.pdf
- Hegedus , S. (2013). Young children investigating advanced mathematical concepts with haptic ... *The Mathematics Enthusiast*. <https://scholarworks.umt.edu/cgi/viewcontent.cgi?article=1261&context=tme>
- Civelek, T., Ucar, E., Ustunel, H., & Aydin, M. K. (2014, December 15). Effects of a haptic augmented simulation on K-12 students' achievement and their attitudes towards physics. *Eurasia Journal of Mathematics, Science and Technology Education*. <https://www.ejmste.com/article/effects-of-a-haptic-augmented-simulation-on-k-12-students-achievement-and-their-attitudes-towards-4340>
- Cardoso, O. I. H., Kus, A., Unver, E., Aslan, P. R., Dawood, M., Kofoglu, M., & Ivanov, V. (2019, August 27). A design-based approach to enhancing technical drawing skills in design and engineering education using VR and AR Tools. *University of Huddersfield Research Portal*. <https://pure.hud.ac.uk/en/publications/a-design-based-approach-to-enhancing-technical-drawing-skills-in->
- Reid, S., Shapiro, L., & Louw, G. (2019). How haptics and drawing enhance the learning of anatomy. *Anatomy Pubs*. <https://anatomypubs.onlinelibrary.wiley.com/doi/10.1002/ase.1807>
- Bölek, K. A., De Jong, G., & Henssen, D. (2021). The effectiveness of the use of augmented reality in anatomy ... - *nature*. *Scientific Reports*. <https://www.nature.com/articles/s41598-021-94721-4.pdf>
- International Student Assessment (PISA) - mathematics performance (PISA) - OECD data. theOECD. (2022). <https://data.oecd.org/pisa/mathematics-performance-pisa.htm>
- Barone, R. (2023). Stem education stats for 2023: Facts on Jobs & Careers, Shortage & Minorities. *iD Tech*. <https://www.idtech.com/blog/stem-education-statistics>

- Aglion, J. (2019, November 25). Where the stem workforce is headed and what society must do to get there. Getting Smart.
<https://www.gettingsmart.com/2019/11/25/where-the-stem-workforce-is-headed-and-what-society-must-do-to-get-there>
- Hoang, T. (2013). The new passive deformable haptic glove - researchgate. ResearchGate.
https://www.researchgate.net/figure/The-new-Passive-Deformable-Haptic-Glove_fig1_260034666
- Jorge, J. (2017). Trivisio optical see-through HMD | Download Scientific ... - researchgate. ResearchGate.
https://www.researchgate.net/figure/Trivisio-optical-see-through-HMD_fig3_221096725
- Duncan, D. E. (2020, April 2). 216 million Americans are scientifically illiterate (part I). MIT Technology Review.
<https://www.technologyreview.com/2007/02/21/37898/216-million-americans-are-scientifically-illiterate-part-i/>
- Sputnik spurs passage of the National Defense Education Act. U.S. Senate: Sputnik Spurs Passage of the National Defense Education Act. (2023, August 8).
https://www.senate.gov/artandhistory/history/minute/Sputnik_Spurs_Passage_of_National_Defense_Education_Act.htm
- Wu, X., Deshler, J., & Fuller, E. (2018, October 25). The effects of different versions of a gateway stem course on student attitudes and beliefs - international journal of STEM education. SpringerOpen.
<https://stemeducationjournal.springeropen.com/articles/10.1186/s40594-018-0141-4>
- Reich, J. (2020). Failure to disrupt: Why technology alone can't transform education. Harvard University Press.
- Doroudi, S. (2021). Shayan Doroudi - learning theories primer. Shayan Doroudi - Learning Theories Primer. <https://sites.google.com/uci.edu/shayan-doroudi/blog/learning-theories>

The Economic Aspects of Schizophrenia Treatments By Abdul Azis Ramzi

1. Introduction

Within the complexities of schizophrenia treatment, understanding its economic implications is crucial for ensuring optimal outcomes for patients. The World Health Organization states that “Schizophrenia is characterised by significant impairments in the way reality is perceived and changes in behaviour” (WHO). This profound impact on individuals diagnosed with schizophrenia necessitates a comprehensive exploration not only of the clinical aspects but also the economic effects shaping the treatment. Economic factors such as income, access to healthcare, and the cost of treatment play a significant role in shaping the course of the illness and the effectiveness of interventions. There is clear and compelling evidence that mental institutions are not effective in providing the care needed for individuals with mental health conditions, including schizophrenia, and, unfortunately, often violate the basic human rights of individuals struggling with this disorder (WHO). This emphasizes the urgent need to consider alternative and more effective approaches within the broader context of schizophrenia treatment, which includes addressing economic barriers and ensuring access to appropriate care that prioritizes the fundamental rights of individuals with this condition.

2. The Impact of Socioeconomic Factors on Patients diagnosed with Schizophrenia

Socioeconomic factors play a significant role in shaping the experiences and outcomes of individuals diagnosed with schizophrenia. According to Dr. Jean Addington a Schizophrenia researcher inducted into the Royal Society of Canada for more than 30 years, and her associates, "For individuals with schizophrenia, recovery from psychotic symptoms is common after the first episode, with 75% to 90% achieving remission from the positive psychotic symptoms--- that is, hallucinations, thought disorder, and delusions--- 1 year after treatment” (Addington et al.). This statistic emphasizes the importance of early intervention and access to adequate treatment resources in facilitating positive outcomes for patients. Recovery from psychotic symptoms is common with early intervention, yet a persistent correlation between poverty and schizophrenia diagnosis exists. This shows the significance of early interventions assets in accomplishing positive results for patients. Nonetheless, the effect of financial variables can't be disregarded, as people with economic disadvantages might confront significant boundaries in getting to ideal and quality treatment, possibly fueling their condition. John Read, a professor of clinical psychology at the University of East London's School of Psychology, states in the New Zealand Journal of Psychology that "For seventy years, however, research has repeatedly demonstrated not only that poverty is a powerful predictor of who develops psychosis, and who is diagnosed 'schizophrenic' (with or without a family history of psychosis), but that poverty is more strongly related to 'schizophrenia' than to other mental health problems” (Read). This link suggests a complex interplay between socioeconomic status and mental health, where poverty serves as a significant risk factor for the emergence and exacerbation of schizophrenia symptoms. Individuals from lower socioeconomic backgrounds experience a heightened burden of stress compared to their

wealthier counterparts (Read). This stress extends beyond financial constraints, encompassing challenges such as hunger, homelessness, powerlessness, isolation, and a lack of self-respect. Upon entering the psychiatric system, individuals from disadvantaged backgrounds are disproportionately more likely to be hospitalized and labeled as 'schizophrenic', irrespective of their behavior (Read). This labeling not only exacerbates their already diminished self-esteem and motivation but also distances them from their loved ones. The unjust labeling and increased hospitalization of individuals from disadvantaged backgrounds perpetuate inequities within the psychiatric system. This not only exacerbates the challenges faced by these individuals but also emphasizes the urgent need for reforms that promote unbiased and equitable mental health care practices. Addressing these disparities is essential for a more inclusive approach to schizophrenia diagnosis and treatment.

3. The impact of Health Insurance and the Pharmaceutical Industry on Schizophrenia Management

The role of health insurance in shaping access to essential interventions and support for individuals diagnosed with schizophrenia is pivotal for their overall well-being. According to a peer-reviewed research conducted by Huey Yi Chong, an MSc in International Pharmacoeconomics and Health Economics and has Ph.D. in Health Economics, and her associates, "The enormous economic burden in schizophrenia is suggestive of the inadequate provision of health care services to these patients" (Chong et al.) This suggests a need for a comprehensive reevaluation of the healthcare services available to individuals with schizophrenia, emphasizing the integral role that health insurance plays in shaping their access to necessary interventions and support. Howard H. Goldman, a professor at the Department of Psychiatry at the University of Maryland, states that "Low levels of employment for people with schizophrenia will no longer create a special barrier to health insurance traditionally obtained as a workplace benefit" (Goldman). While Dr. Chong highlights the economic burden signaling inadequate healthcare provision for those with schizophrenia, Goldman's observation points to a positive shift, wherein low employment levels no longer present a unique obstacle to obtaining health insurance. This shift allows individuals with a schizophrenia diagnosis to benefit from a broader range of public and private plans, ensuring access to health insurance either through employment or alternative means. Notably, those with limited financial means will have access to subsidies, facilitating payment for individual insurance or qualifying for Medicaid if workplace coverage is unavailable (Goldman). This newfound accessibility is particularly significant given the impoverishing effect of schizophrenia, providing affected individuals with the means to access early intervention services before the condition becomes disabling. However, a study conducted by Dr. Elizabeth Khaykin, a professor in the Department of Mental Health at Johns Hopkins Bloomberg School of Public Health, and her associates, introduces a contrasting perspective, stating that "Almost all U.S. adults with schizophrenia were found to receive government health insurance, yet many remained uninsured" (Khaykin et al.). Recent shifts in employment-related barriers to health insurance provide improved access for individuals

diagnosed with schizophrenia. However, challenges persist, as a significant proportion of adults with schizophrenia remain uninsured.

The impact of the pharmaceutical industry on schizophrenia management is a critical aspect that shapes the availability and effectiveness of treatment options for individuals living with this challenging mental health condition. According to research conducted by Joanna P. MacEwan, a Ph.D. Research Economist at Precision Health Economics and an expert in advanced econometric analyses, and her associates, “The level of investment and pharmaceutical innovation in mental disorders was comparatively low, especially relative to the burden of disease” (Joanana et al.). This disparity between the burden of schizophrenia and the level of investment in pharmaceutical innovation emphasizes a gap that may hinder the development of more effective treatment approaches for patients. Furthermore, the Food and Drug Administration (FDA) adds to these concerns, noting that “several large pharmaceutical companies appear to have scaled back their developmental efforts regarding schizophrenia, despite its toll on patients and their families” (FDA). This disparity in investment and innovation relative to the significant burden of schizophrenia underscores the pressing need for increased attention and resources directed towards developing more effective treatment approaches. Addressing these concerns is essential to ensure that individuals with schizophrenia have access to advancements in pharmaceutical research that align with the severity of their condition.

4. Conclusion

In conclusion, acknowledging the economic implications of schizophrenia treatment is paramount in ensuring the best possible outcomes for patients. Socioeconomic factors significantly influence access to care and treatment outcomes, with individuals from lower socioeconomic backgrounds facing increased barriers. While recent health insurance reforms have aimed to improve access to care for individuals with schizophrenia, gaps in coverage still exist, particularly among uninsured individuals who remain vulnerable to inadequate treatment options. Additionally, the pharmaceutical industry's comparatively low investment in schizophrenia treatment presents significant challenges in developing innovative and effective therapies for this complex condition. Therefore, addressing these economic challenges through comprehensive reforms and increased investment in research and development is essential for enhancing the overall quality of care and outcomes for individuals living with schizophrenia.

Works Cited

- Addington, Jean, et al. "Psychosocial Treatments for Schizophrenia." *Current Directions in Psychological Science*, vol. 19, no. 4, 2010, pp. 260-63. JSTOR, www.jstor.org/stable/41038581. Accessed 10 Jan. 2024.
- Chong, Huey Yi, et al. "Global Economic Burden of Schizophrenia: A Systematic Review." *Neuropsychiatric Disease and Treatment*, vol. 12, 2016, p. 357+. Gale Academic OneFile, link.gale.com/apps/doc/A506651222/AONE?u=nysl_we_wnhs&sid=bookmark-AONE&xid=4cb42ce1. Accessed 10 Jan. 2024.
- Goldman, Howard H. "Will Health Insurance Reform in the United States Help People With Schizophrenia?" OXFORD ACADEMIC, 21 July 2010, academic.oup.com/schizophreniabulletin/article/36/5/893/1870571. Accessed 28 Jan. 2024.
- "Improving the Design of Clinical Trials of Drugs to Treat Schizophrenia." FDA, 10 Nov. 2020, www.fda.gov/drugs/regulatory-science-action/improving-design-clinical-trials-drugs-treat-schizophrenia. Accessed 28 Jan. 2024.
- Khaykin, Elizabeth, et al. "Health Insurance Coverage Among Persons With Schizophrenia in the United States." National Library of Medicine, Aug. 2010, www.ncbi.nlm.nih.gov/pmc/articles/PMC7245045/. Accessed 28 Jan. 2024.
- MacEwan, Joanna P., et al. "Pharmaceutical Innovation in the Treatment of Schizophrenia and Mental Disorders Compared with Other Diseases." National Institutes of Health, Aug. 2016, www.ncbi.nlm.nih.gov/pmc/articles/PMC5022985/. Accessed 28 Jan. 2024.
- Read, John. "Can Poverty Drive You Mad? 'Schizophrenia', Socio-economic Status and the Case for Primary Prevention." *New Zealand Journal of Psychology*, vol. 39, no. 2, July 2010, p. 7+. Gale Academic OneFile, link.gale.com/apps/doc/A243877211/AONE?u=nysl_we_wnhs&sid=bookmark-AONE&xid=4c2242b1. Accessed 10 Jan. 2024.
- "Schizophrenia." World Health Organization, 10 Jan. 2022, www.who.int/news-room/fact-sheets/detail/schizophrenia. Accessed 28 Jan. 2024.

Diagnosing and Treating Cancer: An Inspection to Long Non-Coding RNAs and Immunology Therapy By Zehra Desen Çelik & Eden Ng

Abstract

Cancer is referred to as the disease of mutant cells multiplying exponentially, which triggers the formation of malignant tumors. These cells may detach from their primary location in a process called “metastasis,” causing them to circulate in blood and become spread throughout the body. Due to its asymptomatic nature, it is challenging to detect the presence of cancer. However, recent studies have come up with promising methods to diagnose cancer. One of these methods is the use of long non-coding RNA (lncRNA) to detect cancer at an early stage. After diagnosing cancer, treatment stands to be another aspect, as the disease builds resistance to many of the typical treatments. A novel finding shows the advance in cancer treatment using T-cell therapy, found most successful in blood cancers. This review will focus on the role of long non-coding RNAs in cancer progression along with the factors used to create and maintain ideal conditions for T-cell differentiation to occur, potentially aiding efforts on cancer immunotherapy.

Keywords

Cellular and Molecular Biology, Cellular Immunology, Microbial Genetics, Molecular Biology

Introduction

Cancer is the name for diseases that arise due to the abnormal and uncontrolled division of the cells of an organism (“NCI Dictionary of Cancer Terms”). This uncontrolled progression gradually targets healthy cells and spreads all around the body due to cancer’s ability to conduct metabolic activities. Cancer metabolism is adapted in a unique way that allows elevated energy to be obtained in nutrient and oxygen-deficient conditions, often referred to as the Warburg effect, and favors cancerous cells’ malignancy (Chae and Hong, 2022). It has been noted that several non-coding RNAs have an important role in the proliferation of cells as these RNAs impact glycolytic enzyme activities. A special type of RNA named long non-coding RNA (lncRNA for short), consisting of around 200 nucleotides, has been found to have regulatory, catalytic, and structural purposes that facilitate the process of tumorigenesis by cell proliferation (Wang et al., 2020). In 1961, when Jacob and Monod found the existence of mRNA and came up with a repressor-operator model for the regulation of genes, they predicted that the regulatory molecule was also another type of RNA (Kung et al., 2013). In the early 1990s, the first lncRNAs responsible for regulatory functions in distinct genes were discovered. As it became clear that the regulatory purposes were aligned with increased cell proliferation and tumorigenesis, more studies were conducted on lncRNAs to reveal that these non-coding biomolecules influence growth, metastasis, and tumor angiogenesis (Gao et al., 2020). This review will focus on

It has been noted that several non-coding RNAs have an important role in the proliferation of cells as these RNAs impact glycolytic enzyme activities. A special type of RNA

named long non-coding RNA, consisting of around 200 nucleotides, has been found to have regulatory, catalytic, and structural purposes that facilitate the process of

T-cells are a major part of the adaptive immune system. Distributed by the lymphatic system, T-cells are lymphocytes that defend the body against infection and fight cancer. These cells originate in the bone marrow, where they are stem cells, and undergo cell differentiation to become specialized or mature. The process of cell differentiation occurs when an overproduction of a certain protein induces chromatin remodeling—changing of cell chromosomes—which changes the cell's mRNA translation and the types of proteins produced and changing a cell's gene expression. Regular T cells have a certain amount of T cell receptors (TCR)—each binds to a specific type of antigen (from pathogens that activates the immune system)—and a level of antigen specificity—how T cells can identify multiple kinds of antigens. Though T-cells can be differentiated into two groups of effector cell types, the majority of T-cells work in a similar approach by responding to antigen-presenting cells to defend the body from foreign virus/bacteria or pathogens. In this review, we will summarize recent efforts and technological developments to overcome challenges to control T-cell differentiation by modifying the microenvironment, controlling self-reactivity, and limiting certain transcription factor expressions.

Non-coding RNA's Role in Cancer: lncRNAs and Cancer

Research carried out in the past decades indicates that mutations occur in the genome which impact the non-coding regions, resulting in nucleotide expansions and single nucleotide polymorphisms (SNP) (Yang et al., 2014). This suggests that the regions responsible for lncRNA could be altered, causing systems of the body to be regulated in a malfunctioning manner. Therefore, mutations in regions responsible for lncRNA synthesis led scientists to hypothesize that cell proliferation and tumorigenesis are potentially affected by the information located on non-coding genes (Yang et al., 2014). Along with this hypothesis, it is well established that lncRNAs act in different mechanisms, such as controlling mRNA splicing and gene expression (Ali and Grote, 2020). One of the mechanisms that lncRNAs use in cancer is binding to RNA proteins, which causes the transcribed mRNA to be destabilized in certain cases (Ahmad et al., 2023).

Subsequently, lncRNAs regulate cancer metabolism as they have the potential to modulate glycolytic enzyme activities. One of the reasons why cancer cells consume energy quickly and maintain their lives is the Warburg effect: cancer cells generate immediate energy by preventing the regular Citric Acid Cycle and promoting rapid-energy processes such as glycolysis (Kang et al., 2023). Hence, the impact of lncRNA on cancer metabolism and its sustainability is analyzed through in-vitro and in-vivo analysis of lncRNAs that are overexpressed in cancerous tissue.

lncRNA in Hepatocellular Carcinoma

“Highly upregulated in liver cancer” (HULC) is an lncRNA that is overly expressed in the hepatocellular carcinoma tissues. The studies of HULC indicate that it increases cell proliferation and causes tumor angiogenesis through its bindings and interactions with protein molecules. As it was wondered whether lncRNAs could be used as biomarkers of cancer, Panzitt and colleagues detected HULC in the blood of HCC patients as the most upregulated gene. The team generated a gene library to screen the deregulated genes and used techniques such as quantitative RT-PCR and siRNA-mediated knockdown to discover HULC as the most expressed one (Panzitt et al., 2007).

Wang and colleagues carried out several laboratory techniques to investigate HULC as an lncRNA. The team’s use of STRING Protein-Protein Interaction analysis suggested that HULC may take part in numerous functions, including metabolism, response to viruses, and transcriptional regulation (Wang et al., 2020). Further research has demonstrated that HULC regulates the “Fibroblast growth factor receptor type 1” (FGFR1) mediated phosphorylation of two glycolytic enzymes, lactate dehydrogenase A (LDHA) and pyruvate kinase 2 (PKM2). FGFR1 is a membrane protein that catalyzes the phosphorylation of aforementioned glycolytic enzymes and HULC’s interactions with FGFR1 show the potential of HULC to act as an adaptor molecule (Wang et al., 2020). This indicates that HULC presence promotes the Warburg effect and creates habitable conditions for cancerous cells.

To determine HULC’s ability to bind to glycolytic enzymes directly, an RNA pull-down assay using biotinylated HULC and antisense HULC was carried out. Then, western blot analysis and in-vitro His-Tag pull-down assay confirmed that there is a direct binding between HULC and the two glycolytic enzymes, LDHA and PKM2 (Wang et al., 2020). To investigate the phosphorylation of the enzymes and whether HULC played a role, the HULC gene was knocked down and the phosphorylation levels were analyzed. For LDHA, it was observed that the knockdown reduced Y10 phosphorylation, and the overexpression enhanced it while for PKM2, the knockdown reduced Y105 phosphorylation and upregulation enhanced it (Wang et al., 2020). The results show that HULC promotes LDHA activity and phosphorylation through its interactions, and HULC downregulates PKM2 activity since the phosphorylation of Y105 inhibits the tetramer structure of the enzyme (Wang et al., 2015).

Knowing that FGFR1 regulated both enzymes’ phosphorylation and being overexpressed in cancerous cells, it was hypothesized that HULC regulated the FGFR1-mediated phosphorylation. Western blotting indicated a growth in the membrane localizations of the glycolytic enzymes when HULC was overexpressed, and Cell counting kit-8 demonstrated an increase in proliferation (Wang et al., 2015). Thus, the team was able to note that HULC increased LDHA and PKM2 activities by playing a role in FGFR1-mediated phosphorylation.

The Microenvironment and Environmental Sensors

The Microenvironment and Environmental Sensors, in emerging research, has been found to influence T cells’ response through environmental sensing transcription factors like

nutrient-sensing serine/threonine kinases, adenosine monophosphate-activated protein kinase and mammalian target of rapamycin complex 1 that directly activate T-cell differentiation. For example, the degradation of HIF, a metabolic sensor responding to decrease in oxygen levels in the cellular environment, was found to limit the differentiation of CD8 T-cells, which displays that the oxygen tension in the cellular environment borders T-cell differentiation. T-cell differentiation has also been found to be controlled by the bacterial environment because it regulates T-cell responses through pattern recognition receptors provoked by bacterial products, which triggers the immune system and thereby controls the response. Therefore, T-cells are able to manipulate their response by limiting/adding to the expression of nutrient receptors and environmental sensing transcription factors and, by using these forms of treatment/therapy, researchers will be able to train T-cells to respond to certain pathogens/antigens. (Ramsay and Cantrell, 2015). This knowledge can be applied to mass T-cell differentiation as these environmental sensing transcription factors can be implemented to prompt T-cell differentiation, as well as modifying the cell's behavior by modifying nutrient receptors to restrict/allow interpretation of the bacterial environment.

Developmental Self-Reactivity

Developmental Self-Reactivity, a cell's tendency to mistakenly produce antibodies that attack and destroy the body, is a contributing factor in T-cell differentiation, as its harmful nature discourages differentiation. When separating naive CD8⁺ T (TN) cells into 3 distinct subsets: CD5^{lo}Ly6C⁻, CD5^{hi}Ly6C⁻, and CD5^{hi}Ly6C⁺ by levels of self-reactivity, researchers questioned which subset would most likely differentiate into TC17 cells based on its reaction to inflammatory disease through several experiments in murine models. By using Rag1^{-/-} mice induced with IBD, the researchers scored the subsets' histological score and measured the length of the colon after 14 days to show which subset is most self-reactive, then evaluated the amount of cytokines, typically released from TC17 cells, each subset released, which shows how much of each subset was differentiated. Then, from the data, they found that the higher the CD5, the more self-reactive the cell, resulting in it being less likely to differentiate into TC17 cells. Though, unable to variate the culture conditions to increase/decrease CD5 expression, the results are only based on how each subset (with already varying CD5 levels) performed against various inflammatory diseases. (Lee et al., 2024). Because of this study, it can be determined that developmental self-reactivity is correlated to T-cell's ability to differentiate, as high levels of self-reactivity decrease the chances that cell differentiation occurs to result in heterogeneous T cells. Therefore, when looking to induce the likelihood of T-cell differentiation in vitro, self-reactivity should be a factor that is eliminated from the culture; the occurrence of the elimination of CD5 is currently being studied as a proposal has been formed that the knockout (KO) of the CD5, using specific-antigen receptor T-cells to target the CD5 cell ("*A first-in-human clinical trial of CD5*").

Transcription Factors that promote T-cell differentiation

Specific oncogene factors (transcription factors) have recently been found to be directly affiliated with T-cell differentiation. St. Jude's department of immunology found that when removing the transcription factor, ETS proto-oncogene 1 (ETS1), and recombination signal binding protein for immunoglobulin kappa J region (RBPJ), the total of T-cells that conveyed a dramatically increased antitumor response in mouse models (to replicate genetic processes in humans). Through their research, the results gathered provided sufficient data to imply that specific transcription factors are responsible for T-cell differentiation states. ("*St. Jude scientists identify T-cell differentiation nodes*", 2015). Ultimately, these findings will be supplemental in further studies in cancer immunotherapy, as extracting these transcription factors will lead to higher levels of antitumor activity.

Conclusion

Recent research has proven that lncRNAs have a massive role in terms of the progression of cancer through their ability to bind to glycolytic enzymes or regulate certain protein functions. This suggests that lncRNA could be useful in detecting the early development of cancer at an enzymatic level. Analyzing the interactions between lncRNA and proteins indicates the altering nature of enzymes that favor the Warburg effect as well as the rapid proliferation of cells into a malignant tumor.

As for treatment, these three techniques: The Microenvironment and Environmental Sensors, Developmental Self-Reactivity, and Transcription Factors, can promote T-cell differentiation to help treat cancer through Chimeric Antigen Receptor (CAR) T-cell immunotherapy, as a large quantity of T-cells is required to convert into CAR T-cells to effectively treat cancer. Using this knowledge, researchers can further study the effects of these factors as well as exploring the possibility of the application of these factors on immunotherapy in the future. A promising direction for the use of T-cell differentiation is CAR T-Cell Therapy, considered "the new wave of technology", has proven to be successful in liquid tumors due to its 70-94% success rate in B-cell acute lymphocytic leukemia (B-ALL) and significant success in other blood cancers by as targeting specific cancerous ones (Wittibschlager et al., 2023). CAR T-Cells have a specific antigen receptor to cancer cells, which recognizes cancer cells' threat and fights them off effectively, whereas, in an unaltered immune system, T-cells' capability may be limited from the exposure to solid cancers because of a natural response to stress as found in studies (Derewicz, 2022). For this treatment to be made possible, a large amount of T-cells must be collected, capitalizing the treatment's potential with a universal and significant bank of T-cells, which later undergo CAR Cell membrane insertion ("*What is CAR-T cell therapy?*", 2022). To create this bank of T-cells, stem cells can be used to differentiate into T-cells, though various variables must be considered to create ideal conditions for differentiation.

Works Cited

- Ahmad, Mohammad, et al. "Involvement of lncRNAs in cancer cells migration, invasion and metastasis: Cytoskeleton and ECM crosstalk." *Journal of Experimental & Clinical Cancer Research*, vol. 42, no. 1, 18 July 2023, <https://doi.org/10.1186/s13046-023-02741-x>.
- Ali, Tamer, and Phillip Grote. "Beyond the RNA-dependent function of lncRNA genes." *eLife*, vol. 9, 23 Oct. 2020, <https://doi.org/10.7554/elife.60583>.
- Chae, Hee-Suk, and Seong-Tshool Hong. "Overview of cancer metabolism and signaling transduction." *International Journal of Molecular Sciences*, vol. 24, no. 1, 20 Dec. 2022, p. 12, <https://doi.org/10.3390/ijms24010012>.
- Derewicz, Mark. "Why Don't T Cells Destroy Solid Tumors during Immunotherapy?" *Newsroom*, 20 Dec. 2022, news.unchealthcare.org/2022/12/why-dont-t-cells-destroy-solid-tumors-during-immunotherapy.
- "A First-in-Human Clinical Trial of CD5 Knocked-out Chimeric Antigen T Cells for T-Cell Lymphomas." *Leukemia and Lymphoma Society*, www.lls.org/award/first-human-clinical-trial-cd5-knocked-out-chimeric-antigen-t-cells-t-cell-lymphomas. Accessed 8 Sept. 2024.
- Gao, Na, et al. "Long non-coding RNAs: The regulatory mechanisms, research strategies, and future directions in cancers." *Frontiers in Oncology*, vol. 10, 18 Dec. 2020, <https://doi.org/10.3389/fonc.2020.598817>.
- Kang, Hyunkoo, et al. "The Warburg effect on radioresistance: Survival beyond growth." *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, vol. 1878, no. 6, Nov. 2023, p. 188988, <https://doi.org/10.1016/j.bbcan.2023.188988>.
- Kung, Johnny T, et al. "Long noncoding RNAs: Past, present, and future." *Genetics*, vol. 193, no. 3, 1 Mar. 2013, pp. 651–669, <https://doi.org/10.1534/genetics.112.146704>.
- Lee, Gil-Woo, et al. "Developmental self-reactivity determines pathogenic TC17 differentiation potential of naive CD8+ T cells in murine models of inflammation." *Nature Communications*, vol. 15, no. 1, 4 Apr. 2024, <https://doi.org/10.1038/s41467-024-47144-4>.
- "NCI Dictionary of Cancer Terms." *Comprehensive Cancer Information - NCI*, www.cancer.gov/publications/dictionaries/cancer-terms/def/cancer. Accessed 8 Sept. 2024.
- Panzitt, Katrin, et al. "Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA." *Gastroenterology*, vol. 132, no. 1, Jan. 2007, pp. 330–342, <https://doi.org/10.1053/j.gastro.2006.08.026>.
- Ramsay, George, and Doreen Cantrell. "Environmental and metabolic sensors that control T cell biology." *Frontiers in Immunology*, vol. 6, 17 Mar. 2015, <https://doi.org/10.3389/fimmu.2015.00099>.

- “St. Jude Scientists Identify T-Cell Differentiation Nodes to Improve Cancer-Killing.” St. Jude Children’s Research Hospital, 15 Nov. 2023, www.stjude.org/media-resources/news-releases/2023-medicine-science-news/scientists-identify-t-cell-differentiation-nodes-to-improve-cancer-killing.html.
- Wang, Chunqing, et al. “Interactome analysis reveals that lncRNA HULC promotes aerobic glycolysis through LDHA and PKM2.” *Nature Communications*, vol. 11, no. 1, 22 June 2020, <https://doi.org/10.1038/s41467-020-16966-3>.
- Wang, Ping, et al. “Structural insight into mechanisms for dynamic regulation of PKM2.” *Protein & Cell*, vol. 6, no. 4, 4 Feb. 2015, pp. 275–287, <https://doi.org/10.1007/s13238-015-0132-x>.
- “What Is CAR-T Cell Therapy? A New Way to Treat Cancer.” *BioInformant*, 8 Mar. 2022, bioinformant.com/car-t-cell-therapy/.
- Wittibschlager, Valerie, et al. “Car T-cell persistence correlates with improved outcome in patients with B-cell lymphoma.” *International Journal of Molecular Sciences*, vol. 24, no. 6, 16 Mar. 2023, p. 5688, <https://doi.org/10.3390/ijms24065688>.
- Yang, Guodong, et al. “LncRNA: A link between RNA and cancer.” *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, vol. 1839, no. 11, Nov. 2014, pp. 1097–1109, <https://doi.org/10.1016/j.bbagr.2014.08.012>.

The Impact of Hormonal Fluctuations on Brain Structure and Function in Women Across Reproductive Stages By Misha Nasarpuri

Abstract

Throughout a woman's life cycle, she undergoes hormonal changes due to biological aging and reproductive stages. The hormonal fluctuations accompanying the biological complexities during reproductive stages also alter brain size. In this review article, the different reproductive cycles women experience with their bodies will be discussed to understand further how it affects their brains. For example, during menses, it is noted that the brain volume of gray matter goes down from menses to the mid-luteal phase. In contrast, central spinal fluid volume increases simultaneously, indicating solid matter fluctuations in the brain during the different weeks of the menses cycle.

Furthermore, during pregnancy, the ventricles in the female's brain increase during the term and decrease later. At the same time, the overall brain size does the opposite. Lastly, during menopause, women's brain size decreases extremely, not only because of age but also because of a loss in sex hormones. By understanding what is happening during these reproductive stages, it could influence the way doctors advise women in the future, and change the clinical portion of how we look at these cycles.

Introduction

In general, women go through four major reproductive cycles during their lifetime: menstruation, pregnancy, postpartum, and menopause. The first major stage in a woman's reproductive life is menstruation, also known as menses. The menses typically begin around the age of 15 for a female adolescent (De Sanctis). The next stage a woman can experience is pregnancy. During pregnancy, a fetus develops in the woman's uterus after a woman's egg is fertilized by sperm. Generally, pregnancy lasts around 38 weeks and involves significant hormonal changes and physical changes to the mother's body and brain. These drastic changes in the hormonal circadian rhythm play a crucial role in establishing a strong maternal bond. Following pregnancy, some women may experience postpartum depression, which is characterized by a drastic hormonal fluctuation and can last up to a year or longer, causing the mother to feel sad or lose her attachment to her baby (Brunton). The final reproductive cycle is menopause, which typically begins around the early fifties (Richardson). During menopause, hormone levels decrease, and ovulation ceases. When a woman experiences the absence of her menses for 12 consecutive months after their last menses, she is experiencing menopause (Mckinlay).

Menses: The hormones and the brain during puberty

During menses, a woman's hormones undergo a cascade of coordinated hormonal fluctuation, which induces changes in brain volume (Barth). Understanding the hormonal underpinnings of brain function reveals the intricacies of this relationship. To initiate the menses

cycle, five major sex hormones are required: estradiol, progesterone, follicle-stimulating hormone (FSH), luteinizing hormone (LH), and testosterone; which all of which fluctuate during the 28-day menses (Aloufi). Since these five major hormones are essential for a woman's body, it is crucial to understand these more in-depth.

Estrogen is a sex steroid hormone (Chen) that comes in three major forms: estradiol, estriol, and estrone. A steroid hormone helps to maintain and grow muscles (Kim). Although all forms of estrogen are synthesized in the brain, estradiol is the form estrogen has during a woman's menses. During menses, estradiol plays a role in ovulation (Motwani), where estradiol increases the size of the uterine lining during this time (Vigil). Estrogen has other roles during pregnancy and menopause, which will be discussed later.

Progesterone is an endogenous steroid hormone mostly produced in the adrenal cortex (Cable). This hormone prepares a woman's uterus lining, also called an endometrium to help a fertilized egg grow (Cable).

Testosterone is a sex hormone, which can help with bone growth, strength, and libido. Though females have less testosterone than males, it is still present in their bodies. For males, testosterone is high during the morning, which means they are more energetic. As it drops throughout the day, they become more tired and relaxed. Unlike how it stays relatively stable throughout the menses.

Luteinizing hormone, or LH is a glycoprotein hormone that helps control the menses. It does this by binding to receptors in the ovaries, after being released in the blood. It triggers ovulation and aids in the production of estradiol (Filicori).

Finally, follicle-stimulating hormone, or FSH is a glycoprotein hormone. Like LH, it helps regulate the menses by binding to receptors in the ovaries. The difference between the two is that FSH stimulates follicles on the ovary to grow, while also preparing eggs for ovulation (Howles).

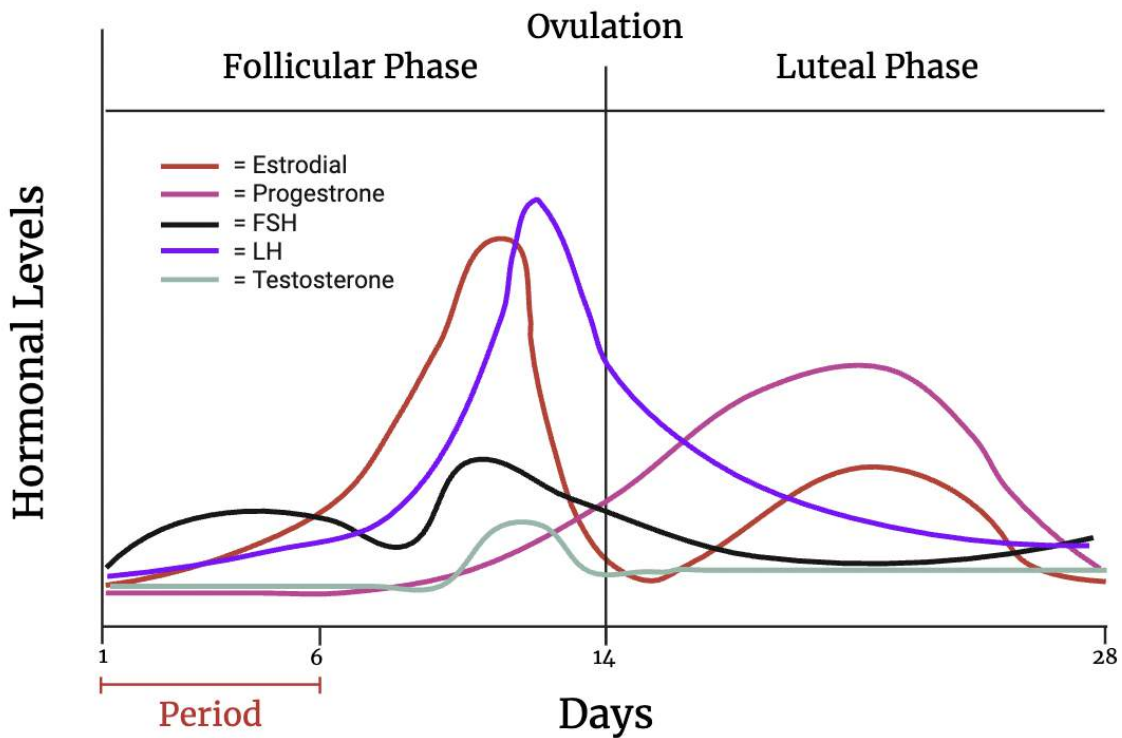


Figure 1 created with BioRender.com - Graph of sex hormones during Menses

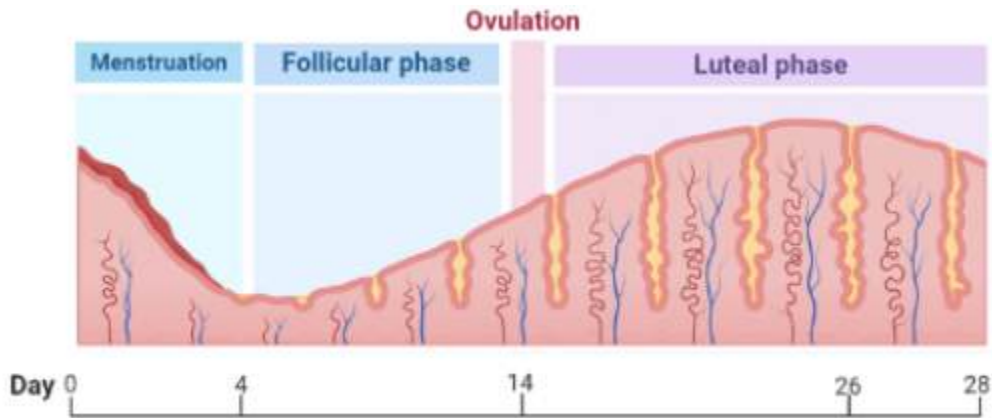


Figure 2 modified image from Kapper - Thickness of uterus lining through Menses

Week One:

Estradiol rises during the first week of the cycle, which can boost mood and energy over time. Though, since it is so low during the first few days it correlates to low energy and fatigue. Additionally, during this time FSH increases (Figure 1), this is in preparation for ovulation as it tries to prepare eggs. Finally, the uterine lining's thickness decreases due to this being where a woman's period would take place (Figure 2).

Week Two:

During the second week, Estradiol continues to rise, and peaks (Figure 1). Additionally, testosterone also starts to rise and peak as well, and the combination of estradiol and testosterone results in women being the happiest throughout their whole cycle during this time. This is the only time testosterone changes (Figure 1). Women have high energy due to this rise in testosterone, which leads to a higher libido and increases in strength. FSH also peaks during this time (Figure 1), as ovulation starts in the middle of week two. Similarly, LH peaks during this time, with the highest hormonal levels of all the hormones (Figure 1). During peak hormone levels for FSH and LH (around day 13 - 14), the ovulation phase enters its final stage. Finally, the uterine lining starts to increase in thickness (Figure 2) due to the increase of these hormones.

Week Three:

In the third week, progesterone starts to rise while the rest of the hormones drop (Figure 1), which leads to a sluggish mood and feeling emotionally low. During the second half of the week, estradiol starts to rise again (Figure 1). Progesterone is a sedating hormone, which is why it is possible to feel low in energy during the first half of the week. When estradiol rises, it puts an end to the low energy. The uterine lining thickness continues to increase (Figure 2).

Week Four:

During the last week, estradiol continues to drop, which leads to body pain and moodiness. Progesterone levels also start to drop at the same time, which makes some women feel energetic. As demonstrated in Figure 1, progesterone after ovulation is due to the fertilized egg implantation into the uterus after ovulation. The endometrium must be prepared to hold it and help it grow. Assuming that there is no fertilized egg, progesterone levels will drop once again after peaking at 22 days to let the ovarian cycle restart. The uterine lining reaches its thickest during this time but shortly starts to decrease again in preparation for the period (Figure 2).

The female cycle lasts 28 days, compared to the male cycle, which lasts only a day. Some studies have shown that during these four weeks, different parts of the brain change, especially the ones that depend on hormones, and they found that this correlated with function. Compared with men, women's brain volume had become less than at the start of the cycle (Hagemann). This is most likely due to the intense hormone fluctuations during the cycle.

These fluctuations can lead to variations in mood, cognitive function, and overall brain efficiency throughout the menstrual cycle. It is noted that gray matter brain volume decreases from menses to the mid-luteal phase, while central spinal fluid volume increases during this time. More research can be done on how the brain changes throughout each week of the menses cycle, to understand how each hormone impacts the brain.

Pregnancy: The brain during Gestation

Pregnancy lasts around nine and a half months (or 40 weeks), while menses typically last for a shorter duration. Even though pregnancy is only 40 weeks long, it triggers significant

changes in the brain. It marks the beginning of substantial hormonal changes that cause the mother's brain to adapt, fostering a bond with her child. These hormonal changes also result in alterations in the brain's structure. Hormones such as progesterone, estriol, which is the primary form of estrogen during pregnancy (Merrill), and oxytocin, a hormone that manages lactation, and contractions during birth (Walter), change often throughout pregnancy. The hormonal changes lead to changes in the brain structure, specifically when the brain is relaxing (Barth). A part of the brain that changed the most was the pituitary gland, which increased in size during pregnancy and then decreased after (Oatridge). This is due to the increase in the release of oxytocin during pregnancy, since there is a greater need for the hormone in the body, the pituitary gland becomes bigger to produce more (Barth). This contrasts with the reduction in gray matter observed in the brain. Yet, after the first pregnancy, structural changes in the brain, and stay like that for at least two years after birth (Luders).

Additionally, several studies have found changes in gray matter in brain regions involved in social cognition. One of the most notable changes occurs in the pituitary gland (Oatridge), which increases in size during pregnancy and then decreases after. Unlike the reduction of gray matter in the brain (Figure 3), there is no change in the white matter. This is important because it can be predicted that due to the reduction in gray matter, functions like the ability to process information become difficult (Driemeyer). Yet, since there was no change in white matter, functions like communicating information stayed the same (Filley). Furthermore, after the first pregnancy, the structural changes in the brain persist for at least two years after birth. These changes can affect mental health, which is why many women are often diagnosed with postpartum depression.

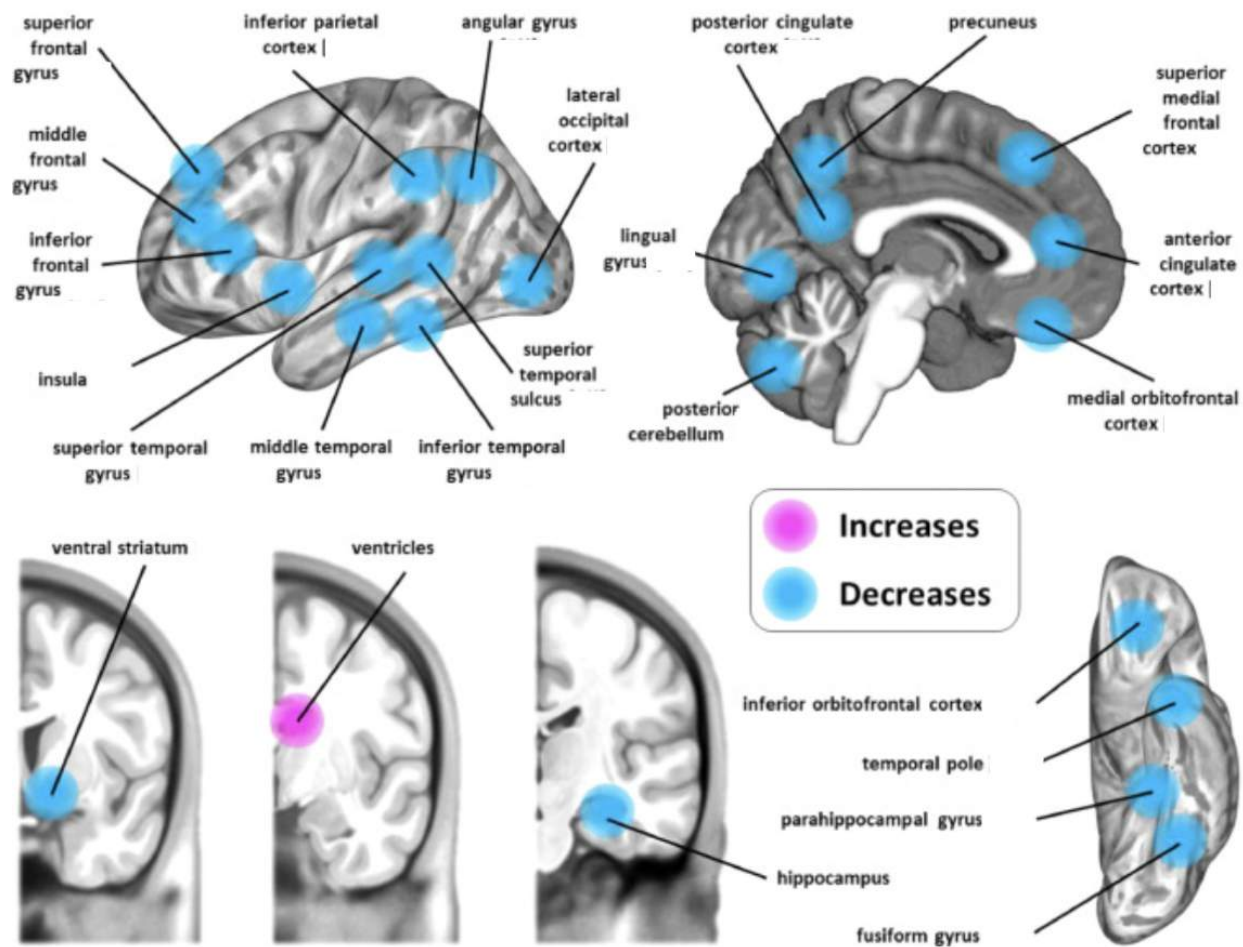


Figure 3 modified image from Luders - How the brain changes during pregnancy

Postpartum: The brain after Gestation

After pregnancy, women go through a period called postpartum. This is where a mother's body starts to return to its pre-pregnant state. It normally lasts 6-8 weeks. During this period, many mothers are in a depressive state, as their bodies are adapting to new things, such as now having a bond with their child, but the child is no longer in their body. Postpartum depression can often evolve into major depression later on in life.

As for the brain, both estradiol and progesterone decreased significantly during postpartum, which could correlate to brain size changes (Luders). In a different study, they compare women's brain size 1-2 days after childbirth vs 4-6 weeks after birth. The results were that on average, they found the difference in their brains was 5 years backwards. Essentially, the researchers found that the brains of women 1-2 days after childbirth were, on average, similar in size to those of women who were 5 years younger. By 4-6 weeks postpartum, the brain size had not fully returned to its pre-pregnancy state, indicating that childbirth and the postpartum period involve substantial brain changes. Overall, the conclusion this study came up with is that pregnancy had a 'negative effect on the brain' as the size of the brain decreased, and the amount of gray matter increased (Luders) and this was the opposite of what happened during pregnancy (Figure 4).

Some examples of the negative effect on the brain include that the postpartum period has been linked with an increased risk of cognitive impairment, which primarily presents as poor memory or recent memory loss, forgetfulness, difficulty concentrating, and distractibility (Zheng). Postpartum-related cognitive impairment may play a major role in various postpartum psychiatric disorders (Henry). Also, significant drops in estrogen and progesterone levels can lead to mood swings, anxiety, and depression. Mental health issues, such as postpartum depression and anxiety are common, with symptoms including severe mood swings, fatigue, and feelings of guilt or shame. Additionally, the demands of caring for a newborn can lead to chronic stress and sleep deprivation, further impacting mental health. There needs to be more research done on the changes of these specific parts of the brain, the reasons behind them, and impacts caused by each change.

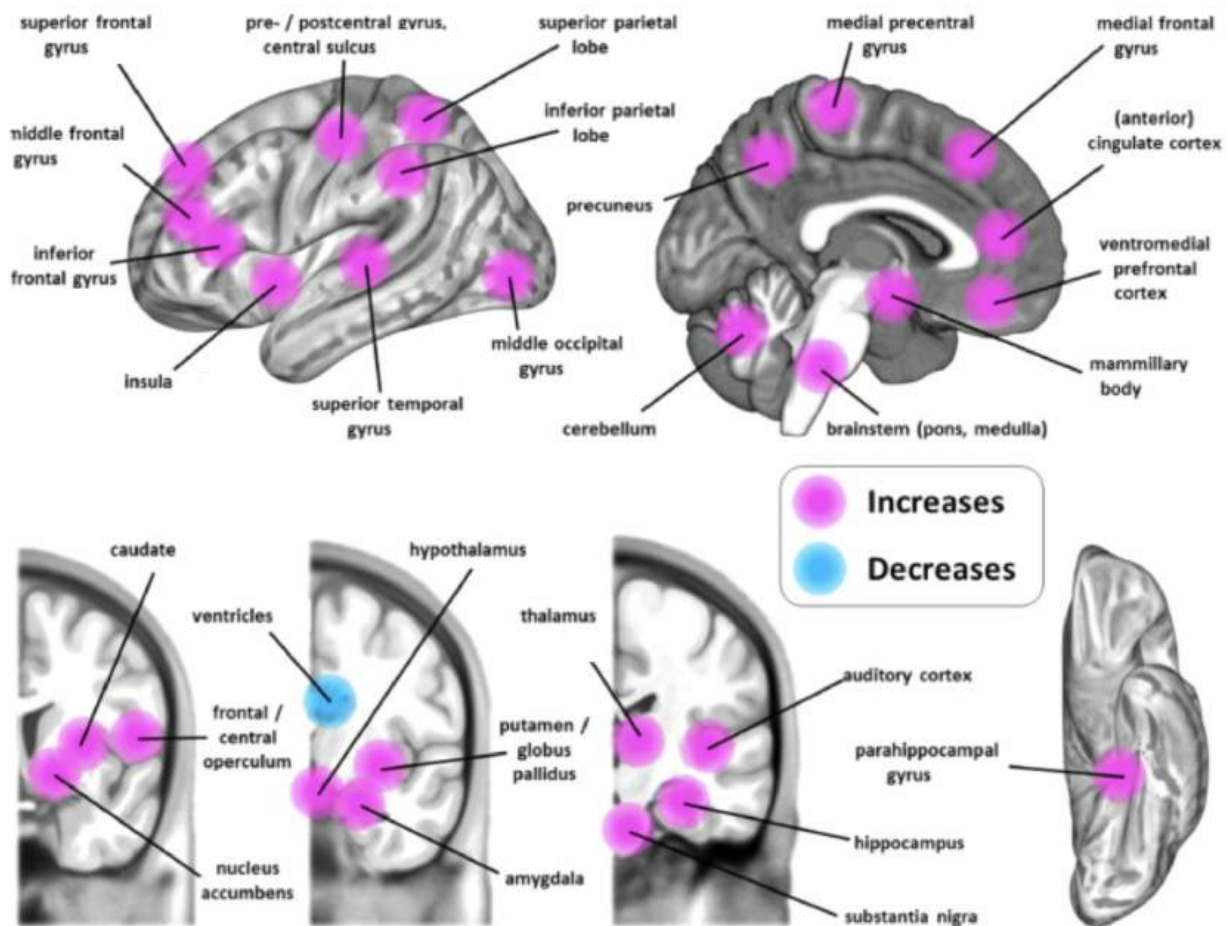


Figure 4 modified image from Luders - How the brain changes during the postpartum period

Menopause: The Brain as Reproductive Hormones Decline

Like in postpartum, key research is missing on why specific parts of the brain are changing during Menopause. Menopause is when a woman experiences the absence of her

menses for 12 consecutive months after their last menses. During this time, estrone, the primary form of estrogen during menopause, and progesterone decrease immensely. During menopause, women often are uncomfortable, and have major symptoms, such as hot flashes and night sweats.

Because of these major shifts in a woman's hormones, their brain changes as well. For example, it was found that the total volume of the brain as a woman ages goes down by 5 years. In postmenopausal women, this reduction in brain volume is equivalent to the brain aging by an additional 58 years (Then). This makes them more vulnerable to aging effects. Additionally, menopause affects the brain because estrone is involved with more than reproduction. When estrone doesn't activate, the brain is not able to regulate temperature in the body through the hypothalamus correctly (Zhang). This causes the symptoms of trouble sleeping, causing sleep patterns to change.

Conclusion

When a woman experiences the four events mentioned above, it's evident that significant brain changes occur, often due to hormonal fluctuations. These brain changes can make women feel disconnected from their usual selves. Though support resources exist to help women understand their bodies better, these are often overlooked. Healthcare providers must provide a more comprehensive approach to providing women with the necessary knowledge during the various phases experienced by women. It is crucial to share this information with women at any life stage, as they may experience or be experiencing dysphoria without understanding the underlying causes.

For example, if someone you know has menopause, some types of treatments can help. Such as eating a Mediterranean diet, doing weight-bearing activities, or even hormone replacement therapy (HRT), can all aid the symptoms (Valtueña). To focus more specifically, HRT replaces estrogen that the body stops making during menopause (Palacios). This can help when women get symptoms such as hot flashes and nightly pains.

However, in the bigger picture, it is important to understand that the scientific community has research gaps on women's health. Research gaps exist around brain changes during reproductive cycles and how environmental and lifestyle factors may contribute to them. Factors such as stress, pollution, diet, physical activity, and exposure to endocrine-disrupting chemicals can influence hormonal balance and brain function. Understanding these influences is crucial for developing strategies to mitigate their impact on women's health during different reproductive stages (Bala).

It was only in 1993 that the NIH mandated that there should be equal representation for women and minorities in research. This can be seen in women's mental health, as they are often diagnosed with anxiety or depression, often due to underlying causes, in some cases it can be due to the immense amount of hormonal fluctuations they go through. Another example of delays in research because of historical biases in women's healthcare, is that only after 2002, the Women's Health Initiative (WHI) study revealed that prolonged use of HRT could increase the risk of heart disease, stroke, and breast cancer; leading to fewer HRT prescriptions. Understanding that there

is a bias in women's health, which makes a big difference. Often, women aren't taken seriously when they have a problem and therefore have to advocate for themselves. If we start to understand and bridge this knowledge gap, it influences others to start talking about these important topics. One topic where research can be done is how the brain changes throughout each week of the menses cycle, to understand how each hormone impacts the brain.

An area where research can be done is when HRT is the most impactful for women with menopause. This can help in preparing healthcare professionals to provide better and more timely treatments to prevent different mental health issues. Clinicians can look into newer non-hormonal options like selective serotonin reuptake inhibitors (SSRIs) or gabapentin may be recommended for managing symptoms, for those who cannot take HRT (Davis). Actionable items for women themselves can include asking and researching approaches to help with their reproductive worries.

Healthcare providers must provide a more comprehensive approach. Apart from HRT, This may include looking into lifestyle changes such as regular exercise, a balanced diet rich in protein, and staying hydrated can help reduce symptoms and recommend better sleep hygiene and stress management exercises and practices like yoga, meditation, and deep-breathing exercises. All of which can help in any phase of a woman's reproductive life. Physicians can also recommend avoiding triggers like caffeine, alcohol and spicy foods can also be beneficial (Sturdee) during menopause.

In this paper, a variety of studies are used. Though most are newer to the field, the older studies are ones that are vital to understanding how brain structure changes during these reproductive cycles. They are good jumping off points for researchers to do more work on these topics, as next steps are clear. Yet, they also demonstrate how often this research is really happening. We as a community should be more aware of this bias in science, and astigmatism around this topic, and try to help women, and healthcare providers talk about reproductive health.

Work Cited

- Aloufi, Nada, et al. "Sex differences and the effect of female sex hormones on auditory function: a systematic review." *Frontiers in Human Neuroscience* 17 (2023): 1077409.
- Bala, Renu, et al. "Environment, lifestyle, and female infertility." *Reproductive sciences* 28 (2021): 617-638.
- Barth, Claudia, Arno Villringer, and Julia Sacher. "Sex hormones affect neurotransmitters and shape the adult female brain during hormonal transition periods." *Frontiers in neuroscience* 9 (2015): 37.
- Brunton, Paula J., and John A. Russell. "The expectant brain: adapting for motherhood." *Nature Reviews Neuroscience* 9.1 (2008): 11-25.
- Cable, Jessie K., and Michael H. Grider. "Physiology, progesterone." (2020).
- Chen, Changmei, et al. "The roles of estrogen and estrogen receptors in gastrointestinal disease." *Oncology letters* 18.6 (2019): 5673-5680.
- Davis, S. R., et al. "The 2023 practitioner's toolkit for managing menopause." *Climacteric* 26.6 (2023): 517-536.
- De Sanctis, Vincenzo, et al. "Onset of menstrual cycle and menses features among secondary school girls in Italy: A questionnaire study on 3,783 students." *Indian journal of endocrinology and metabolism* 18.Suppl 1 (2014): S84-S92.
- Driemeyer, Joanna, et al. "Changes in gray matter induced by learning—revisited." *PloS one* 3.7 (2008): e2669.
- Filicori, Marco. "The role of luteinizing hormone in folliculogenesis and ovulation induction." *Fertility and Sterility* 71.3 (1999): 405-414.
- Filley, Christopher M., and R. Douglas Fields. "White matter and cognition: making the connection." *Journal of neurophysiology* 116.5 (2016): 2093-2104.
- Hagemann, Georg, et al. "Changes in brain size during the menstrual cycle." *PloS one* 6.2 (2011): e14655.
- Henry, Julie D., and Peter G. Rendell. "A review of the impact of pregnancy on memory function." *Journal of clinical and experimental neuropsychology* 29.8 (2007): 793-803.
- Howles, Colin M. "Role of LH and FSH in ovarian function." *Molecular and cellular endocrinology* 161.1-2 (2000): 25-30.
- Kapper, Celine, et al. "Minerals and the Menstrual Cycle: Impacts on Ovulation and Endometrial Health." *Nutrients* 16.7 (2024): 1008.
- Kim, Yong Jin, et al. "The role of sex steroid hormones in the pathophysiology and treatment of sarcopenia." *Osteoporosis and Sarcopenia* 2.3 (2016): 140-155.
- Luders, Eileen, et al. "Potential brain age reversal after pregnancy: younger brains at 4–6 weeks postpartum." *Neuroscience* 386 (2018): 309-314.
- Mckinlay, Sonja M., Donald J. Brambilla, and Jennifer G. Posner. "'Reprint of' the normal menopause transition." *Maturitas* 61.1-2 (2008): 4-16.
- Merrill, Reynold C. "Estriol: a review." *Physiological reviews* 38.3 (1958): 463-480.
- Motwani, Rohini, and Hare Krishna. "Ovulation Process." (2024).

- Oatridge, Angela, et al. "Change in brain size during and after pregnancy: study in healthy women and women with preeclampsia." *American Journal of Neuroradiology* 23.1 (2002): 19-26.
- Palacios, Santiago. "Advances in hormone replacement therapy: making the menopause manageable." *BMC Women's Health* 8 (2008): 1-5.
- Richardson, Sandra J. "1 The biological basis of the menopause." *Baillière's clinical endocrinology and metabolism* 7.1 (1993): 1-16.
- Sturdee, David W. "The menopausal hot flush—anything new?." *Maturitas* 60.1 (2008): 42-49.
- Then, Stephanie, et al. "Interactions between age, sex, menopause, and brain structure at midlife: a UK Biobank study." *The Journal of Clinical Endocrinology & Metabolism* 106.2 (2021): 410-420.
- Vigil, Pilar, et al. "Ovulation, a sign of health." *The Linacre Quarterly* 84.4 (2017): 343-355
- Valtueña, Silvia, Stefania Sette, and Francesco Branca. "Influence of Mediterranean diet and Mediterranean lifestyle on calcium and bone metabolism." *International journal for vitamin and nutrition research* 71.3 (2001): 189-202.
- Walter, Michael H., Harald Abele, and Claudia F. Plappert. "The role of oxytocin and the effect of stress during childbirth: neurobiological basics and implications for mother and child." *Frontiers in endocrinology* 12 (2021): 742236.
- Zhang, Zhi, et al. "The effects of estrogens on neural circuits that control temperature." *Endocrinology* 162.8 (2021): bqab087.
- Zheng, Jin-Xia, et al. "Disrupted spontaneous neural activity related to cognitive impairment in postpartum women." *Frontiers in psychology* 9 (2018): 624.

The Business of K-Pop: The Role of Angel Investments and Venture Capital in Global Success By Tung-Tung Yeh

Abstract

"The Business of K-Pop: The Role of Angel Investments and Venture Capital in Global Success" is an essay that examines the financial strategies that have contributed to the worldwide rise of K-pop, specifically focusing on the role of angel investments and venture capital. With groups like BTS, BLACKPINK, and SEVENTEEN gaining global recognition, the essay discusses how these funding helped transform K-pop from a regional trend to a global phenomenon. It outlines the historical background of K-pop's rise, detailing how companies strategically used investments to improve music production, choreography, and international marketing efforts. Through the analysis of specific companies like Big Hit (rebranded as HYBE in 2021), JYP Entertainment, and YG Entertainment, the essay demonstrates how these companies use financial support to create high-quality content, engage with fans, and boost sales.

The essay also explores the challenges faced by investors in the K-pop industry, including social and cultural differences, intense competition, and high production costs. Despite these risks, it emphasizes the continued growth potential of K-pop, supported by innovative business models and government support. In conclusion, the paper illustrates that angel and venture capital investments have played an essential role in K-pop's success and will likely continue to support its global expansion

Introduction

As K-pop rises worldwide and gains status in the global market, it has attracted more and more fans in recent years. Groups like BTS, BLACKPINK, and SEVENTEEN have progressively stood in the world's spotlight, even reaching Billboard's Top 100. These achievements, made in just a few recent years, show K-pop's transition from being popular only in Asia to gaining global recognition. How did this happen? Investment has played a crucial role in this success. In the entertainment industry, there are several ways to invest, including government funding, crowdfunding, angel investment, and venture capital, and more. This essay will explore specifically how angel investment and venture capital have contributed to the K-pop industry's worldwide success. By discussing specific case studies and analyzing the strategies used by investors, we can understand the important role these investments have played in transforming K-pop into a global phenomenon.

Background of K-pop

Korean pop music, known as K-pop, originated in the 1950s. If you are new to K-pop, you may only be familiar with BTS (which is perfectly fine), but the concept of group singers was already popular back in the 1950s in Korea. However, the era that truly attracted international attention began in the 1990s. Seo Taiji and Boys, the first generation of K-pop groups (Fury), gained attention from other Asian countries with their hip-hop choreography,

catchy melodies, and fashionable clothes. Their debut was a significant step for Korea to go global. In the 2000s, the next generation of groups like Big Bang and Super Junior debuted and received even more recognition. By the 2000s, K-pop had become a trend everyone needed to know about and listen to; otherwise, they were seen as not keeping up with society. In the third generation, which started in the 2000s and continues to the present day, you may recognize more groups such as BTS, BLACKPINK, and SEVENTEEN. These groups are all modern K-pop icons. Besides the investments made by angel investors and venture capital, their music, unique choreography, and high-quality music videos are significant reasons for their success.

Factors that led to K-pop's success

First, the music. After the success of some early K-pop groups, entertainment companies realized that to succeed globally, they needed to make their music attractive and catchy. What does this mean? If you listen to any K-pop songs released in recent years, you will notice that their music often combines English and Korean, especially in the chorus. This allows international fans to sing along and memorize the songs easily without needing to know Korean. For instance, BTS's first English EP, "Dynamite," reached the Billboard Top One upon its release, becoming the first Korean song to achieve this milestone (Pascual and Sheeran).

Next, the unique choreography. K-pop dances are a major attraction for the public. Their synchronized dance moves provide visual satisfaction, and when combined with their catchy songs, it becomes a significant advantage.

Lastly, the high-quality music videos. K-pop music videos are always meticulously produced, with attention to details such as clothing, scenery, and film editing, making audiences want to watch them over and over again. For example, PSY's "GANGNAM STYLE" music video is ranked fifth in the top ten most-viewed YouTube music videos of all time, with 5,220,850,078 views ("YouTube - Most Viewed Music Videos of All Time").

In a company, if all the staff know their tasks but there is no executive to guide and make the team cohesive, it is not going to work out. Similarly, understanding the factors that contribute to K-pop's success is essential before discussing their approach. In the next section, we will explore how K-pop companies strategize and implement their methods.

K-pop company's strategies

Korean entertainment companies have implemented several approaches to increase their idols' exposure, cement the relationship between fans and idols, and generate revenue. These companies often use social media platforms such as Instagram, YouTube, and TikTok to boost international exposure. K-pop companies understand their audience well and frequently use short videos for promotion to both current and potential fans.

In addition, fan clubs are a tradition in K-pop. Almost every K-pop group has its platform (often called "Bubbles") that creates a bridge between idols and fans. On Bubbles, idols can communicate directly with their fans in a group chat format where every fan can see the idols' messages and reply with limited words. Access to these apps usually costs between 5 to 10 US

dollars per month. This allows idols to foster their relationships with fans by giving them a sense of participation in their lives and making money at the same time.

Beyond these engagement strategies, idols also generate significant revenue through album sales, with "random photo cards" being a particularly lucrative portion. A typical K-pop album contains an outbox, a CD, a lyric book, a photo book, and two to three random photo cards (the quantity of the cards depends on the group's size). These photo cards are included randomly in albums, encouraging fans to purchase multiple albums to collect the complete set. Each member in a group will have at least five different versions of their photo cards, and each album often has about three versions. For example, if a group has five members and their latest album has three versions, with each member having five different cards per version, there would be a total of 75 photo cards.

Many loyal fans purchase multiple albums to collect all the different photo cards. Considering that most loyal fans might buy twenty albums per version, and there are countless loyal fans, the revenue generated can be substantial. According to the International Federation of the Phonographic Industry (IFPI), 19 of the 20 best-selling albums came from K-pop. SEVENTEEN's album "FML" was the best-selling album in 2023, with 6.4 million copies sold (McIntyre). With the "random photo card" strategy, K-pop albums have successfully sold in large quantities, making significant profits.

Involvement of Angel investment Venture capital

Angel investment plays a crucial role in funding the early stage of an entertainment company, as investors can provide the initial capital that is necessary to develop their projects, secure talent, and money for keeping staff and idols. Angel investments usually are given by personal funds in exchange for equity-convertible debt. The advantage of angel investment is that it allows startups to bypass traditional funding methods, which might make it hard to get funds due to the high risks associated with a new company. Likewise, venture capital provides the same opportunity to startups but slightly different. Venture capital (VC) firms invest by investing others's money, not their own money. However, VC also plays a crucial role in startups. Venture capital provides substantial financial resources for growing operations, reaching new markets, and improving production quality in startups with high growth potential in exchange for equity. This investment allows entertainment companies to access not only the necessary capital but also strategic guidance, connections with the industry, and business expertise from experienced venture capitalists. With the help of angel investment and venture capital, they not only offer financial aspects but also provide valuable expertise, networking opportunities, and mentorship, which can be vital for the growth and success of emerging entertainment companies. Here are some real-life examples of how VC investment helps some top K-pop entertainment companies to thrive.

Big Hit Entertainment (rebranded as HYBE in 2021), the company of BTS, is a successful example of a K-pop company that has benefited significantly from venture capital investment. SV Investment, a South Korean venture capital firm, invested 3.3 million in Big Hit

before BTS became the international star they are today. The return on this early investment was spectacular, with a return of over 25 times the initial amount. As Big Hit Entertainment grew, the company's revenue increased from \$30 million in 2016 to \$200 million in 2018, highlighting the impact of venture capital on its expansion and success (“Top 20 Most Active Korean Venture Capital Firms & Private Equity Firms”).

JYP Entertainment, which manages groups of TWICE and ITZY, has also received substantial venture capital funding. Atinum Investment, a famous venture capital firm in South Korea, has invested in JYP Entertainment, helping in its expansion and the enhancement of its global marketing strategies. This financial support helps JYP Entertainment maintain its competitive edge in the entertainment industry (PGD Capital Advisors).

YG Entertainment, the company of BLACKPINK and BIGBANG, has also benefited from venture capital investments. These funds have allowed YG to invest in high-quality production, international exposure, and well-produced music, solidifying its global presence and commercial success.

Both angel investment and venture capital play critical roles in the success of K-pop groups by providing the necessary financial resources and strategic support. Although no existing document has recorded a specific example of a K-pop company receiving funds from angel investments, some entertainment startups have benefited from such support. Angel investments, typically funded by individuals, may not always be substantial enough to be noticed by the public. However, these investments, along with venture capital, enable companies to produce high-quality content, reach international markets, and build strong fan bases, ultimately contributing to the global success of K-pop.

Process of the investment

In order to invest successfully, there are plenty of steps that need to be done. *Due diligence:* Investors, both angel and venture capitalists, conduct thorough due diligence before committing funds. This involves evaluating the company’s management team, business model, and financial health. They look for the talent involved, production quality, and the potential audience. Due diligence helps in mitigating risks associated with entertainment investments (Segal).

Business plans: Entertainment companies present detailed business plans to potential investors. These plans typically include a project overview that describes the K-pop group and their unique points, as well as financial projections detailing expected revenue derived from album sales, concerts and merchandise. Additionally, the business plan outlines the marketing strategy, highlighting how the group will be promoted inside Korea and globally through social media and other platforms. Risk management strategies are also included in the plan to address potential challenges and alleviate risks. A comprehensive business plan shows the viability and potential profitability of the project, therefore it is extremely important in investment (“Business Plan: What It Is, What's Included, and How to Write One”).

How investments work

Investments by either individual or venture capital companies are allocated to various aspects such as training, production, and marketing. Funding from investors supports the development of new talent by supporting costs for vocal, dance, and language training. Production investments are used to create high-quality music videos and albums. Significant funds are also used in marketing campaigns, social media, and world tours to enhance the group's visibility. This comprehensive financial support enables K-pop companies to produce top artists, attract local and global fans, and make considerable revenue from album sales, concerts, and merchandise.

Challenges and Risks

By listening to the praise for investments in K-pop, you may feel an impulse to invest in a K-pop startup. However, not every group that successfully debuts will necessarily make money and attain international attention. There are potential risks and challenges you may face if your investment does not turn out well.

First, social differences: For some multinational groups, members may face social differences. It is possible that after they debut, news emerges about internal conflicts, like bullying within the group. Such controversies can seriously damage their reputation, leading to a loss of fan support and potential disbandment before making any profit.

Second, competition: Numerous K-pop groups debut every year. Unless a group debuts from one of the major K-pop entertainment companies, it can be challenging to thrive in such a competitive environment. Investors need to recognize that not all groups will achieve the same level of success as BTS, even they did come from small entertainment companies.

Lastly, high production costs: Creating and promoting a K-pop group can be very expensive. There is a risk that the investment might run out before the group's debut or before they become famous. Therefore, it is absolutely essential to understand the details of a company by thoroughly reviewing their business plan and maintaining clear communication with the CEO.

Future trend

Will K-pop continue to thrive? The public is certainly curious about this. According to the historical revenue growth of HYBE (the biggest K-pop entertainment company), HYBE's revenue increased from \$504 million in 2019 to \$1.674 billion in 2023, reflecting a 34% growth over four years (PGD Capital Advisors). Therefore, there is a significant chance that K-pop will continue to gain increasing profits in the coming years. In addition, the support from the government allows those companies to facilitate the promotion of K-pop culture, further increasing its global recognizability and profitability.

Conclusion

In summary, the success of K-pop can be attributed to several key factors, including strategic investments, innovative marketing, and strong government support. Angel investments

and venture capital have played crucial roles in the K-pop industry. Although there are potential risks and challenges such as cultural differences, high competition, and production costs, these investments have given K-pop startups the opportunity to grow, make profit, and achieve worldwide fame.

Investment from both angel investors and venture capital has been pivotal in making K-pop a global success. The financial strategies employed by investors have allowed K-pop companies to create high-quality content, effectively train members, and build strong relationships between fans and idols.

K-pop has significant potential for continued growth. With innovative content, attractive merchandise, and idols who capture the hearts of fans, K-pop may very fast become a main force in global entertainment. Perhaps in the future, people on the street will be discussing K-pop idols more than they talk about stars like Taylor Swift!

Works Cited

- “Business Plan: What It Is, What's Included, and How to Write One.” Investopedia, <https://www.investopedia.com/terms/b/business-plan.asp>. Accessed 28 July 2024.
- Fury, Rick. “A Brief History of K-pop – The Los Angeles Film School.” LA Film School, 6 April 2021, <https://www.lafilm.edu/blog/a-brief-history-of-kpop/>. Accessed 8 September 2024.
- McIntyre, Hugh. “K-Pop's Global Impact: 19 Of The 20 Bestselling Albums In 2023 Are From South Korea.” Forbes, 27 March 2024, <https://www.forbes.com/sites/hughmcintyre/2024/03/27/k-pops-global-impact-19-of-the-20-bestselling-albums-in-2023-are-from-south-korea/>. Accessed 24 July 2024.
- Pascual, Danielle, and Ed Sheeran. “BTS' Top Songs on the.” Billboard, 13 June 2023, <https://www.billboard.com/lists/bts-top-songs-billboard-hot-100/>. Accessed 21 July 2024.
- PGD Capital Advisors. “The K-Pop Industry: General Overview and Outlook.” Power Point, 2023, <https://pgpcapital.com/wp-content/uploads/2023/09/K-Pop-Industry-Overview-Fall-2023-vF.pdf#:~:text=URL%3A%20https%3A%2F%2Fpgpcapital.com%2Fwp>. Accessed 5 August 2024.
- Segal, Troy. “How to Invest in Movies.” Investopedia, 1 August 2022, <https://www.investopedia.com/financial-edge/0512/how-to-invest-in-movies.aspx>. Accessed 28 July 2024.
- “Top 20 Most Active Korean Venture Capital Firms & Private Equity Firms.” Seoulz, 18 July 2023, <https://www.seoulz.com/top-20-most-active-korean-venture-capital-firms-private-equity-firms/>. Accessed 25 July 2024.
- “YouTube - Most Viewed Music Videos of All Time.” Kwordb.net, <https://kwordb.net/youtube/topvideos.html>. Accessed 21 July 2024.

Responsible AI for Nonprofits: Ensuring Data Privacy and Mitigating Bias in LLMs

By Keshav Rangan

Abstract

Large Language Models (LLMs) are revolutionizing artificial intelligence, enabling unprecedented advancements in natural language processing, understanding, and generation. These models have vast potential in the nonprofit sector to drive innovation, improve efficiency, and enhance overall impact. However, nonprofit organizations are behind the curve in adopting AI tools due not only to limited resources and technical expertise but also ethical concerns about how to utilize these tools in an appropriate manner. This paper explores two major ethical concerns of LLMs: data privacy and bias, diving into the primary risks and solutions for each. Under data privacy, we discuss personally identifiable information (PII) leakage, privacy attacks, legal issues, and subjective privacy harms, offering privacy-enhancing techniques such as synthetic data, differential privacy, and homomorphic encryption and novel ideas for new legal regulations and guidelines. Regarding bias, we examine the origin, nature, and types of bias. Then, we propose an architecture for reducing bias in LLMs, including introducing an operational definition of fairness, recommending metrics for measuring bias, discussing algorithms to mitigate bias, and introducing novel ideas to personalize bias mitigation strategies and eliminate bias at its source. We then explore the Pareto frontier to balance and optimize for model accuracy, data privacy, and fairness. The paper concludes with applications for the nonprofit sector, proposing the guidelines and standards for handling LLMs and best practices for using AI tools.

1) Introduction

Large Language Models (LLMs), which are designed to understand, generate, and manipulate human language based on vast datasets, are at the forefront of artificial intelligence applications today. These models are capable of producing coherent and contextually relevant text, making them valuable for applications ranging from automated customer service to content creation and beyond. Many LLMs exist on the market today, including OpenAI's ChatGPT (Radford et al.), Google Gemini (Anil et al.), Anthropic Claude ("Introducing the next Generation of Claude"), and Meta's Llama (Touvron et al.). These models empower business operations across various industries such as healthcare, education, and finance in various ways such as customer service, content generation, and data analysis and insights.

This paper examines the use of LLMs specifically in the context of the nonprofit sector. Nonprofits often have limited resources and staff, making it essential for them to operate as efficiently as possible. The potential benefits of nonprofit use of AI are vast, offering opportunities to enhance fundraising, improve donor engagement, and optimize program effectiveness. AI can also help nonprofits analyze large amounts of data to identify trends, personalize communication, and streamline administrative tasks, thus maximizing the impact and efficiency of these organizations. However, less than 30% of nonprofits have started using or

exploring AI tools compared to an industry standard of 72% (Cardillo). Although the lack of resources may be a factor, this is mostly due to an underlying fear of ethical concerns surrounding AI as nonprofits highly prioritize the safety of their community.

Data privacy is often listed as one of the largest concerns of LLMs (Busch; Das et al.; Yao et al.). As LLMs are trained on large swaths of data, the data they learn from may inadvertently include sensitive personal information. This is especially problematic for nonprofit organizations where the privacy of community members is paramount. For example, in my work with shelters for unhoused youth, I've helped to develop an application of LLMs that uses data about what worked best for past residents to inform decisions about the best-fitting programs for future residents. It's essential that personal information concerning vulnerable minors is not leaked to the public. Bias is another primary concern with generative AI. The data on which LLMs are trained carry the same biases present in human language and writing. If these biases are not properly addressed, the models may perpetuate or even amplify harmful stereotypes and prejudices. This too is particularly troubling for nonprofit organizations that aim to serve diverse and often marginalized communities. In the same unhoused youth example, the model might favor certain groups over others, either recommending different services or having higher prediction accuracies for those groups. This inadvertently reinforces existing inequalities and fails to equitably serve all residents. Despite the risks, the potential for AI at nonprofits cannot be overlooked. By better understanding these concerns and learning how to address them, we can come up with the proper guidelines for using AI responsibly and effectively.

Section 2, Background, will cover transformers, the ML algorithms behind LLMs, and the use cases of LLMs in industry. Data Privacy (Section 3) covers data privacy concerns in LLMs and privacy-enhancing techniques (PETs) to combat them. Bias (Section 4), covers the origin, nature, and types of bias, alongside a strategic plan to mitigate bias in LLMs. The Pareto Frontier (Section 5), covers an approach to simultaneously prioritize data privacy, bias, and model accuracy. Nonprofit Guide (Section 6), applies this research into actionable steps and guidelines for nonprofits to make best use of AI.

2) Background

2.1: The Foundations of LLMs: The Transformer Model

The transformer model (Muñoz) is a deep-learning algorithm that serves as the basis of most LLMs. Transformers represented a major milestone in ML models with their ability to retain information. In this section, we'll dive into the architecture behind the transformer model without getting too deep into the mathematical intricacies.

2.1.1: The Self-Attention Mechanism

The transformer architecture is based on the self-attention mechanism. Self-attention allows a model to calculate the relative importance of different parts of an input with respect to itself. The input is a sequence of vectors, a numeric representation of words, and the output is

also a sequence of vectors. Each input vector is linearly transformed into three weighted matrices: Query, Key, and Value. A dot product is taken between the Query of each input vector and all Keys of all other input vectors to measure how much focus or attention should be given to each part of the input with respect to a certain part of the input. The results are converted into attention weights, used to take a weighted average of each Value vector, which is returned as an output vector (Muñoz).

2.1.2: The Encoder

The encoder architecture serves the primary purpose of understanding and processing input text. Although not as popular in LLMs today, many models are encoder-only including Google BERT (Devlin et al.). Some models use a combination of the encoder and the decoder such as Google Gemini. The encoder is made of six identical layers with two sub-layers: a multi-head self-attention mechanism and a fully connected feed-forward network (linear transformation with activation function). In an encoder-decoder model, the output of the sixth layer is sent to the decoder (Muñoz).

2.1.3: The Decoder

The decoder architecture serves the primary purpose of generating text. The decoder-only model is popular in most generative LLMs today such as OpenAI's ChatGPT, Anthropic Claude, and Meta Llama. The decoder is similar to the encoder in many ways with six identical layers with three sub-layers each. First, masked multi-head attention ensures that predictions at a certain position in the input only depend on values before that position. Second, encoder-decoder attention performs multi-head attention over the output of the decoder. The third sub-layer is a fully connected feed-forward network. At the end of the six layers, a linear layer transforms the output into one large vector, followed by a softmax layer that turns this vector into probabilities. The cell with the highest probability is chosen for the output (Muñoz).

2.2: LLM-Enhancing Techniques

Aside from large tech companies and research institutions, most organizations, including nonprofits, will not develop their own LLMs, as they require significant amounts of computational resources including large datasets and advanced hardware and software, alongside energy, time, and financial requirements. The most common approach to implementing AI is the usage of 3rd party APIs such as OpenAI's API. When used safely, 3rd party APIs are vastly beneficial to nonprofits. There are several LLM-enhancing techniques that can be used to prioritize certain aspects of the model and improve certain features.

Retrieval-Augmented Generation (RAG) ('Retrieval-Augmented Generation') enhances language models by integrating information retrieval into the generation process, allowing the model to access specific information. It works by first retrieving relevant information based on the user's prompt from an external database and then using this information as context to generate more accurate and specific responses (Ferrara).

Few-shot prompting (Brown et al.) enhances language models by training them to solve a specific task. It works by including a small number of input-output pairs in the prompt, guiding the model on how to handle similar inputs and produce the desired outputs (Ferrara).

Fine-tuning (Dodge et al.), a technique used widely in ML, enhances language models by training them on specific datasets tailored to a particular domain, improving their performance and accuracy for a specific use case. It works by starting with a pre-trained model (such as ChatGPT) and then adjusting its parameters through additional training on the specialized dataset, allowing the model to better understand and generate relevant responses (Ferrara).

3) Data Privacy

3.1: Concerns and Harms

3.1.1: Personally Identifiable Information (PII) Leakage

By nature, generative AI generates content by learning from patterns in pre-existing data. Large amounts of data are required for both training and fine-tuning LLMs so that the model can recognize patterns and make adjustments. The more data we use, the more likely sensitive personal information gets revealed (Busch). When PII exists in a training dataset, it may find its way to the model output via user-inputted prompts (Francis). Pre-trained models such as ChatGPT can retain or “memorize” personal data and their link to its owner, so data is retrievable through prompting with the relevant context (Das et al.). When PII is leaked, a person’s personal information may be used against them: we refer to this as an objective privacy harm (Valentino-DeVries).

3.1.2: Malicious Attacks

Attack Pathways: Attack pathways are characterized by the place in which a malicious attack occurs in the context of a system infrastructure. The pathways are ordered from the most abstracted and human-facing layer (user-level) to the most physical and concrete layer (hardware-level), as classified in the paper “A Survey on Large Language Model (LLM) Security and Privacy: The Good, the Bad, and the Ugly” (Yao et al.).

- User-Level Attacks: User-level attacks aim to manipulate other users, leveraging LLMs to generate deceptive content. They are the most common form of attack with LLMs and can be applied in several ways. This includes exploiting PII leakage issues to infer personal attributes from generated text and generating fraudulent content to facilitate cyberattacks.
- Network-Level Attacks: Network-level attacks compromise the infrastructure or communication within a network, such as generating realistic and convincing personalized phishing emails.
- Software-Level Attacks: Software-level attacks involve using malicious software to exploit vulnerabilities in software applications, such as using LLMs to generate malware.

- OS-Level Attacks: OS-level attacks target and exploit vulnerabilities in the operating system, such as using LLMs to analyze system information and develop attack strategies. These are less common, however, as LLMs lack low-level system access.
- Hardware-Level Attacks: Hardware-level attacks exploit vulnerabilities in physical devices, such as analyzing unintended information leakage to extract sensitive data. These too are less common, as LLMs cannot access physical devices, only data associated with them.

Types of Attacks: Here, we'll break down several different types of attacks that can be used with LLMs. However, many attacks remain theoretical due to the fact that many mainstream LLMs today are well-secured and able to censor their outputs. The large scale amount of parameters, well into the billions, in mainstream LLMs also makes them difficult to attack. Nevertheless, these attacks are still feasible and worthy of looking into.

- Adversarial Attacks: Adversarial attacks aim to intentionally deceive ML models by manipulating the data used to train the model or the model itself. For example, by inserting malicious data into the training dataset, the model may lose its effectiveness, as pre-trained models are often susceptible to even small-scale data injection. Additionally, some attacks may alter the model processes themselves, triggering the model to behave or respond in various ways (Yao et al.).
- Inference Attacks: Inference attacks aim to gain sensitive information by making specific queries and observations about the model behavior (Yao et al.). Attribute inference attacks deduce personal information by analyzing the model's outputs (Yao et al.). Membership inference attacks determine whether or not a data sample is in a model's training dataset, exploiting the model's tendency to overfit training data (Das et al.).
- Leakage and Extraction Attacks: Extraction attacks directly access confidential information such as the model's gradient and the training data through strategic querying (Yao et al.). Gradient Leakage Attacks access gradients, which control model optimization, to manipulate and extract data from the training dataset. PII Leakage Attacks aim to reveal personal information from a training dataset by presenting the relevant context, as discussed in 3.1.1 (Das et al.).

3.1.3: Subjective Privacy Harms

Although privacy harms generally stem from access to personal information, privacy harms may still occur without access to PII: we call these harms subjective privacy harms. Even with privacy-preserving language models (PPLMs), you may still be targeted due to the nature of LLM computation. Although your data is not associated with your personal identifiers, the data is still reflected in the model and may give you the feeling as if you are being watched. Subjective privacy harms are not emphasized in many laws and regulations (Yao et al.). We further detail this at the end of 3.1.4.

3.1.4: Legal Regulations and Concerns

In light of their recent popularity, Large Language Models have brought up a number of legal concerns around information privacy. For example, some aspects of certain LLMs violate the GDPR (General Data Protection Regulation), which is the EU's (European Union) regulation on information privacy. Italy temporarily banned ChatGPT and many other EU nations were considering it. However, the ban was lifted after OpenAI implemented new privacy measures ('ChatGPT vs Privacy'). The EU AI Act classifies AI applications by their risk level. The classification includes three categories: "unacceptable" (banned), "high-risk" (strict legal requirements), and "low-risk" (unregulated) ('ChatGPT vs Privacy'). Here are a few more legal concerns around LLMs. Section 6 will address how to approach these concerns.

User Consent: Big data is generally used without consent. If data used to train LLMs includes sensitive information, the corporation must receive affirmative consent (Busch). Here are consent regulations that apply to LLMs, relevant in the United States and the EU.

- 1) Notice and Disclosure: affirmative consent to use personal data (Busch).
- 2) Opt-Out: the ability to opt out of data collection (Busch).
- 3) Deletion and Minimization: the ability to delete personal data, also known as the "Right to be Forgotten (Francis)" (Busch).
- 4) Data Localization: data must be stored where the user is located (Francis).
- 5) Data Subject Access Requests (DSARs): ability to request access to personal data (Francis).

Regulations 4) and 5) apply specifically to the EU, but are also relevant for global LLM systems that are connected with the EU (Francis). Regulations 3) and 4), which require deletion and access of personal data, are complicated in LLMs, as once the model has been trained, it is difficult to "unlearn" that information or access pre-processed data (Francis).

Data Processing: All applications that involve data processing require a legal basis. In addition to user consent, legal bases for data processing include meeting a contract or legal obligation, saving one's life, or performing a task for legitimate or public interests. User consent is the most common legal basis for data processing, but the user must be aware of how the data is being processed and what they are consenting to. If the scope of data processing changes, user consent must be requested again. There is further regulation around the processing of children's data and the processing of data for profiling purposes. Children's data must be well-protected and profiling can only be performed under consent, authorization, or obligation ('ChatGPT vs Privacy'). Despite heavy regulation on data processing, there is minimal regulation for data scraping. By adding guardrails and limits to how we scrape information, we can target the root of many issues around data privacy and bias (Busch). Guardrails may include preventing copyrighted information and information from untrustworthy sources from being scraped and used to train LLMs. Furthermore, we can introduce limits to measure the quality of a training dataset and ensure that it meets pre-determined industry standards.

Copyrighted Information: As LLMs scrape a large amount of publicly available data, some of this data may be copyrighted (Busch). However, it is difficult to determine whether or not this is fair use due to the ambiguity of copyright. The AI industry relies on the fact that the

Copyright Act does not explicitly prohibit the use of copyrighted material to train LLMs. However, if a generative LLM were to output “memorized” copyrighted text from its training data, as discussed in 3.1.1, this could bring about more direct copyright violations (Moreno). Generally, legislation on privacy considerations in LLMs is tied to personal data. As discussed in 3.1.3, objective privacy harms flow from individual data, but subjective privacy harms do not.

Subjective Privacy Harms: There are currently few regulations for subjective privacy harms. Some potential measures that may be implemented are making explicit definitions as to what defines a “subjective privacy harm” and requiring explicit consent to understand the risks associated with data collection. Furthermore, we can implement digital literacy guidelines to help people become more aware of subjective privacy harms and how to best address them.

3.2: Privacy-Enhancing Techniques (PETs)

3.2.1: Less Effective Solutions

Banning or Limiting Access to LLMs: As mentioned earlier, EU nations such as Italy have banned access to LLMs for short periods of time. However, in banning or limiting access to LLMs, we lose its capabilities, which can be vastly beneficial for various applications. Using this approach is a short-term solution that does not get to the root of the issue.

Privately Owned LLMs: One popular approach to protecting privacy has been the use of privately owned LLMs. If you run an LLM on your own machine, then you can train it with private data without worrying about others accessing that data. Running a private LLM, however, is extremely computationally expensive and leaves the user stuck managing updates. This approach also does not get to the root of the issue, as if the model were to be shared with anyone else, it would be subject to the same data privacy issues (Francis).

3.2.2: Synthetic Data

The idea behind synthetic data is that by removing personal identifiers in the data used to train an LLM, the model will no longer reveal PII. Synthetic data approaches use similar artificial data without the PII in the original data (Anand). For LLM developers, to implement synthetic data, you first use a named-entity recognition (NER) model to identify the PII in a dataset. You then “de-identify” the PII with either redaction, the removal and replacement of PII with placeholder data, or synthesis, the removal and replacement of PII with non-sensitive data. Let’s say an original training example read “John Smith came to our shelter on 8/11/24.” With redaction, the synthetic data would read “{NAME} came to our shelter on {DATE}” and with synthesis, it would read something like “Joe Thatcher came to our shelter on 7/1/23” If you are an LLM user, synthetic data is as simple as replacing any PII in your input prompt with artificial data, either through redaction or synthesis (Ferrara). However, one drawback to synthetic data is that the usage of artificial data may cause the model to lose effectiveness, as it no longer has the same context as before (Francis). The data privacy vault offers a solution to this very issue.

Data Privacy Vault: The data privacy vault, proposed by IEEE, stores sensitive data in an isolated “vault.” The data is stored with respect to the user’s location, in compliance with data localization regulations (see 3.1.4). All other occurrences of sensitive data in the application are de-identified to prevent PII leakage. De-identification occurs during both model training (in the training data) and model usage (in the users’ prompts). Unlike with redaction or synthesis, de-identification in this architecture replaces PII with an identifiable marker that references specific data located in the vault. In this manner, LLM outputs containing de-identified data can be re-identified based on access control policies (which control how and by who the data vault is accessed). Under the data privacy vault, no PII is ever shared directly with the LLM, but the system can still generate contextually relevant and accurate outputs (Francis).

Synthetic data offers a promising straightforward solution to data privacy issues without changing the model itself. However, it is hard to guarantee privacy with this approach, as adversaries may still be able to gain auxiliary information pertaining to individual data (Papertnot, ‘Privacy and Machine Learning’). The next two sections take a more technical approach to solving data privacy by altering the model architecture itself.

3.2.3: Differential Privacy

Differential privacy involves the use of cryptographic algorithms to mask PII (Anand). Although it is slightly less cost-efficient than standard approaches and may slightly decrease model accuracy, it does a great job of maximizing data privacy while minimizing loss in effectiveness (Das et al.). Because LLMs inherently learn specific details in a training dataset, the more queries we make, the higher the likelihood that sensitive information in the training dataset is leaked. To ensure that this data is not leaked, differential privacy obfuscates the outputs of an LLM to mask PII. Differential privacy is characterized by the following definition: given two datasets (D and D') that differ only by a single data point and a mechanism (M) to compute over that data, the output should never differ by more than $\exp(\epsilon)$ where ϵ is the pre-set privacy budget (Testuggine and Mironov). In simpler terms, differential privacy reduces the impact of one individual data point on the overall result (Parmaar). A higher privacy budget reduces the privacy guarantee and a lower budget strengthens privacy, at the potential cost of effectiveness (Testuggine and Mironov). To meet this definition, differential privacy adds fine-tuned noise to model outputs during computation, allowing the final result to differ slightly from the base model, removing specific details such as PII. In this manner, model accuracy is inverse to privacy, as the more noise we add, the results become more secure but also further from the original output. However, in practice, increasing noise to a certain extent has little effect on accuracy (Parmaar). Below, we discuss specific differential privacy methods and how they work.

Differentially-Private Stochastic Gradient Descent (DP-SGD): DP-SGD is based on a mathematical concept called gradient descent, which estimates where a function outputs its lowest value or local minima (‘Gradient Descent’). Applied to ML, gradient descent uses a loss function to solve for which “direction” the model weights should go to minimize the results of the loss function in the next iteration (Tamboli). This value is called a gradient and is associated

with each parameter in the model. DP-SGD modifies the optimization process to ensure privacy. The model parameters are split into “mini-batches,” where the gradient is computed for each sample in the batch. The gradients are then “clipped” to a predefined maximum value to prevent the model from learning too much from any one data point. Finally, noise is added to the largest gradient in the minibatch, which represents the data point with the highest risk of information leakage. The noise value is carefully calibrated to ensure the best balance of privacy and utility (Testuggine and Mironov).

Private Aggregation of Teacher Ensembles (PATE): The PATE approach is based on the idea that models should classify based on patterns in the data, not specifics. It uses an aggregation of several standard ML models to inform decision-making so that you do not need to sacrifice accuracy for privacy. The goal of PATE is that two classifiers trained on different datasets should classify a new input in the same way, therefore not leaking specific information in the process. To accomplish this, a dataset is partitioned to avoid overlap, and a “teacher” model is trained on each partition of the data. When a query is made, the predictions of each teacher are aggregated, alongside the addition of noise to obfuscate specific information, to form a single common prediction. These predictions are used to label “fake data” in a new student model. The student uses the labeled data to train itself and can be published open source to answer queries. In this way, the teacher models and the original data are completely removed from the out-facing model, strengthening the privacy guarantee. The stronger the consensus between the teachers, the stronger the privacy guarantee, as it is unlikely that the prediction was made based on a single training data point. Therefore, by improving the supervised learning procedures in the teacher models, we improve the model’s ability to correctly label data, thereby improving the consensus between teachers and further strengthening privacy (Papernot, ‘Privacy and Machine Learning’).

3.2.4: Homomorphic Encryption (HE)

Homomorphic encryption (HE) is a method that enables computational operation on encrypted data (Anand). With HE, data can be transferred, analyzed, and returned without compromising privacy (‘What Is Multiparty Computation?’). There are three major types of HE: 1) partial HE, which allows only one type of operation, 2) somewhat HE, which restricts the number of operations, and 3) fully HE, which permits more than one type of operation with no restrictions on the number of operations (Anand). We’ll be focusing on Fully Homomorphic Encryption (FHE), the most common type of HE in applications today. FHE occurs during the inference phase of the model, after training has occurred (Bradley). The inputs to the model are encrypted and sent into the LLM. Linear operations such as addition and multiplication can be performed on the encrypted data through homomorphism, a mathematical transformation that maintains the meaning of an input. (‘360 Privacy for Machine Learning’). Nonlinear operations are implemented with a Programmable Bootstrapping (PBS) operation. PBS includes a table lookup (TLU) operation to compute over encrypted data and refreshes the ciphertext to level out the noise (Frery). When the model returns its output, the user can decrypt the data to obtain their

results. Computing over ciphertext increases the computational cost, so we quantize model weights and activations to integers to reduce the complexity of intermediate operations (Frery). Homomorphic encryption can be performed alongside training time methods such as differential privacy, as it occurs only after the model has been trained.

Secure Multiparty Computation: Secure multiparty computation is a subfield of homomorphic encryption that facilitates computation from multiple encrypted data sources (Anand). Each party shares a random portion or “share” of its data with the other parties involved, keeping one share for itself. When each party aggregates the shares of all other parties, they have a piece of data from each party, but no way to decipher each individual data point. Once each aggregation of shares is shared with a third party, computations may be performed over the aggregated data to train an LLM. Then, much like with homomorphic encryption, during the inference phase, encrypted text may be sent to the aggregated LLM to produce safe outputs that do not contain sensitive information from other parties involved. Unlike the methods above, which are primarily targeted at protecting data from external adversaries, multiparty computation also protects data from other participating parties. Much like FHE, multiparty computation reduced the trade-off between data privacy and utility, for even less computational cost (‘What Is Multiparty Computation?’).

4) Bias

4.1: The Origin and Nature of Bias

Bias refers to systematic favoritism or prejudice in decision-making. In the context of LLMs, bias occurs when the model's outputs unintentionally reflect these prejudices. The potential harms of bias in LLMs are significant: they can perpetuate stereotypes, lead to discriminatory practices, and worsen existing social inequalities. As AI is being increasingly used in the industry today, bias in LLMs can have real-world consequences such as unfair treatment in hiring processes, biased law enforcement practices, and unequal access to opportunities (Mahoney et al.). But LLMs are not naturally biased, rather, they pick up on biases present in their training data. Training datasets contain bias because they are gathered from human-generated content. As humans are naturally biased and shaped by our society, bias is present in our language itself, and therefore in Large Language Models as well (Piers). Due to inherent vulnerabilities in AI such as overfitting to specific data points, these biases may even be amplified in model outputs (Ray). For example, models such as ChatGPT tend to be skewed towards Western views and perform best in English (‘Is ChatGPT Biased?’). We as humans, however, have the capability to change our biases once we become aware of them. If we work towards eliminating all bias in LLMs, we not only create an AI tool for fair decision-making, but we also create a tool to help us become less biased at the personal and societal level (Piers).

4.2: Types of Bias

Biases in Datasets: There are three major types of bias that show up in training datasets as classified in the paper “AI Fairness” (Mahoney et al.). Biases present in datasets may be reflected in the outputs further described below.

- **Sample Bias:** With sample bias, one population is either overrepresented or underrepresented.
- **Label Bias:** With label bias, the qualitative or quantitative information associated with a specific data point introduces bias (for example, data that reveals that a certain group of people were less likely to receive job offers than another group).
- **Proxy Bias:** With proxy bias, an ML model uses a specific piece of information to make a decision (for example, using arrests as a proxy for committing a crime).

Biases in Outputs: There are many types of biases that occur in model outputs. Here, we categorize these biases based on their source (Ray).

- **Representation and Favoritism Biases:** Representation biases often occur when a particular group or idea is favored or overrepresented in a training dataset while another is unfavored or underrepresented. Examples include cultural, racial, gender, and linguistic biases, which discriminate against particular groups of people, and ideological and positive/negative sentiment biases, which favor particular ideas, often leaning either overly optimistic or pessimistic on certain topics.
- **Prominence and Trend Biases:** Prominence and trend biases favor certain portions of the dataset. Examples include authority, recency, availability, and outlier biases, which favor information from authorities, current events, available sources of information, and extreme viewpoints respectively. Attention and novelty bias favor popular and trending items in the dataset.
- **Alignment Biases:** Alignment biases tend to base decision-making on alignment with particular information and consensuses in the model’s data. For example, confirmation bias aligns with the training data and content recommendation bias aligns with the user’s known beliefs. Groupthink and false consensus biases generate content that reflects potentially incorrect consensus views.
- **Source Biases:** Source biases emerge from the source of training data. Examples include source bias, whereby data is used from untrustworthy sources, and cognitive bias, whereby data is used from human-generated and potentially biased content.
- **Time-Based Biases:** Time-based biases occur when content is skewed due to the timing of various information. Examples include temporal bias, where content generated is only relevant to specific times, and hindsight bias, which portrays a biased view of historical events.
- **Internet Pattern Biases:** Internet pattern biases emerge when a model copies a behavior learned from internet patterns. Examples include clickbait, sensationalism, and commercial biases.
- **Implicit Bias:** LLMs are capable of implicit bias as well by inferring potentially biased relationships between various ideas and concepts present in the training data.

4.3: Combating Bias in LLMs

4.3.1: Defining Bias and Fairness

It's important to define the various facets of bias in society and the potential harms they create to serve as a baseline for fairness in LLMs (Gallegos et al.). However, bias and fairness are complex topics that manifest in a number of ways and oftentimes depend on the context or culture in which they are brought up. Each organization should take the time to craft its own definitions of bias and fairness by examining what is most important to achieve equitable decision-making for its specific use cases of LLMs (Mahoney et al.). To define bias, organizations should take the time to review the various types of bias present in LLMs (4.2) and identify which are most likely to occur for their application. Each organization should define what biased outputs may look like as a benchmark for what to avoid. Fairness, on the other hand, is often defined in two categories: individual and group fairness. Individual fairness refers to treating individual people or data points that contain similar data similar to one another. This avoids decision-making based on just one or two attributes, emphasizing the importance of considering the individual as a whole. Group fairness prioritizes making the model's predictions and outcomes equitable across groups. This means ensuring that a particular group, such as one that identifies with a particular race or gender, is not statistically favored to have a more positive outcome than another group. Defining group fairness often depends on the organization's philosophy about differences between groups. The WAE (we're all equal) philosophy states that all groups have the same abilities and that any statistical differences between groups depict structural biases in society. On the other hand, the WYSIWYG (what you see is what you get) philosophy states that observations in the data reflect ability and that differences between groups in the data are not always biased. Regardless of the application, it's important to define what the organization prioritizes most in terms of mitigating bias to approach fairness in LLMs from a common angle (Mahoney et al.).

4.3.2: Measuring Bias: Fairness Metrics

Measuring bias manifests in two different ways: bias detection and bias monitoring. Bias detection measures the discrimination and fairness of the dataset and the LLM at a static point in time. It assesses the extent of bias during the development and training phases of model deployment. Bias monitoring, on the other hand, uses the same metrics to measure bias, but does so after the model has been deployed as a continuous process. This is important because bias can emerge or evolve due to changes in the data's distribution or other external factors (Papageorgiou and Mougan). Now we'll cover important terminology around measuring bias in LLMs. Fairness metrics, which quantify unwanted bias in the LLM or its training data, can be applied in both the bias detection and monitoring phases. We quantify unwanted bias or discrimination in LLMs as models that place privileged groups at an advantage over unprivileged groups. Privileged groups are segments of a dataset that are often societally favored such as personal demographics like race and gender or ideas and beliefs reflected by various media platforms. Privileged groups are

often characterized by favorable labels in the dataset, which are values that correspond to an advantage in the model outcome. These values, often referred to as privileged values, historically have a systemic advantage. To preserve fairness, we define protected attributes as partitions of the population that should be treated equally such as race, gender, and ethnicity (Mahoney et al.). Now we will cover the various types of fairness metrics as classified in the paper “Bias Mitigation for Machine Learning Classifiers: A Comprehensive Survey” (Hort et al.).

- **Dataset Metrics:** Dataset metrics are based on labels assigned in the training dataset. They determine the degree of bias in the underlying data, often between privileged and unprivileged groups. The Mean Difference algorithm, for example, calculates the difference in positive labels between population groups.
- **Parity-Based Metrics:** Parity-based metrics are based on predicted outcomes and focus on whether or not different groups are classified equitably. One popular algorithm, demographic parity, measures whether or not privileged and unprivileged groups receive an equal percentage of positive labels.
- **Outcome-Based Metrics:** Outcome-based metrics compare how well the model generates outcomes for privileged versus unprivileged groups. Some of these metrics are based on the predicted outcomes of the LLM and the actual outcomes backed by the data. For example, the confusion matrix compares true positive, false positive, false negative, and true negative rates between groups. Other metrics are based on predicted probabilities of the model and actual outcomes backed by the data. Strong pairwise demographic parity, for example, measures the difference in probability between groups that the predicted probabilities are above certain thresholds.
- **Similarity-Based Metrics:** Similarity-based metrics meet individual fairness definitions (see 4.3.1), ensuring that statistically similar individuals receive similar outcomes. Consistency metrics, for example, compare the outcomes of an individual with the outcomes of similar individuals. The similarity of individuals is calculated based on their associated data, disregarding protected attributes.
- **Causal Reasoning:** Causal graphs evaluate relationships between protected attributes and outcomes, ensuring that decision-making is not discriminatory or determined by protected attributes. Normalized Prejudice Index, for example, measures information that may be in common between protected attributes and outcomes.

4.3.3: Reducing Bias: Mitigation Techniques

After measuring bias, the final step is to apply algorithms to mitigate bias. The goal of this step is to balance the accuracy of the model with minimizing bias against less privileged groups represented in the data (Fernandez). Bias mitigation techniques occur in three different places in the ML pipeline: pre-processing, to modify the data used to train the model, in-processing, while the model is being trained, and post-processing, to modify the model outputs (Mahoney et al.). In-processing techniques are specific to the model they are used in while pre-processing and post-processing techniques are independent of the model and act solely

on the inputs and outputs. The data collection phase, before pre-processing, is one area that has received less focus in terms of bias mitigation. By adopting data collection guidelines (further discussed in 3.1.4) that prevent biased and prejudiced data from being collected, keeping only bias-free data, we can target the cause of bias in LLMs at its source. For the most comprehensive approach to bias mitigation, multiple techniques may be used in conjunction throughout multiple phases (Hort et al.). Now, we'll dive into some specific bias mitigation techniques, as classified in the paper "Bias Mitigation for Machine Learning Classifiers: A Comprehensive Survey".

Pre-Processing Algorithms: Pre-processing algorithms aim to mitigate the bias that is intrinsic to the training data itself. This is a difficult task as bias can show up in data in numerous complex ways. Furthermore, there are legal complications around modifying data (Mahoney et al.). The following are various types of pre-processing algorithms (Hort et al.).

- **Relabelling:** Relabelling changes the ground truth labels in the dataset to construct a more balanced and equitable dataset across groups. This is often done in instances close to the decision boundary, where classification could go either way (referred to as "massaging"). Relabelling is also performed to meet individual fairness definitions that similar individuals should receive similar outcomes.
- **Perturbation:** Perturbation modifies other features in the dataset to ensure that non-protected attributes for privileged and unprivileged groups are similar, through algorithms such as the Disparate Impact Remover. Perturbation is often an iterative process aimed at changing the most bias-prone attributes in a dataset.
- **Sampling:** Sampling changes the distribution of data points in a number of different ways. Reweighting modifies the impact of data points on training procedures for a more equitable data distribution. For example, the model could give higher weight to instances with positive outcomes for unprivileged groups. Upsampling and downsampling add or remove data points. For example, the model could upsample minority groups with synthetic data (see 3.2.2).
- **Representation Learning:** Representation learning involves training an ML model to modify a training dataset to reduce bias while retaining maximum information from the original dataset. Learning Fair Representation, for instance, optimizes for removing information about protected attributes while keeping as much information regarding non-protected attributes as possible. Other applications are trained to remove discriminating features from the dataset and augment the training data with new unbiased features.

In-Processing Algorithms: In-processing algorithms introduce various constraints to ensure fairness by influencing various aspects of the model during its training process (Mahoney et al.). Now, we'll cover different types of in-processing algorithms (Hort et al.).

- **Regularization and Constraints:** These methods apply changes to the loss function, a key portion of the training process of ML models. Regularization adds terms to the loss function to penalize discriminatory decision-making. The Prejudice Remover, for instance, reduces overfitting by mitigating the relationship between protected attributes

and other information. Constraints define the degree of bias that cannot be reached during training according to the loss function. The Meta-Fair Classifier, for example, takes a fairness constraint as input for a meta-algorithm to optimize constraints. Regularization and constraint methods are highly dependent on the definitions each organization sets for bias and fairness priorities in their application.

- **Adversarial Learning:** Adversarial learning models are trained alongside the primary LLM to exploit bias and fairness issues. Both models compete with and learn from each other throughout the process. In Adversarially Reweighted Learning, for example, the adversary adjusts the weights of specific regions in the input area with the highest loss, informing changes in the primary model as well.
- **Compositional Classification:** Compositional approaches involve the use of multiple classification models, either through classifying different groups within different models or using an ensemble approach similar to PATE (see 3.2.3), where all models vote on classification. With the former approach, decoupled classifiers are trained primarily on the data of a specific group to classify that group without bias and with high accuracy. With the latter, multiple models with varying fairness definitions are used in an ensemble fashion to generate a Pareto-front of solutions (discussed in depth in section 5) and select the best trade-off between all fairness objectives.
- **Adjusted Learning:** Adjusted learning methods either modify learning algorithms or create new learning algorithms. The Approximate Projection onto Star Sets (APStar) method, for example, trains deep neural networks, like LLMs, to maximize fairness among all groups involved and find the best trade-off with the Pareto frontier (see section 5 for more).

Post-Processing Algorithms: Post-processing algorithms work best when you do not have access to the model itself. They manipulate the outputs of the model after training procedures, sometimes at the cost of model accuracy (Mahoney et al.). The following are various types of post-processing algorithms (Hort et al.).

- **Input Correction:** Input correction modifies the data that you input into the LLM similar to how you would modify training data with pre-processing modifications like relabelling, perturbation, and representation learning.
- **Classifier Correction:** Classifier correction modifies classification models to maximize fairness. Approaches include optimizing a classifier with respect to various fairness metrics and splitting the classifier for privileged and unprivileged groups.
- **Output Correction:** Output correct modifies the outputs of the LLM. These approaches are also similar to pre-processing approaches. Examples include favoring unprivileged groups at the decision boundary and modifying outputs most likely to receive biased outcomes.

5) Pareto Frontier

Throughout this paper, we've discussed three primary objectives for an LLM: privacy (section 3), fairness (section 4), and accuracy. We now explore a way to solve for all three of these objectives equally. The Pareto frontier is a set of solutions that represents the best trade-off between all objectives. It uses a graph with a variable amount of dimensions for each objective (in our case, three), to plot the best solutions for each objective. The solutions present on the Pareto frontier are such that improving one objective cannot be done without worsening another (Datta). Most approaches today to combine privacy and fairness are a combination of privacy-enhancing techniques such as differential privacy (section 3) and bias mitigation methods (section 4). However, with this combined approach, one objective is often optimized, while the other is given less focus. The Pareto frontier solves this issue with the principle of impartiality which states that the model architecture should not favor one objective to another. To do so, privacy and fairness must be optimized simultaneously, as opposed to in different parts of the ML pipeline (Yaghini et al.).

Two novel algorithms were introduced in the paper "Learning to Walk Impartially on the Pareto Frontier of Fairness, Privacy, and Utility" (Yaghini et al.). FairDP-SGD extends DP-SGD with the Demographic Parity Fairness Regularizer, which depends on the fairness violations over a dataset (Yaghini et al.). FairPATE extends PATE's GNMax algorithm, which rejects queries if they have high privacy costs (Yaghini et al.). Confidential&Fair-GNMax also rejects queries based on their impact on fairness. At low privacy and fairness budgets, where mitigating information leakage and bias is more important, these approaches outperform the baseline model for utility and cost. At higher privacy and fairness budgets, where mitigating information leakage and bias is less important, these approaches perform similarly to the baseline model. FairPATE generally outperforms FairDP-SGD. FairPATE is especially better with higher privacy budgets and lower fairness budgets, but the models perform similarly with lower privacy budgets and higher fairness budgets (Yaghini et al.). The applications of the Pareto frontier in the world of AI are just scratching the surface and are an interesting field of research for the future.

6) Nonprofit Guide

6.1: Implementing Data Privacy and Bias Mitigation Techniques

6.1.1: Data Privacy

In this section, we cover the best use cases for synthetic data, differential privacy, and homomorphic encryption, along with their sub-methods such as the data privacy vault, DP-SGD, PATE, and secure multiparty computation.

Synthetic Data: Synthetic data is best used for smaller-scale use cases and AI applications, as it is less computationally expensive and more simple to implement than the following approaches. It's good practice to use synthetic data software when inputting sensitive information into LLMs like ChatGPT or designing a simple or small-scale application that involves sensitive data such as a chatbot that implements RAG. For slightly larger-scale

applications in which the organization wants to avoid the more complex methods below, the data privacy vault is a great approach, keeping all sensitive data safe throughout the entire software pipeline. Overall, synthetic data is a great choice for teams without extensive technology backgrounds looking to get involved in the world of AI.

Differential Privacy: Differential privacy is a strong approach for more advanced applications with an abundance of back-end data. It's also great for applications for which keeping the original data intact is important. These algorithms can often be implemented with third-party packages, but we recommend this approach for teams with a background in technology. DP-SGD and PATE often produce similar results, but we recommend DP-SGD for tasks that involve generating content and PATE for tasks that involve classifying outcomes.

Homomorphic Encryption: Homomorphic encryption is similar to data privacy in terms of use cases and third-party packaging. It's great for teams with a technical background developing more advanced applications with large amounts of data. One key difference is that all information processed by LLMs is encrypted with this approach and then decrypted after output. With HE, there is more feasibility to control whether or not the final output contains PII or not, similar to the data privacy vault and its access control policies. We recommend secure multiparty computation as a strong approach for collaborative projects between multiple organizations.

6.1.2: Bias Mitigation

As described in 4.3, bias mitigation techniques are highly dependent on the organization and the use case of AI. We recommend the architecture proposed in 4.3 that begins with defining bias and fairness, measuring bias with metrics, and applying bias mitigation algorithms. Teams with smaller-scale applications, looking for simplicity, may find that using third-party bias mitigation software is the best solution. Teams with more technical background looking to develop specialized applications with AI may find that specifying their own priorities and implementing their own unique approach to bias mitigation is best for them.

6.2: More Concerns and Limitations with LLMs

Concerns: Apart from the major topics covered in this paper—data privacy and bias—there are a number of other concerns with LLMs to be aware of for organizations planning to implement AI. Some of these concerns are practical issues, while others are ethical conundrums. In terms of practical issues, the next largest concern is hallucinations. As LLMs are not one hundred percent accurate, they are susceptible to unreliable outputs and may sometimes output outright incorrect information or “hallucinations.” Furthermore, when these tools are used irresponsibly, unintended consequences may arise that lead to a loss of reputation or trust in the organization. One last practical concern is the fact that larger corporations like OpenAI do not release their data sources, which may be an issue for organizations that want to know how the model they are using has been trained (Bommasani et al.). In terms of ethical issues, one common concern is the concentration of power in the hands of just a few big tech companies in the age of generative AI. Involving vast amounts of data, this often promotes data “quantity over

quality” and in some cases, unlawful access of internet data (Bommasani et al.). Another issue is the computational costs of AI applications and their potential environmental harm. Looking more into the future, there is also a lingering fear that AI will replace humans and that one day AI might become sentient and more intelligent than us (Gordon).

Limitations: Although LLMs are quite powerful and diverse in terms of their applicability, there are some limitations to be aware of when using these tools. We discuss these limitations so organizations can be aware of when not to over-rely on AI in their work. Firstly, LLMs are sensitive to input phrasing and may have difficulty answering ambiguous prompts. At times, their outputs can be either too vague or overly verbose. They also struggle with personalizing outputs to the user, handling multilingual queries, and navigating long conversations effectively. They have limited domain-specific knowledge as they often lack access to external resources. As discussed in the “Concerns” section, their quality is inconsistent at times and they often require significant computational costs. Finally, LLMs have a number of inherent issues such as limited contextual awareness, moral reasoning, emotional intelligence, creativity, human intuition, self-awareness, and specificity. However, as LLMs continue to develop, we believe that they will improve on a number of these fronts. LLMs have already broken through a number of limitations of the past such as the inability to check facts, produce visual content, and access real-time data (Ray).

6.3: Recommended Guidelines and Standards

6.3.1: Safe Use Guidelines

In this section, we cover some guidelines that organizations can follow to approach AI implementation safely and effectively.

1. **Stay Organized:** It’s important to stay organized as you think about how you want to implement AI into your work. By setting goals, choosing the right tools, conducting thorough testing, and revising as you go, your organization can stay core to its mission and remain effective (Schohl).
2. **Staff Training:** When you start to integrate AI tools into the workflow, it’s important to train staff to be aware of its potential good and harm (‘Quick Guide’). As some people tend to fear using AI, the hardest part here is learning to contain people’s fears and helping them feel comfortable using the technology safely and effectively (Schohl).
3. **Vet Vendors:** If your organization is using third-party vendors, it’s important to ensure that they are reputable and do a good job at ensuring data privacy and mitigating bias (‘Quick Guide’).
4. **Data Management:** With any AI application, it’s essential to manage data wisely. Ensure that the underlying data you use for LLM tools are high quality, up-to-date, and bias-free (‘Quick Guide’). Also, be sure to avoid inputting sensitive information directly into LLMs. Instead, anonymize that PII with synthetic data approaches beforehand (there are third-party tools available that do this for you) (Graciano).

5. **Stay Human-Centered:** Perhaps the most important takeaway here is to remain human-centered in your work with AI. Although AI does replicate human reasoning in many ways, it's far from perfect. It's essential to retain human oversight by reading over and editing AI outputs, ensuring that they are aligned with the core values of your organization ('Quick Guide'). Instead of a replacement for humans, AI should be seen as a tool to help augment our work (Schohl).
6. **Take It Slow:** Because generative AI is relatively new, you should take AI implementation slow at first, especially if your organization is new to this technology. By lowering the stakes first to truly understand AI and how it can be used, you can develop more useful, and safe, AI applications (Schohl).

6.3.2: Industry Standards

As the field of generative AI is rising, nonprofits are in a unique position to lead by example and establish norms and standards for the industry to approach safe AI implementation for social good. By publishing social impact reports that include details on bias and privacy and open-sourced documentation for your AI applications, you can remain transparent with the public on how you use AI. Some proposals have been made for labels for each model that qualify and quantify the data, the model itself, alongside its known capacities, weaknesses, and biases. As development is an iterative process, setting up a structure to collect feedback is also a great approach to improving your tools. Before implementing or releasing your applications, it's important to establish auditing policies to test them and ensure that they achieve their goals. Finally, one avenue for future norms is establishing guidelines on when not to use AI. Organizations should ensure that their applications meet community needs and ethical standards and do more good than harm (Bommasani et al.).

6.3.3: Legal Compliance

As discussed in 3.1.3, there are numerous legal concerns with LLM applications. In this section, we cover how to comply with those regulations. First and foremost, it's essential that your organization's employees become aware of safe and legal AI use (Graciano). In line with copyright laws, your organization should set guidelines on what qualifies as original work when LLM tools are used to generate content (Wong). In terms of consent, you must receive customer consent when using their data in your applications. The customer must be aware of what they are complying with and how their data is being processed, giving them the option to opt out at any time. Finally, for data processing, create data processing agreements and ensure that all processing has a legal basis (most commonly user consent). Provide extra protection for children's data and only conduct profiling under explicit consent. Data Protection Impact Assessments (DPIAs) are useful documents that identify and minimize privacy risks. They detail what data you are processing, how it is being processed, whether or not PII is used, and whether or not it is sent to third parties ('ChatGPT vs Privacy').

6.4: Empowering Nonprofits with AI

Now that we've covered some important considerations and guidelines for AI use, we now explore some of the ways AI can be applied to empower nonprofit organizations. As nonprofits have limited resources and staff, AI and LLM technology has the potential to significantly boost efficiency. Nonprofits often have funding issues, and struggle to retain donors and recruit volunteers. Furthermore, complex legal regulations make it harder to innovate new ideas. AI has the ability to solve these issues (Khawaja). Here are some ideas to get your mind rolling on how to apply AI in the nonprofit sector:

- Fundraising: AI can help you write and edit grants (Khawaja). It can also help plan and write content for fundraising events (Schmid).
- Community Inquiries/Customer Support: Using RAG or other LLM tools, it's quite simple to design a specialized chatbot with your company data to respond to community inquiries, provide customer support, or raise awareness for your cause (Schmid).
- Compliance: Real-time AI tools can help you stay up to date on new regulations that are relevant to your AI applications (Khawaja).
- Spreading Awareness/Marketing: AI can help you generate and edit long-form content in the form of blog posts or short-form content in the form of social media (Schmid). It can also be used to collect customer data to inform marketing strategy (Graciano).
- Personalization/Sales: AI can be used to personalize product recommendations and give purchasing assistance (Graciano).
- Volunteer/Donor Relations: AI tools can help organize databases of stakeholders, segmenting various types of donors and volunteers. They can also help create personalized appeals and thank-you notes to establish long-term donor relationships. AI can be used for volunteer management to organize events and keep track of participants (Schmid).
- Reports and Statistics: AI can be used to generate reports and statistics by looking for trends in the provided data and locating more data from online sources (Schmid).

7) Conclusion

We hope that our work has laid the foundation to empower nonprofit organizations to use AI both safely and productively. We tackle two primary concerns in LLMs (data privacy and bias), explaining all their associated risks and offering strategies to mitigate them. We conclude with a section relevant to the nonprofit sector that takes research into application, offering strategies, limitations, and guidelines for AI use. Future steps for our work include implementing these guidelines and information to inform AI development at shelters for unhoused youth. As we dive into a rapidly changing technological landscape, it is imperative that we stay on top of the latest advancements while remaining aware of their potential limitations by continuing research on the development of fair, safe, and effective AI systems.

Works Cited

- ‘360 Privacy for Machine Learning’. *Zama*.
<https://www.zama.ai/post/360-privacy-for-machine-learning-with-homomorphic-encryption>.
- Anand, Chetan. ‘More Responsible Data Usage Through Privacy Enhancing Technologies’.
<https://www.isaca.org/resources/news-and-trends/isaca-now-blog/2024/more-responsible-data-usage-through-privacy-enhancing-technologies>.
- Anil, Rohan, et al. ‘Gemini: A Family of Highly Capable Multimodal Models’. 2023,
<https://doi.org/10.48550/arXiv.2312.11805>.
- Bommasani, Rishi et al. ‘On the Opportunities and Risks of Foundation Models’. 2022,
<https://doi.org/10.48550/arXiv.2108.07258>.
- Bradley, Jeremy. ‘The (r)Evolution of FHE’. <https://www.zama.ai/post/the-revolution-of-fhe>.
- Brown, Tom B., et al. ‘Language Models Are Few-Shot Learners’. 2020,
<https://doi.org/10.48550/arXiv.2005.14165>.
- Busch, Kristen E. ‘Generative Artificial Intelligence and Data Privacy: A Primer’. *Congressional Research Service*, no. R47569, 2023.
- Cardillo, Anthony. ‘How Many Companies Use AI? (New Data)’.
<https://explodingtopics.com/blog/companies-using-ai>.
- ‘ChatGPT vs Privacy—How Concerned Should We Be?’. *Osano*.
<https://www.osano.com/articles/chatgpt-privacy>.
- Das, Badhan C., et al. ‘Security and Privacy Challenges of Large Language Models: A Survey’.
2024, <https://doi.org/10.48550/arXiv.2402.00888>.
- ‘Data Protection and ChatGPT Use’. University of California, Santa Barbara.
<https://it.ucsb.edu/news/data-protection-and-chatgpt-use>.
- Datta, Subham. ‘Defining Multiobjective Algorithms and Pareto Frontiers’.
<https://www.baedlung.com/cs/defining-multiobjective-algorithms-and-pareto-frontiers#:~:text=The%20Pareto%20frontier%20is%20a,be%20on%20the%20Pareto%20frontier>
- Devlin, J., et al. ‘BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding’. 2018, <https://doi.org/10.48550/arXiv.1810.04805>.
- Dodge, Jesse, et al. ‘Fine-Tuning Pretrained Language Models: Weight Initializations, Data Orders, and Early Stopping’. 2020, <https://doi.org/10.48550/arXiv.2002.06305>.
- Fernandez, Franklin Cardenoso. ‘Bias Mitigation Strategies and Techniques for Classification Tasks’.
<https://www.holisticai.com/blog/bias-mitigation-strategies-techniques-for-classification-tasks>.
- Ferrara, Joe. ‘Safeguarding Data Privacy While Using LLMs’.
- Francis, Alexandra. ‘Privacy in the Age of Generative AI’.
<https://stackoverflow.blog/2023/10/23/privacy-in-the-age-of-generative-ai/>.
- Frery, Jordan. ‘Towards Encrypted Large Language Models with FHE’.
<https://huggingface.co/blog/encrypted-llm>.
- Gallegos, Isabel O., et al. ‘Bias and Fairness in Large Language Models: A Survey’.
Computational Linguistics, vol. 50, no. 3, 2023, <https://doi.org/10.48550/arXiv.2309.00770>.
- Gordon, Andy. ‘10 Reasons Nonprofits, Charities and Purpose-Led Organisations Shouldn’t Use Large Language Models’.
<https://ai.torchbox.com/thinking/2023-04-21-ten-reasons-to-say-no-to-llms>.

Graciano, Patricia. ‘ChatGPT Privacy Concerns: How to Leverage LLMs While Maintaining Data Privacy’. <https://www.private-ai.com/en/2023/05/01/chatgpt-privacy-concerns/>.

‘Gradient Descent’. *Khan Academy*.
<https://www.khanacademy.org/math/multivariable-calculus/applications-of-multivariable-derivatives/optimizing-multivariable-functions/a/what-is-gradient-descent>.

Hort, Max, et al. ‘Bias Mitigation for Machine Learning Classifiers: A Comprehensive Survey’. 2023, <https://doi.org/10.48550/arXiv.2207.07068>.

‘Introducing the next Generation of Claude’. *Anthropic*.
<https://www.anthropic.com/news/claude-3-family>.

‘Is ChatGPT Biased?’. *OpenAI*. <https://help.openai.com/en/articles/8313359-is-chatgpt-biased>.

Monique Ogburn, et al. ‘Homomorphic Encryption’. *Procedia Computer Science*, vol. 20, 2012, pp. 502–509, <https://doi.org/10.1016/j.procs.2013.09.310>.

Khawaja, Ruhma. ‘Empowering Non-Profit Organizations – The Game-Changing Potential of Generative AI and LLMs’. <https://datasciencedojo.com/blog/generative-ai-for-non-profit/>.

Mahoney, Trisha, et al. *AI Fairness*. O’Reilly Media, Inc., 2020,
<https://krvarshney.github.io/pubs/MahoneyVH2020.pdf>.

Moreno, Ivan. ‘AI Models’ Link To Nonprofit Data Raises Fair Use Question’.
<https://www.morganlewis.com/-/media/files/news/2024/ai-models-link-to-nonprofit-data-raises-fair-use-question-law360.pdf>.

Muñoz, Eduardo. ‘Attention Is All You Need: Discovering the Transformer Paper’.
<https://towardsdatascience.com/attention-is-all-you-need-discovering-the-transformer-paper-73e5ff5e0634>.

Papageorgiou, Ioanna, and Mougan, Carlos. ‘Necessity of Processing Sensitive Data for Bias Detection and Monitoring: A Techno-Legal Exploration’. *NeurIPS 2023 Workshop on Regulatable ML*, 2023, <https://openreview.net/forum?id=inZoxsEpYn>.

Papernot, Nicolas, et al. ‘Scalable Private Learning with PATE’. *ICLR*, 2018,
<https://doi.org/10.48550/arXiv.1802.08908>.

Papernot, Nicolas, and Ian Goodfellow. ‘Privacy and Machine Learning: Two Unexpected Allies?’ <https://www.cleverhans.io/privacy/2018/04/29/privacy-and-machine-learning.html>.

Parmaar, Neeraj Rajkumar. ‘Differential Privacy in Deep Learning’.
<https://towardsdatascience.com/differential-privacy-in-deep-learning-cf9cc3591d28>.

Piers, Craig. ‘Even ChatGPT Says ChatGPT Is Racially Biased’.
<https://www.scientificamerican.com/article/even-chatgpt-says-chatgpt-is-racially-biased/>.

‘A Quick Guide to Using AI Responsibly at Your Nonprofit’. *Soapbox Engage*.
<https://www.soapboxengage.com/blog/2492-a-quick-guide-to-using-ai-responsibly-at-your-nonprofit>.

Radford, A., et al. ‘Language Models Are Unsupervised Multitask Learners’. *OpenAI Blog*, 2019.

Ray, Partha Pratim. ‘ChatGPT: A Comprehensive Review on Background, Applications, Key Challenges, Bias, Ethics, Limitations and Future Scope’. *Internet of Things and Cyber-Physical Systems*, vol. 3, 2023, pp. 121–154, <https://doi.org/10.1016/j.iotcps.2023.04.003>.

‘Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks’. *NeurIPS*, 2020,
<https://doi.org/10.48550/arXiv.2005.11401>.

- Schmid, Kara. ‘ChatGPT for Nonprofits: Maximizing Impact with the Power of AI’.
<https://donorbox.org/nonprofit-blog/chatgpt-for-nonprofits>.
- Schohl, Lisa. ‘How Nonprofits Can Use A.I. Well — and Avoid Pitfalls’.
<https://www.philanthropy.com/article/how-nonprofits-can-use-a-i-well-and-avoid-pitfalls>.
- Song, Shuang, et al. ‘Stochastic Gradient Descent with Differentially Private Updates’. 2013
IEEE Global Conference on Signal and Information Processing, 2013, pp. 245–248,
<https://doi.org/10.1109/GlobalSIP.2013.6736861>.
- Tamboli, Nasima. ‘Understanding Gradient Descent Algorithm and the Maths Behind It’.
<https://www.analyticsvidhya.com/blog/2021/08/understanding-gradient-descent-algorithm-and-the-maths-behind-it/>.
- Testuggine, Davide, and Mironov, Ilya. ‘Differential Privacy Series Part 1 | DP-SGD Algorithm Explained’.
<https://medium.com/pytorch/differential-privacy-series-part-1-dp-sgd-algorithm-explained-12512c3959a3>.
- Touvron, Hugo, et al. ‘LLaMA: Open and Efficient Foundation Language Models’. 2023,
<https://doi.org/10.48550/arXiv.2312.11805>.
- Valentino-DeVries, Jennifer. ‘Q&A: How Do You Define ‘Privacy Harm’?’
<https://www.wsj.com/articles/BL-DGB-17390>.
- Vaswani, Ashish, et al. ‘Attention Is All You Need’. 2023, <https://doi.org/10.48550/arXiv.1706.03762>.
- ‘What Is Multiparty Computation?’ . *IEEE Digital Privacy*
<https://digitalprivacy.ieee.org/publications/topics/what-is-multiparty-computation>.
- Williams, J., and L. Nee. ‘Privacy Engineering’. *Computer*, vol. 55, no. 10, IEEE Computer Society, Oct. 2022, pp. 113–118, <https://doi.org/10.1109/MC.2022.3192948>.
- Wong, Stephanie. ‘Let’s Look at LLMs: Understanding Data Flows and Risks in the Workplace’.
<https://fpf.org/blog/lets-look-at-llms-understanding-data-flows-and-risks-in-the-workplace/>.
- Yaghini, Mohammad, et al. ‘Learning to Walk Impartially on the Pareto Frontier of Fairness, Privacy, and Utility’. *NeurIPS 2023 Workshop on Regulatable ML*, 2023,
<https://openreview.net/forum?id=R5MTSLPyYZ>.
- Yao, Yifan, et al. ‘A Survey on Large Language Model (LLM) Security and Privacy: The Good, the Bad, and the Ugly’. 2023, <https://doi.org/10.48550/arXiv.2312.02003>.
- Yew, Rui-Jie, et al. . ‘You Still See Me: How Data Protection Supports the Architecture of ML Surveillance’. *NeurIPS Workshop on Regulatable ML*, 2023,
<https://doi.org/10.48550/arXiv.2402.06609>.
- Zhao, Chaun, et al. ‘Secure Multi-Party Computation: Theory, Practice and Applications’.
Information Sciences, vol. 476, 2019, pp. 357–372,
<https://doi.org/10.1016/j.ins.2018.10.024>.

Nietzsche's Nihilism and How to Lead a Meaningful Life

To What Extent Is Nietzsche's Theory on Finding Meaning in Life True and Applicable to Our Modern Lives? By Benjamin Ooi

Abstract:

In this essay I will explore Nietzsche's concept of the Übermensch and its relevance today. In *Thus Spake Zarathustra*, Nietzsche calls for creating our own meaning in a world with no inherent purpose, encouraging us to transcend societal and religious norms. I will compare his argument to modern theories like Maslow's Hierarchy of Needs and will contrast it with critiques from Tolstoy, as well as other philosophies like Hedonism and Asceticism. Through this analysis I will argue that Nietzsche's focus on self actualization and individual will is increasingly relevant from a modern perspective, as a result of a decline of traditional religious morals, a rise of technology addiction, and an overemphasis on political conformity. I will conclude that Nietzsche's philosophy provides a solution to modern nihilism, urging us to reject external moral authorities and pursue authentic self-fulfillment, guiding us to become the Übermensch in a society that encourages conformity.

Introduction

In Friedrich Nietzsche's *Thus Spake Zarathustra*, he concludes that in lives with no inherent meaning, instead of succumbing to the apathy of nihilism, we must create our own meaning and become the "Übermensch".⁷⁹

How does this theory align with current psychology, and is it actually feasible in our lives today? In this essay, I will argue that Nietzsche's theory is supported by the psychological concept of wants and needs. I will then examine how Nietzsche would advise us to avoid the pitfalls of modern day society and live a life with meaning and purpose.

In the aftermath of COVID, existential anxiety seems to be on the rise.⁸⁰ This phenomenon is indicative of a broader issue about the growing confusion and lack of direction that many are experiencing.⁸¹ Factors contributing to this issue include technological advancements, like the internet and cellphones, which constantly reshape our lives.⁸² Economic instability and a widening gap between the ultra-wealthy and the impoverished leads to feelings of alienation and separation. Additionally, social media heightens individuals' feelings of

⁷⁹ Friedrich Wilhelm Nietzsche, Adrian Del Caro, and Robert B. Pippin, *Thus Spoke Zarathustra: A Book for All and None*, Cambridge Texts in the History of Philosophy (Cambridge ; New York: Cambridge University Press, 2006). 7-8

⁸⁰ John Cottone, "The Profound Challenge of Existential Anxiety," *Psychology Today*, 2 Dec. 2022, <https://www.psychologytoday.com/us/blog/the-cube/202212/the-profound-challenge-of-existential-anxiety>.

⁸¹ Jack Kelly, "Why Is It So Hard for Recent College Graduates to Find a Decent Job?" *Forbes*, February 15, 2024, <https://www.forbes.com/sites/jackkelly/2024/02/15/why-is-it-so-hard-for-recent-college-graduates-to-find-a-decent-job/>.

⁸² Emily A. Vogels, Risa Gelles-Watnick, and Navid Massarat, "Teens, Social Media and Technology 2022," *Pew Research Center*, August 10, 2022, <https://www.pewresearch.org/internet/2022/08/10/teens-social-media-and-technology-2022/>.

isolation and inadequacy. This angst requires a deeper look at its causes and the underlying issues plaguing society.

The Death of God

In his book *Thus Spake Zarathustra*, Nietzsche declares that “God is dead”.⁸³ Nietzsche claims that as science progresses, God becomes less and less important as an explanatory force, and therefore is losing his authority over society. Assuming that objective morals cannot exist without God, by killing God, we are also killing a source of meaning in our lives. A good Christian of the past might have proudly given aid to the poor, knowing that “Whoever is kind to the poor lends to the Lord, and he will reward them for what they have done”.⁸⁴ By offering a reward, God artificially adds value to this principle and attempts to guide the actions of his followers. Conversely, someone without an external source of morality would have to feel compelled on their own to give aid. With the advancement of science, God as an explanatory force has lost his credibility which Nietzsche argues will lead to people becoming lost.

Nietzsche’s declaration is even more relevant today. Church attendance (and attendance in other religious spaces) has steadily declined since data was first collected in the 1940s.⁸⁵ While in the past most people were diligently practicing a religion, today, only a small minority incorporate God into their daily lives, despite the fact that most people still identify as religious. The difference between active participation and passive identification is crucial. This loss of religious agency has caused a separation between religion and society, resulting in the death of most traditional guidance in life. In the past religion extended beyond individual practice and was deeply ingrained in every part of society. Nietzsche imagined a world like this one, where traditional values disappeared, leading to this idealless void. As a result, humans would be forced to create their own meaning for themselves, because there would no longer be an intrinsic purpose for them to fulfill.

However, Nietzsche actually saw this loss of meaning as positive for society, as he believed Christian universal maxims were limiting and inauthentic to the individual. In his view, these societal norms merely controlled us, eliminating creativity and individuality. These norms promoted the herd mentality. Nietzsche believed that true greatness could be found when breaking free from these norms, imposing one’s own will, and creating personal meaning. This process is at the center of his theory about the *Übermensch*.

Finding Meaning in a Meaningless World:

Rather than give up in the overwhelming face of nihilism, Nietzsche offers a solution: The creation of meaning through the pursuit of the *Übermensch*, or overman. The *Übermensch* is a person who transcends the will of conventional society and religion, asserting their individual desires.⁸⁶ After realizing the arbitrariness of the values and ideals of others, the *Übermensch* sees

⁸³ Nietzsche, Del Caro, and Pippin. 1-5

⁸⁴ *The Holy Bible: New International Version* (London: Hodder and Stoughton, 1980).

⁸⁵ Jeffrey M. Jones, "U.S. Church Membership Falls Below Majority for First Time," *Gallup*, March 29, 2021, <https://news.gallup.com/poll/341963/church-membership-falls-below-majority-first-time.aspx>.

⁸⁶ Nietzsche, Del Caro, and Pippin. 7-8

an opportunity for freedom. According to Nietzsche, the Übermensch does not seek validation from others and is driven purely by their Will to Power, which often expresses itself as a furthering of one's creativity and authentic desires.

Nietzsche vs Other Philosophers

Tolstoy's Critique of Moral Relativism

One prominent critic of Nietzsche was Leo Tolstoy, who disagreed with many of his arguments surrounding religion in particular. In his 1894 book *The Kingdom of God is Within You*, he analyzed several aspects of Nietzsche's philosophy, including the idea that God is dead, and the ideas of the Übermensch and the Will to Power. To start, Tolstoy believed that Christianity was not a set of principles and ethics, but a path to spiritual fulfillment, which he believed necessary to the human condition.⁸⁷ Regardless of religion's "true" purpose, both philosophers agree that it provides guidance for its followers. Therefore, their critical disagreement is whether or not God is dead. If Tolstoy were able to see the world today, he would likely acknowledge how God has lost his influence in society.

Tolstoy also criticized the theory of the Übermensch as individualistic, which reflects his differing beliefs about the relativity of morals. After arguing that there is no purpose for humans in life, Nietzsche contends that in order to be fulfilled and live a meaningful life, we must become the Übermensch. Tolstoy criticized this ethos as extremely selfish, instead believing in social responsibility, and arguing that the true sign of greatness is one who acts selflessly out of service to others.⁸⁸ Ultimately, their disagreement derives from a difference in each philosopher's opinions about the relativism of morals. Tolstoy thinks that morals are objective because they come from a religious source. Given that this paper is examining a current context where God is no longer a credible source, Tolstoy's argument isn't applicable. Outside of Tolstoy, there are many arguments for objective morality, but this paper will proceed under the assumption that morality is totally subjective in order to further explore Nietzsche's claims and their implications.

Similarly, Tolstoy disagrees with Nietzsche's idea of the Will to Power. Nietzsche believed that the driving force behind all human beings is the Will to Power, the desire to assert oneself over one's surroundings and to not be subservient to anything.⁸⁹ Though it is often misinterpreted as one's desire to have power over others, in reality, having power over others is a byproduct, as the best way to not be subservient is to be in control. That ethos also becomes a problem for Tolstoy, because by asserting oneself, one is by definition also asserting their individual values. Tolstoy advocates for the universality of values, so love, compassion, and nonviolence are what should control human action.

The debate is fundamentally about whether there are universal morals. Even though most people would agree that love and compassion are important moral values, the way each person

⁸⁷ Leo Tolstoy and Leo Wiener, *The Kingdom of God Is within You* (Place of publication not identified: publisher not identified, 2017). Chapter XI

⁸⁸ Tolstoy and Wiener. Chapter V

⁸⁹ Nietzsche, Del Caro, and Pippin, *Thus Spoke Zarathustra*. 7-8

interprets those values as more important than the same terminology used to describe them. For example, some people would argue that love should be for all humans, and others would argue that it is only for those close to you. Effectively those values are different which can be seen through the actions they inspire. There is little evidence for truly universal values.

Hedonism and the Value of Pleasure

Another challenger to Nietzsche would be the philosophy of Hedonism, which attacked the idea that one should strive to become the *Übermensch*. To understand the concept of Hedonism, another philosopher to look at is Robert Nozick. Nozick disagreed with Hedonism, and proposed the idea of an Experience Machine to illustrate its flaws. Nozick's machine offers any experience with no cost, calling to question the value of pleasure as a life pursuit. While in the machine, we would forget that what we are experiencing is not real, making it feel like reality from our perspective. Nozick argues that most people would choose not to plug in, as they would see some inherent value in either reality or in the satisfaction of overcoming challenges.⁹⁰

However, upon further examination, people in our modern world would definitely opt to spend time in the experience machine, which Nietzsche would argue is a reflection of the meaninglessness of our modern lives. Millions of people spend endless time on social media, sacrificing real experiences for momentary dopamine hits. This addiction to pleasure is frighteningly similar to the concept of Nozick's machine. That Hedonistic reality in our world is a direct counter to Nietzsche's philosophy. Nietzsche claims that life is primarily about the suffering and struggle on the path to becoming the *Übermensch*, and the pleasure one experiences is merely a byproduct of achieving one's goals.⁹¹ According to this argument, although we might enjoy the experience machine and become trapped by it, we would not be living a fulfilling or meaningful life, as pure pleasure is not the goal.

Asceticism vs Nietzsche

Another philosophy that potentially counters Nietzsche is Asceticism, which argues that desires lead to unhappiness and suffering, and promotes lives of extreme modesty.⁹² On a fundamental level, this contradicts Nietzsche's philosophy. The *Übermensch* and the Will to Power are built upon the fulfillment of one's authentic desires, which Asceticism claims should be suppressed.

Nietzsche's philosophy is somewhere between Asceticism and Hedonism. While Nietzsche doesn't call for pleasure, he also doesn't call for the elimination of it. To him, whatever a person authentically believes to be the right action is the action that should be taken. Nietzsche was once again a moral relativist, meaning he believed that all persons have their own individual belief on what morality is. Ascetic religions like Buddhism assert that one should try to eliminate all pain in their lives, an absolute standard that Nietzsche would reject, and also a

⁹⁰ Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974. 42

⁹¹ Nietzsche, Del Caro, and Pippin. 16-17

⁹² Jonathan Haidt, *The Happiness Hypothesis: Finding Modern Truth in Ancient Wisdom* (New York: Basic Books, 2006). Chapter 5

rejection of self-fulfillment that he fundamentally argues for. Buddhists believe that desires are the ultimate source of pain and through asceticism one can eliminate all their desires. Nietzsche does not support the suppression of one's desires, but promotes the pursuit of them. Ultimately Nietzsche's philosophy lands between Asceticism and Hedonism, in that he claims that one shouldn't suppress their desires but also shouldn't be guided solely by them.

Although the *Übermensch* takes a different form for each person, a critical commonality is that it always takes the shape of each person's subjective ideals. What is a more important achievement than becoming the best person you can be? With the recent push towards prioritizing mental health in the modern world, psychology focused on finding meaning has been at the forefront. While modern society is incorrectly prioritizing happiness, psychology says we should be prioritizing self actualization.

Nietzsche vs Modern Psychology

Nietzsche's theories align surprisingly well when compared to modern psychological theories like Maslow's Hierarchy of needs. Essentially, this theory organizes a person's needs based on levels of importance. At the base are the essential needs like food and shelter. Following those needs are safety needs and then social needs like love and belonging. Next is esteem needs, and finally self actualization. Maslow defined self-actualization as those who "live creatively and fully using their talents, capacities, and potentialities."⁹³ While Nietzsche might not agree with the terminology, his concept of the *Übermensch* shares key similarities with Maslow's idea of self-actualization, both involving freedom from basic, constricting needs and emphasizing individual higher-level desires. Nietzsche's *Übermensch* is an individual who transcends conventional morality to create their own values and purpose. The core of both ideas is very similar, representing the pinnacle of human development where one is authentic and lives up to their full potential.

Although this idea of self actualization is exactly what Nietzsche advocates for, Nietzsche would likely disagree with Maslowian therapy, which contradicts his beliefs about how excellence and growth occur. To start, Nietzsche heavily criticized an overly safe mentality, believing that it led to conformity and mediocrity. While he was most likely referring to societal and social safety, the same still holds as a philosophical maxim; meaning that those who take risks will also reap the rewards. Even though Maslow believes we will be comfortable and happy with a low-risk life, according to Nietzsche, we will never be able to achieve greatness or become the *Übermensch*. Self-esteem is particularly important to Nietzsche's philosophy. He believes that individuals should derive their esteem from personal accomplishments, rather than from external validation. He criticized the idea of seeking approval from others, as he once again believed that it led to dependency, conformity, and mediocrity. Instead, he advised that one develops a strong sense of pride and self-respect in their own unique talents and achievements. For him, true esteem comes from within one's ability to live authentically and with courage.

⁹³ Pierre Pichère, *Maslow's Hierarchy of Needs*, trans. Carly Probert (Namur: Lemaitre Publishing, 2015). 32, 45, 58, 70, 82

How does Nietzsche align with modern scientific data? Modern research suggests that mindless pleasure is not a fundamental desire of humanity. In a 2010 study, researchers reviewed the Rat Dopamine Experiment, to study the ways in which the brain feels pleasure. Ultimately, Kringelbach and Berridge found that there are several different parts of the brain that regulate behavior and feelings of pleasure.⁹⁴ During the Rat Dopamine Experiment, researchers found that after neurologically linking the press of a button to the release of dopamine in a rat's brain the rat would obsessively press the button until death. This behavior led the researchers to believe that the rat was gaining immense pleasure out of pressing the button. Notably, it was found later in human patients that although dopamine motivated people to repeatedly do the inciting action, it caused no increase in happiness.⁹⁵ This can also be seen through drug addicts. Many of them report becoming adapted to the drugs and even wanting to stop, but being unable to do so.⁹⁶

As technology progresses, a future where humans are all hooked up to machines that can feed us infinite pleasure is not far away. We already have social media like TikTok which is designed to repeatedly give us dopamine hits in order to increase watchtime.⁹⁷ What this experiment shows is that there is actually no true desire for effortless pleasure beyond their exploitation of a chemical process in our brains. Nietzsche says that we should only do what we want to do. Addictive and conformative behavior gets in the way of what Nietzsche believes is our true potential: becoming the *Übermensch*. Ultimately, he is solely focused on what we truly want, and while there might have been a societal shift towards individualism, at least compared to the past, there are other needs that have been manufactured which we are subservient to.

Herd Mentality in Our Own Lives:

For Nietzsche, herd mentality is a problem ultimately because one is being controlled by something other than themselves. We are fundamentally controlled by the things we need, and although some are unavoidable like food and water, Nietzsche would argue that it is imperative that we minimize these needs in our own lives in order to pursue our wants, and become fully realized. Our own capitalist society has motivated us to create more needs for ourselves such as addictions like drugs or alcohol, but also for products like video games and social media. After all, why bother making a product that consumers want, when you can manipulate them into needing it? According to economist Adam Alter, products like social media and mobile games are specifically designed to be as addicting as possible.⁹⁸ Almost 97 percent of teens use the

⁹⁴ Morten L. Kringelbach and Kent C. Berridge, "The Functional Neuroanatomy of Pleasure and Happiness," *Discovery Medicine* 9, no. 49 (June 2010): 579–87.

⁹⁵ Kringelbach and Berridge.

⁹⁶ Nora D. Volkow, George F. Koob, and A. Thomas McLellan, "Neurobiologic Advances from the Brain Disease Model of Addiction," *New England Journal of Medicine* 374, no. 4 (2016): 363–71, <https://doi.org/10.1056/NEJMra1511480>.

⁹⁷ April C. S. Zheng, "The Neuroscience of TikTok: Does Short-Form Video Content Harm Our Attention Span?," *Brown Public Health Review*, December 13, 2021, <https://sites.brown.edu/publichealthjournal/2021/12/13/tiktok/>.

⁹⁸ Adam L. Alter, *Irresistible: The Rise of Addictive Technology and the Business of Keeping Us Hooked* (New York: Penguin Press, 2017). Chapter 1

internet daily.⁹⁹ Based on this description, it seems clear that for many people a want has transformed into a need. Nietzsche would not be surprised that young people have the highest rates of depression and suicide. We are being oppressed by our needs.

Are All Our Gods Dead?

Similarly to herd mentality, Nietzsche's fundamental issue with God was that he created universal ideals and morals for us to follow, preventing us from following our own desires. While God is no longer a major force in American society, there may be other forces that have a similar influence.

Modern political parties function in many comparable ways to religions. What specifically promotes the herd mentality is the promotion of specific values or morals. For example, political parties nowadays are very focused on promoting what they see as correct opinions and virtues. In the past, some religious people might have seen dissenters as evil or hell-bound, and politicians and people part of the political groups we have today often assert similar sentiments about the opposition. One common practice is the labeling of people as bigots, which can be likened to the proclamation of someone as a heretic. People use these political parties and the strong feelings and convictions surrounding them to suppress diversity of thought. Nietzsche would hate this emphasis on prescribed unity, as he constantly advocated for creativity and individualism, yet even with religion on the decline, collectivism has lived on in other forms. Nietzsche would say that we need to kill these new gods, or at least undermine their power over us.

Conclusion

The goal of this essay was to evaluate Nietzsche's claims about finding purpose in life, and to explore their relevance in modern society. His examination of the human psyche and the necessity of the freedom to individually express oneself seems to align with other philosophers and modern research. Additionally, his ideas of the importance of moral self guidance and following one's own values would directly address critical problems facing society, especially young people. Nietzsche would argue that in order to find happiness in modern society, we must reject political parties, collectivist morals, and break free from addiction to enable ourselves to pursue our own purpose. He would tell modern young people not to use social media because of its addictive design. He would encourage them to avoid chemical dependency by not using drugs or alcohol. He would argue that young people should actively develop their own opinions outside of rigid groups or organizations. He would urge people to take these first steps to become the Übermensch and live fully realized lives.

⁹⁹ Emily A. Vogels, Risa Gelles-Watnick, and Navid Massarat, "Teens, Social Media and Technology 2022," *Pew Research Center*, August 10, 2022, <https://www.pewresearch.org/internet/2022/08/10/teens-social-media-and-technology-2022/>.

Works Cited

- Alter, Adam L. 2017. *Irresistible: The Rise of Addictive Technology and the Business of Keeping Us Hooked*. New York: Penguin Press.
- Cottone, John. "The Profound Challenge of Existential Anxiety." *Psychology Today*, 2 Dec. 2022, <https://www.psychologytoday.com/us/blog/the-cube/202212/the-profound-challenge-of-existential-anxiety>.
- Csikszentmihalyi, Mihaly. 2009. *Flow: The Psychology of Optimal Experience*. Nachdr. Harper Perennial Modern Classics. New York: Harper [and] Row.
- Frey, Bruno S., and Alois Stutzer. 2002. *Happiness and Economics: How the Economy and Institutions Affect Well-Being*. Princeton Paperbacks. Princeton, NJ: Princeton Univ. Press.
- Haidt, Jonathan. 2006. *The Happiness Hypothesis: Finding Modern Truth in Ancient Wisdom*. New York: Basic Books.
- Jones, Jeffrey M. "U.S. Church Membership Falls Below Majority for First Time." Gallup. March 29, 2021. <https://news.gallup.com/poll/341963/church-membership-falls-below-majority-first-time.aspx>.
- Kelly, Jack. "Why Is It So Hard for Recent College Graduates to Find a Decent Job?" *Forbes*. February 15, 2024. <https://www.forbes.com/sites/jackkelly/2024/02/15/why-is-it-so-hard-for-recent-college-graduates-to-find-a-decent-job/>.
- Kringelbach, Morten L., and Kent C. Berridge. 2010. "The Functional Neuroanatomy of Pleasure and Happiness." *Discovery Medicine* 9 (49): 579–87.
- Nietzsche, Friedrich Wilhelm, Adrian Del Caro, and Robert B. Pippin. 2006. *Thus Spoke Zarathustra: A Book for All and None*. Cambridge Texts in the History of Philosophy. Cambridge ; New York: Cambridge University Press.
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.
- Pichère, Pierre. 2015. *Maslow's Hierarchy of Needs*. Translated by Carly Probert. Namur: Lemaitre Publishing.
- Russell, Bertrand, and A. C. Grayling. 2006. *The Conquest of Happiness*. Routledge Classics. London: Routledge.
- Ryan, Richard M., Veronika Huta, and Edward L. Deci. 2008. "Living Well: A Self-Determination Theory Perspective on Eudaimonia." *Journal of Happiness Studies* 9 (1): 139–70. <https://doi.org/10.1007/s10902-006-9023-4>.
- Sun, Wujun, Lei Liu, Yuan Jiang, Ping Fang, Xiaosheng Ding, and Guangjun Wang. 2023. "Why Are Hedonists Less Happy Than Eudaimonists? The Chain Mediating Role of Goal Conflict and Mixed Emotions." *Frontiers in Psychology* 14:1074026. <https://doi.org/10.3389/fpsyg.2023.1074026>.
- The Holy Bible: New International Version. 1980. London: Hodder and Stoughton.

Tolstoy, Leo, and Leo Wiener. 2017. *The Kingdom of God Is Within You*. Place of publication not identified: publisher not identified.

Vogels, Emily A., Risa Gelles-Watnick, and Navid Massarat. "Teens, Social Media and Technology 2022." Pew Research Center. August 10, 2022.

<https://www.pewresearch.org/internet/2022/08/10/teens-social-media-and-technology-2022/>.

Volkow, Nora D., George F. Koob, and A. Thomas McLellan. "Neurobiologic Advances from the Brain Disease Model of Addiction." *New England Journal of Medicine* 374, no. 4 (2016): 363-371. <https://doi.org/10.1056/NEJMra1511480>.

Zheng, April C. S. "The Neuroscience of TikTok: Does Short-Form Video Content Harm Our Attention Span?" *Brown Public Health Review*. December 13, 2021.

<https://sites.brown.edu/publichealthjournal/2021/12/13/tiktok/>.

Coagulation Clarification of Microplastics: Affordable, Natural, and Available Way to Treat Water for Rural Communities By Thiyagesh Venkatesan

Abstract

Water availability and cleanliness is a pressing issue—rural communities often lack clean water. One method used to clarify water is coagulation and flocculation, whereby a material called a coagulant is added to a colloid, adhering particles together into flocs. Later in flocculation, these flocs are pulled down. However, the industrial coagulation processes are expensive and unavailable for rural communities. Alum and other chemical coagulants used in this process can cause significant health and environmental hazards. Another problem is microplastics, small, tiny plastics, that may lead to anywhere from interference with reproduction to neurotoxicity. This paper explores use of natural coagulants in removing microplastics from drinking water. Natural coagulants are easily available for rural communities and are safe to use without complex industrial coagulation processes. Using household microplastic sources, a microplastic contaminated mixture is created with turbidity of 20 NTU (Nephelometric Turbidity Units). Natural coagulants — *Strychnos Potatorum* (Nirmali), chitosan, *Moringa Oleifera*, starch, and okra are prepared by mixing measured quantities of each on a 1 Molar NaCl solution. This solution is then used at different concentrations to study their effectiveness in removing microplastics from the water. This study found that *M. Oleifera* and Okra are the best biocoagulants at removing Microplastics. *Moringa* reduced turbidity to 5.02 NTU while Okra reduced turbidity to 4.77 NTU performing better than control which reduced turbidity to 5.85 NTU. Alum the industrial coagulant reduced turbidity to 0.67 NTU. This study proves the effectiveness of using natural readily available substances in removing microplastics from water.

Introduction

Microplastics have been a more concerning pollutant in Earth's environment, that may have profound impacts on human lifestyle and marine environments. Synthetic clothing with polyester and car tires have been major sources of microplastics [8][4]. Microplastics have been found widely, while humans ingest microplastics on a large scale. Health impacts of microplastics are generally unknown, but research has shown they may lead to respiratory disease, interference with reproduction, and neurotoxicity [8]. When marine organisms consume microplastics, they inflict upon themselves complications. Marine organisms have been harmed by the influx of microplastics, and transfer their complications across the food chain. From seafood to humans, microplastics have been found [8]. This experiment evaluates how a method called coagulation/flocculation, normally used in water treatment plants, can purify microplastics from water.

Industrial chemicals called coagulants like aluminum sulfate (commonly known as alum) have been used on a wide scale in coagulation [3][13]. Coagulation is the process used in industrial water treatment in which particles of a colloid are no longer suspended by the zeta potential due to the addition of a coagulant disrupting the system of electrostatic charges of these

particles called the Electron Double Layer (EDL) [12][6][13][1][2]. When coagulation happens, particles clump together into aggregates. This enables flocculation, where these aggregates settle. Later steps like filtration can remove the contaminants from water more easily. Coagulation is useful as methods like filtration can not always remove small particles such as microplastics [12].

The World Health Organization estimates that 2.2 billion people worldwide do not have access to clean drinking water [7]. However, the processes and materials used in industrial coagulation are neither available nor cheap for rural communities. Alzheimer's disease, among other illnesses, have been linked with aluminum based coagulants, rendering it harmful to use in rural water purification, where advanced methods to properly dispose of sludge and chemical removal of alum are unfeasible. Due to the adverse environmental and health impacts of chemical coagulants, natural coagulants are better suited for rural communities.

Sanskrit texts from India, dating back over 4,000 years, mention that the seeds of *Strychnos Potatorum* were used to clarify cloudy surface water. This practical use led to the species being aptly named the "clearing nut tree," or "Nirmali." Nirmali seeds contain anionic polyelectrolytes that aid in coagulation. Kalekar et al [15], show coagulation activity of *Potatorum* seed powder in purifying turbid well water. Alenazi et al [9], used *S. Potatorum* to study kaolinite turbidity reduction.

Moringa Oleifera is one of the most studied biocoagulant. *Moringa* is a tropical plant grown widely in impoverished parts of Africa and India. *Moringa* is known to contain a compound Moringin [5] which makes it an effective coagulant. Mohamed et al [18] studied both *Moringa Oleifera* and *S. Potatorum* seeds in purifying car wash wastewater. Asrafuzzaman et al [10] studied locally available natural coagulants - *Moringa Oleifera*, *Cicer arietinum* and *Dolichos lablab* on clay suspended water.

Starches from natural sources are known to contain polysaccharides. Starch is biodegradable and is easily available over the world. Ortiz et al [21] studied reduction of turbidity using cassava starch. Sibiya et al [20] studied magnetized rice starch as a coagulant. Chitosan is another natural biocoagulant that is shown to be effective. Rao [19] shows the effectiveness of chitosan in coagulation and flocculation of industrial wastewater using Chitosan. Macczak et al [17] studied the use of chitosan and Starch based flocculants. Okra was studied by Freitas et al [14] on the treatment of textile waste water.

While many studies have been published showing effectiveness of both chemical and biocoagulants on different types of waste water, very few have studied on microplastic contaminated water. Azizi et al [11] studied alum to remove different types of microplastics. Khan et al [16], studied removal of microplastics through coagulation/flocculation using alum, eggshells and ferromagnetic substances. Different microplastic sizes at varying solution pH levels were studied.

The aim of this study is to find if natural coagulants are suitable for removing microplastics from drinking water. The different coagulants studied in this paper are *M. Oleifera*, Okra, *S. Potatorum*, chitosan and starch. Microplastics from household items were used to create

microplastic contaminated water. Different biocoagulants were then applied at different concentrations to find the most suitable coagulant.

Methods

Materials:

Different coagulants used are: 1. *Moringa Oleifera* seed powder bought from an online vendor, sieved. 2. Okra seed powder (*Abelmoschus esculentus*). Seeds bought from the garden store are ground using a coffee grinder. 3. *Strychnos Potatorum* seed powder bought from an online vendor, sieved. 4. Chitosan - food grade chitosan bought from an online vendor. 5. Corn starch powder from the grocery store. 6. Alum (potassium aluminum sulfate) is bought from the grocery store. Microplastics were sourced from household dryer lint and synthetic textiles like polyester stuffings and acrylic strings. Turbidity is a measure of how clear the water is. This is measured in NTU - Nephelometric Turbidity Units. Turbidity is measured using a Turbidity meter. The NTU measurements after applying the different coagulants are compared to show the best coagulant.

Collecting Microplastics:

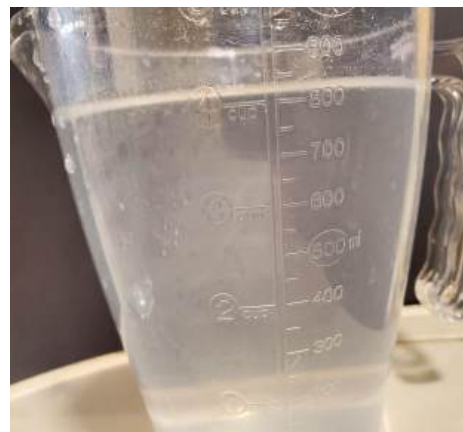
1. Collect dryer lint from a home clothes-dryer over time and store it in a dry place until at least 10.0 grams are collected. Ensure the dryer lint is collected only from cycles where polyester clothes and beddings are dried. This lint will be processed into microplastic for the experiment.
2. Separate large chunks from the collected dryer lint using a sieve. Grind the separated chunks with a coffee grinder until the lint breaks into smaller pieces.
3. Additional microplastics were derived from polyester stuffings and acrylic strings cut into small pieces.



(a) Different sources of microplastics



(b) Electric mixer



(c) Prepared turbid solution

Figure 1: Preparing microplastic contaminated turbid solution

Setup of Microplastic Contaminated Water:

1. Pour water and microplastic fibers into a blender Figure 1b. Run the blender for 2 minutes.
2. Place a filter in a funnel and into a beaker.
3. Pour the mixture into the 190 micron filter so that the filtrate goes into the beaker (Figure 1c. This filtrate contains very fine pieces of polyester lint that act as microplastic (all pieces larger than 190 micros are filtered out).
4. Measure the turbidity of the microplastic water. If the turbidity is close to 20 NTU, continue to the next step. Otherwise, repeat the blending process.
5. Label seven beakers with the coagulant name and the amount needed for the trial.
6. Add 100 ml of microplastic water to each of the seven beakers. For each set of trials, follow the instructions “Setup of microplastic contaminated water” for each set of trials(the experiment is run 7 trials at a time) to prepare the microplastic water to filter.

Coagulant Preparation:

1. For okra, use a coffee grinder to grind the seeds to fine powder. For *Moringa* and Nirmali, store bought powder is used.
2. For each type of coagulant, use a 190 micron filter to sieve out large chunks and fibers into a beaker. (skip this step for alum, chitosan and starch)
3. Prepare a 1 molar NaCl solution by adding 5.85g of salt to 100 ml of distilled water and stir with the magnetic mixer for 2 minutes. This is used as a solvent for *Moringa*, Nirmali, starch and okra.
4. For chitosan use 50ml vinegar and 50 ml distilled water mixed together as solvent. For alum just use 100 ml distilled water as solvent.
5. Pour 200 mg of each coagulant into the beaker containing 100 ml of solvent.
6. Mix each beaker individually for 2 minutes with a magnetic mixer, at its highest speed.



(a) Magnetic mixer

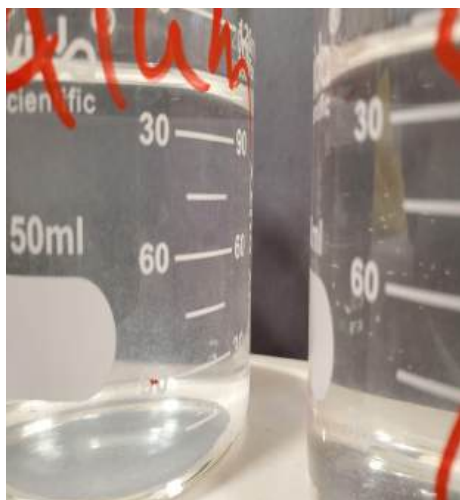


(b) Prepared coagulant solution for each coagulant

Figure 2: Coagulant preparation

Coagulant/Flocculation

1. For each set of trials, follow the instructions in the section “Setup of Microplastic Contaminated Water” to prepare the microplastic water. Note that the experiment is run 7 trials at a time
2. Follow the instructions in the section “Coagulant Preparation” to prepare the coagulant solution for each set of trials.
3. Measure and record the turbidity of each of the 7 beakers (labeled Control, Alum, Chitosan, Starch, *Moringa*, Okra and Nirmali).
4. Using a pipette, add a specific quantity (0.5ml, 1ml, 2.5 ml, 5 ml, or 10 ml) of the coagulant solution to each beaker. If a trial is a control, apply no coagulant solution (instead plain distilled water is added).
5. Stir each beaker individually for 2 minutes at the lowest speed (approximately 400 rpm) using the magnetic mixer to ensure better distribution of the coagulant and promote coagulation.
6. Allow 60 minutes for the suspension to settle.
7. Measure and record the turbidity of each beaker’s water with a turbidimeter.
8. Repeat steps 4-7 for 2 more hours.
9. Measure the turbidity of each beaker’s water after overnight settling.
10. Repeat steps 1-9 for each concentration for 3 trials each.



(a) Floccs forming in Alum treated microplastic water



(b) Experiment setup

Figure 3

Data

NTU is a measure of turbidity, or the cleanliness of water. The turbidity meter readings were taken every hour for 3 hours and overnight (12 hours). The NTU readings over time is shown in Figure 4. NTU for each coagulant at concentration is shown.

Table 1: NTU of coagulant treated microplastic water over time, organized by concentration

mL	Coagulant	Initial NTU	NTU 1hr	NTU 2hr	NTU 3hr	Overnight
0.5ml	alum	25.93	17.27	17.27	16.60	2.93
	chitosan	25.07	21.77	18.83	17.63	8.20
	control	24.30	20.87	21.17	20.30	5.17
	<i>Moringa</i>	23.83	25.43	24.27	23.27	5.03
	nirmali	23.40	21.03	20.60	20.57	6.37
	okra	24.87	22.03	21.67	21.53	5.13
	starch	25.00	20.60	21.23	19.97	5.50
	1ml	alum	23.27	13.93	12.37	9.93
chitosan		25.23	17.80	13.80	12.07	4.77
control		22.80	19.17	17.60	17.63	4.73
<i>Moringa</i>		23.60	16.43	22.77	21.97	3.73
nirmali		21.93	17.63	16.73	16.17	5.77
okra		22.90	21.20	21.77	21.77	3.67
starch		25.00	18.63	18.90	16.83	4.57
2.5ml		alum	22.07	11.60	10.83	9.83
	chitosan	23.53	17.17	14.67	13.23	6.20
	control	21.03	16.70	15.47	15.93	5.33
	<i>Moringa</i>	22.97	27.93	26.63	25.80	4.87
	nirmali	20.83	16.33	16.73	16.43	6.10
	okra	21.10	25.70	24.33	24.07	3.77
	starch	24.33	17.70	16.40	16.30	4.93
	5ml	alum	22.73	9.27	7.97	8.53
chitosan		23.43	17.27	16.17	15.53	8.53
control		22.37	17.03	17.40	16.43	7.07
<i>Moringa</i>		23.07	32.50	31.43	32.60	5.13
nirmali		21.30	17.70	16.27	16.03	7.83
okra		21.33	30.07	26.77	28.07	4.53
starch		24.13	18.33	15.47	15.57	6.33

10ml	alum	22.80	11.23	10.23	9.77	0.67
	chitosan	24.43	22.67	20.97	21.17	10.40
	control	22.90	17.30	16.63	16.20	6.97
	Moringa	21.90	32.07	31.03	30.20	5.07
	nirmali	23.27	23.70	23.10	22.10	9.03
	okra	23.53	30.80	26.10	28.07	5.37
	starch	23.83	13.20	12.57	13.87	5.97

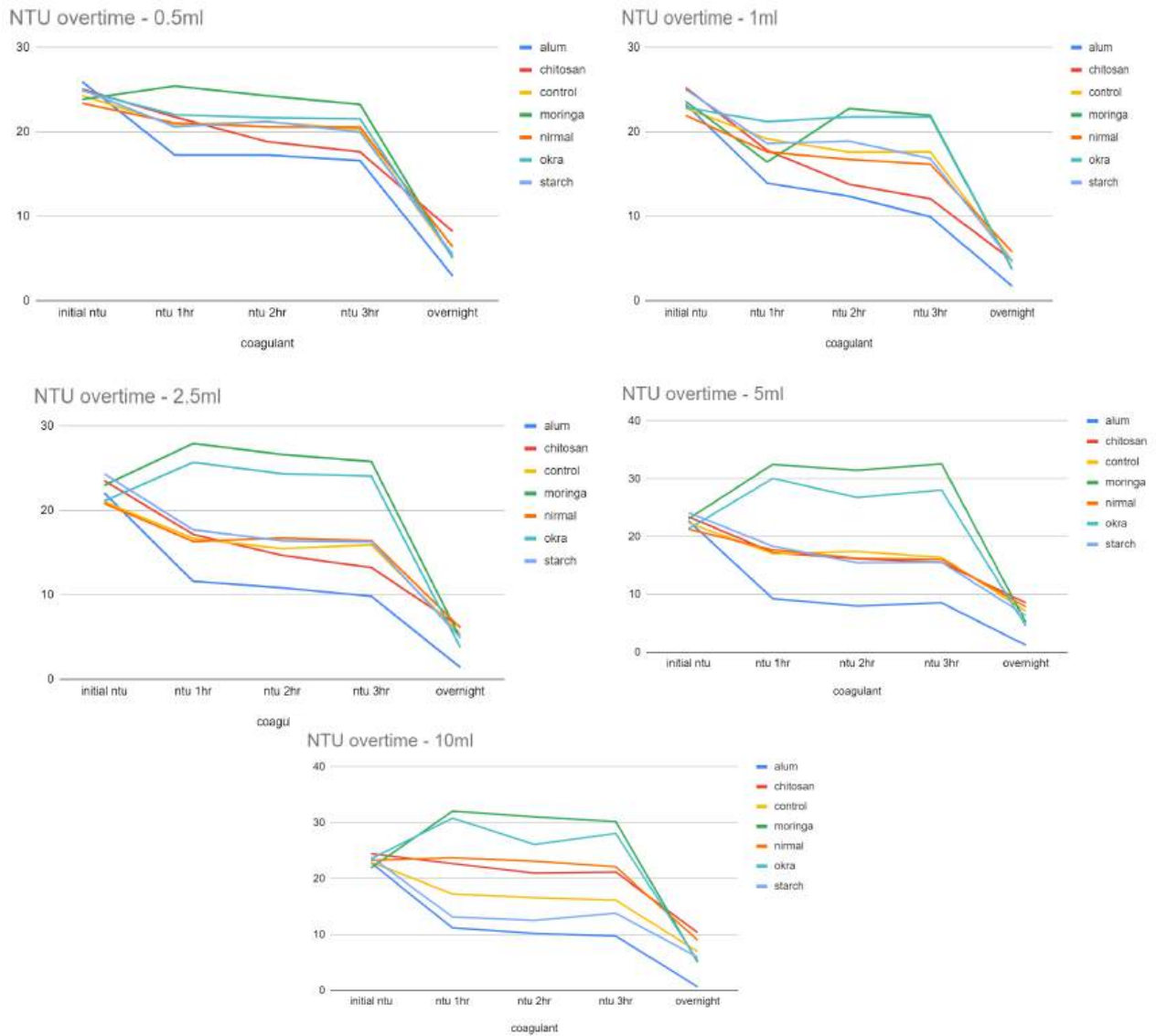


Figure 4: NTU reduction overtime

To show the overall effectiveness of the different coagulants, the NTU readings after letting coagulation continue for 12 hours are shown in Figure 5.

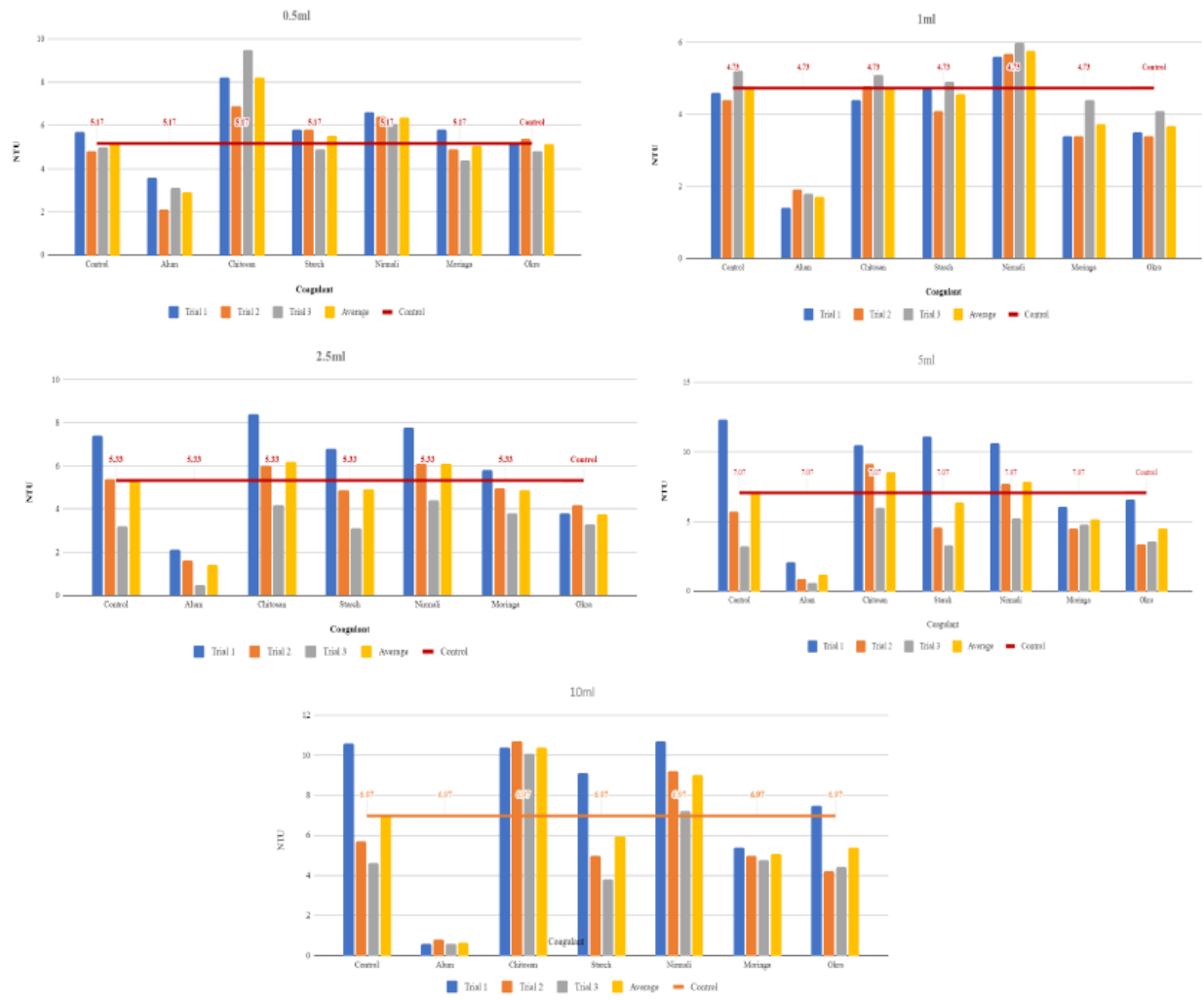


Figure 5: NTU reduction overnight

Results and Discussion

Coagulant type:

The chemical coagulant alum is the best coagulant by far at an average final (overnight) NTU of 0.67 NTU at 10 ml concentration. Alum is fast-acting and shows significant reduction in NTU even within one hour. Since it is a salt, it completely dissolves in water and does not add turbidity on its own.

The control, the trials where no coagulant was applied and reduction of turbidity was caused only by settling due to gravity, had an average final NTU of 5.85.

Among the natural coagulants, okra was the best, reducing turbidity to 4.77 NTU on average. *M. Oleifera* also reduced turbidity to 5.02 NTU on average. Both these are slow-acting, only reducing turbidity after coagulating overnight. One significant drawback is that these seed powders themselves add significant turbidity to the solution. For the first three hours, they actually increase the turbidity of the solution. Given more time, they eventually cause significant purification when compared to control. If the active ingredient was extracted or if the fiber impurities were removed before adding the coagulants, they will likely perform similarly as alum.

Starch, an unconventional coagulant, had an average final NTU of 5.46. Though performing better than control, Okra and *Moringa* were still significantly better than Starch.

However, *Strychnos Potatorum* and chitosan were both worse than the control, at 7.02 and 7.62 respectively. Chitosan's active ingredient was likely destroyed during preparation of food grade chitosan by the manufacturer. *S. Potatorum*'s seed powder was likely oxidized due to long term storage losing its coagulation abilities.

Coagulant concentration:

The best concentration of alum was at 10ml, resulting in a final NTU of 0.67. Coagulation strength increases as more alum is added. It also works faster at higher concentrations.

Okra and *Moringa* have significant impurities in them that it actually caused an increase in turbidity. This is especially true if added at higher concentration. If too little is added, coagulants don't have enough active compounds to cause coagulation. 0.5ml trial did not achieve good coagulation. If too much is added, there are too many impurities. Okra performed best at 1 ml concentration, achieving a final NTU of 3.67. The best concentration of *Moringa* was also at 1ml, achieving a final NTU of 3.73. Starch was best at 1 ml concentration achieving NTU of 4.57.

Most natural coagulants themselves were quite turbid. If the coagulants had not been as turbid, they would have been more effective. This highlights the importance of purifying the coagulant and extracting its active compounds that aid in coagulation. Also, okra seeds were grounded after being bought, and they performed better than the *Moringa* seeds which were already in powder form when bought. This additionally indicates the importance of the proper storage of biocoagulants, and more research is needed in that aspect. Despite the extra turbidity from okra and *Moringa*, the two coagulants were able to achieve significant purification. The poor performance of chitosan and *S.Potatorum* shows the importance of proper processing and storage of the biocoagulants, though even with proper storage, the two biocoagulants would likely still not outdo okra or *Moringa*.

Conclusions

In conclusion, natural coagulants are proven effective against removing microplastics from water. Alum, the chemical coagulant, removed microplastics effectively, but can cause adverse health and environmental effects. Okra and *Moringa Oleifera* were proven to be quite effective in removing microplastics from water, despite themselves being turbid. This was achieved by simple processing of organic materials that are available for rural communities. This research shows the promising potential of these natural materials in removing microplastics that are easily available and safe for the environment.

Works Cited

- "Everything You Want to Know..." Zeta Meter, Retrieved October 16, 2023, from <http://www.zeta-meter.com/coag.pdf>.
- "Flocculation." Oregon Health Authority, Retrieved October 16, 2023, from <https://www.oregon.gov/oha/PH/HEALTHYENVIRONMENTS/DRINKINGWATER/OPERAT>.
- "Comparative Study of the Use of Natural and Artificial Coagulants for the Treatment of Sullage (Domestic Wastewater)." August 2017.
- International Union for Conservation of Nature Annual Report 2017, June 2017, Retrieved October 16, 2023, from <https://portals.iucn.org/library/sites/library/files/documents/2018-007-En.pdf>.
- "Moringin, a Stable Isothiocyanate from Moringa Oleifera, Activates the Somatosensory and Pain Receptor TRPA1 Channel In Vitro." *Molecules*, February 22, 2020.
- "Development of Bio-Based Material from the Moringa Oleifera and Its Bio Coagulation Kinetic Modeling – A Sustainable Approach to Treat the Wastewater." *Heliyon*, September 1, 2022.
- "Drinking-Water." World Health Organization, September 13, 2023, Retrieved December 20, 2023, from <https://www.who.int/news-room/fact-sheets/detail/drinking-water>.
- "Health Effects of Microplastic Exposures: Current Issues and Perspectives in South Korea." April 20, 2023, Retrieved October 15, 2023, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10151227/>.
- Alenazi, M., et al. "Turbidity Removal Using Natural Coagulants Derived from the Seeds of Strychnos Potatorum: Statistical and Experimental Approach." *IOP Conference Series: Materials Science and Engineering*, vol. 888, IOP Publishing, 2020.
- Asrafuzzaman, M., A. N. M. Fakhruddin, and M. A. Hossain. "Reduction of Turbidity of Water Using Locally Available Natural Coagulants." *ISRN Microbiology*, vol. 2011, no. 632189, Dec. 2011.
- Azizi, N., M. Pirsahab, N. J. Haghighi, et al. "Removal of Most Frequent Microplastic Types and Sizes in Secondary Effluent Using Al₂(SO₄)₃: Choosing Variables by a Fuzzy Delphi Method." *Scientific Reports*, vol. 13, no. 20718, 2023.
- Bradley, E. "Wastewater Coagulation." November 29, 2022. Retrieved October 16, 2023, from <https://www.dober.com/water-treatment/resources/wastewater-coagulation>.
- Droste, R. L. *Theory and Practice of Water and Wastewater Treatment*. Wiley, 1997.
- Freitas, T., V. Oliveira, M. de Souza, H. Geraldino, V. Almeida, S. Fávoro, and J. Garcia. "Optimization of Coagulation-Flocculation Process for Treatment of Industrial Textile Wastewater Using Okra (A. Esculentus) Mucilage as Natural Coagulant." *Industrial Crops and Products*, vol. 76, 2015, pp. 538-544.
- Kalekar, S. V., and N. W. Shinde. "Coagulation Activity of Strychnos Potatorum L. Seed Powder." *International Journal of Sciences & Applied Research*, 2017.

- Khan, M., M. Ahmad, M. F. Hossain, A. Nawab, I. Ahmad, K. Ahmad, and S. Panyametheekul. "Microplastic Removal by Coagulation: A Review of Optimizing the Reaction Conditions and Mechanisms." *Water Emerging Contaminants & Nanoplastics*, vol. 2, no. 22, 2023.
- Maćczak, P., H. Kaczmarek, M. Ziegler-Borowska, K. Wegrzynowska-Drzymalska, and A. Burkowska-But. "The Use of Chitosan and Starch-Based Flocculants for Filter Backwash Water Treatment." *Materials*, vol. 15, no. 1056, 2022.
- Radin Mohamed, R. M. S., N. Rahman, and A. H. M. Kassim. "Moringa Oleifera and Strychnos Potatorum Seeds as Natural Coagulant Compared with Synthetic Common Coagulants in Treating Car Wash Wastewater: Case Study 1." *Asian Journal of Applied Sciences*, vol. 2, 2014.
- Rao, L. N. "Coagulation and Flocculation of Industrial Wastewater by Chitosan." *International Journal of Engineering and Applied Sciences*, vol. 2, no. 7, 2015.
- Sibiya, N., G. Amo-Duodu, E. Tetteh, and S. Rathilal. "Effect of Magnetized Coagulants on Wastewater Treatment: Rice Starch and Chitosan Ratios Evaluation." *Polymers (Basel)*, vol. 14, no. 20, 2022, p. 4342.
- Villabona-Ortíz, et al. "Reduction of Turbidity in Waters Using Cassava Starch as a Natural Coagulant." *Prospectiva*, 2021.

Stomach Cancer in South Korea: How Diet, H. pylori and Alcohol Impact Incidence Rates

By Lenna Yoon

Abstract

Stomach cancer has been a longstanding issue in South Korea, with the country currently having the third-highest incidence rate of gastric cancer in the world. Dietary factors may contribute to Korea's high stomach cancer incidence rate. Infection rates for H. pylori, an established bacterial carcinogen, are relatively high in Korea, which could be due to certain hygiene and communal eating practices. Additionally, different strains of H. pylori may have different strengths of virulence, possibly explaining discrepancies between Western and East-Asian infected populations. Korean cuisine is often rich in salt, which in high levels can enhance the carcinogenic effects of H. pylori and N-nitroso compounds found in processed meat, fermented foods, and commercially available seasonings. Furthermore, men are found to have twice the incidence and mortality rate of women in Korea, which may be partially due to higher rates of heavy drinking in men. While alcohol's effect on gastric cancer risk is debated, heavy drinking does appear to be a risk factor, particularly for individuals with variants of the ALDH2 gene that are uniquely common in East Asian populations.

Introduction

Cancer is a significant problem in our society today, where nearly one out of six people die from cancer worldwide, making it the second leading cause of death (WHO, "Cancer"). Stomach cancer is the fifth most common cancer worldwide and has the fourth-highest mortality rate (Sung et al.; WCRF). Across the globe, countries and populations are widely impacted by stomach cancer, one country in particular being South Korea. Cancer is South Korea's leading cause of death, with 27 percent of deaths being caused by cancer in 2020 (K.W. Jung et al.). While other cancers such as lung, colon, thyroid, and breast cancer are common in Korean populations (K.W. Jung et al.), South Korea has long had an issue with gastric cancer.

In 2002, the Korean Gastric Cancer Association (KGCA) reported gastric cancer as the most common cancer in the country (KGCA). South Korea currently has the third-highest incidence rate of gastric cancer in the world, only behind Mongolia and Japan (WCRF). Compared to the global incidence rate of 11.1 per 100,000, South Korea has a rate of 27.9 per 100,000 (WCRF). Additionally, men are much more likely to develop gastric cancer than women in Korea (K.W. Jung et al.). Although the crude stomach cancer incidence rate for women is 32.4 per 100,000, the incidence rate for men is almost double, 63.9 per 100,000 (Figure 1). This trend is seen in mortality rates as well, with a crude mortality rate of 6.7 per 100,000 for women and 12.7 per 100,000 for men (K.-W. Jung et al.). While stomach cancer is a prevalent problem worldwide, Korea has struggled with the disease in particular. Given stomach cancer's impact on Korean society, there is a need to investigate its causes and effects.

Gastric cancer incidence by age groups

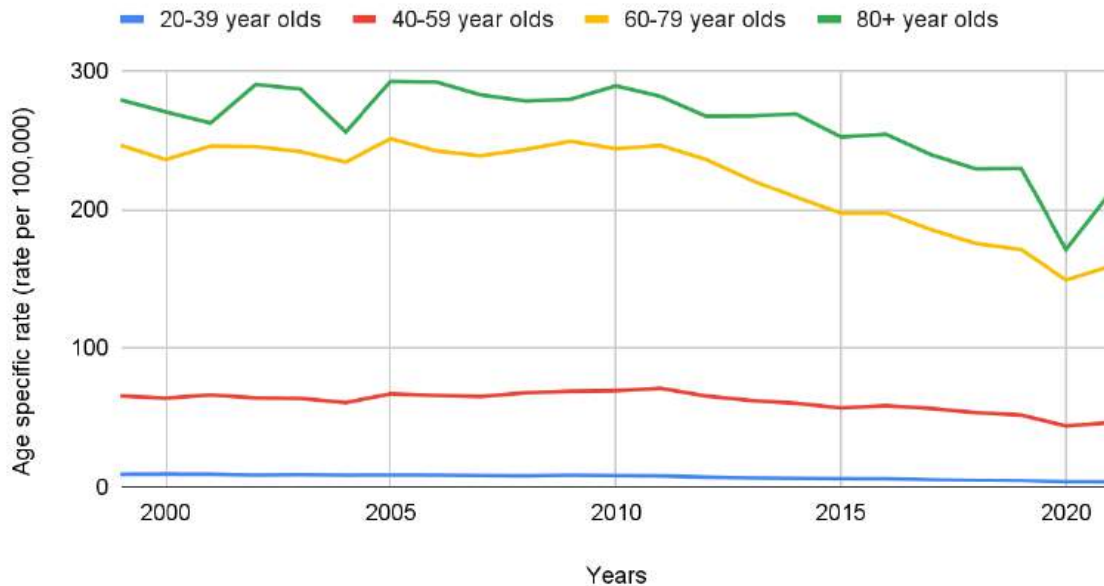


Fig 1: Graph shows gastric cancer incidence in the age groups of 20-39, 40-59, 60-79, 80+ in South Korea by year from 1999-2021. (reproduced from https://kosis.kr/statHtml/statHtml.do?orgId=117&tblId=DT_117N_A00023&conn_path=I2&language=en)

Numerous studies have investigated the risk factors and causes of stomach cancer. Research has shown direct links between higher stomach cancer risk and diet, pathogens, genetics, environment, and tobacco use (NCI, “Causes of Stomach Cancer - NCI”). Not only does South Korea have a high incidence rate of stomach cancer in the world, but in the USA, ethnic Koreans also disproportionately contribute to the number of stomach cancer cases. Out of all the ethnic groups in the United States, Korean Americans had the highest incidence rate in 2017, followed by Japanese Americans (E. Lee et al.). Diet is a risk factor shared by most South Koreans and Korean Americans with commonly consumed foods including kimchi, stew, bean paste, chili paste, sesame oil, garlic, and green onions (S. K. Lee et al.). To explore why the incidence rate is so high in ethnic Koreans, this review paper seeks to answer the following question: what is the role of diet in contributing to high incidence rates of stomach cancer in South Korea? Looking into the contributions that diet has on the development of gastric cancer is vital in preventing and lowering gastric cancer risk. Focusing on Korea can reveal more information helpful to combating stomach cancer in general.

Transmission of *H. pylori* and its role in gastric cancer

Helicobacter pylori (*H. pylori*) is a type of bacteria that grows in the mucus layer of the stomach. The bacteria is most likely spread through fecal and oral contact and is typically gained during childhood (NCI, *Helicobacter Pylori* (*H. Pylori*) and Cancer - NCI; Graham). Since its

discovery in 1982 by Australian scientists Barry Marshall and Robin Warren (Buckley and O'Morain), *H. pylori* was deemed a class I carcinogen in 1994 by the International Agency for Research on Cancer (IARC), a division of the World Health Organization (WHO). By causing gastritis and even gastric atrophy in the later stages of infection, the growth of *H. pylori* in the stomach produces and encourages inflammation, a well-known contributor to the formation of gastric cancer (Correa; Graham; Miftahussurur et al.). The stomach's acid barrier which protects the stomach from bacteria can also deteriorate due to an *H. pylori* infection (Miftahussurur et al.). *H. pylori* is known to have considerable genetic diversity, with varying effects being observed over different strains (Israel et al.).

One study conducted on gerbils by Israel et al. found that gastric ulcer strain B128 caused more severe symptoms and proliferation than duodenal ulcer strain G1.1 (Israel et al.). The *cagA* gene, which is acknowledged to be a major risk factor for gastric cancer, has distinctions between Western-type *cagA* and Eastern-type *cagA*. Western-type *cagA* contains motif EDIYA-C and East-Asian-type *cagA* motif EDIYA-D. East-Asian type *cagA* is considered to be more virulent, and is thought to have higher carcinogenic effects than Western-type *cagA*. However, Tserentogtokh et al. speculated that *cagA* types may not be a factor in geographic differences as in Mongolia, Western-type *cagA* was more common in gastric cancer patients instead of East-Asian type *cagA* (Tserentogtokh et al.). Additionally, several *H. pylori* strain types were found to affect Western and East Asian populations in different ways (Israel et al.; Ito et al.; Maeda et al.). For example, the *iceA1* gene, found in some strains of *H. pylori*, was identified as a contributor to the development of peptic ulcers in Western populations, while in Japanese subjects no increase in risk was observed (Ito et al.). Studies conducted in western populations concluded that *cagA*-positive increased risk for gastric cancer, however no increase in risk was found in studies conducted in East Asian populations (Maeda et al.).

Korea historically has had a problem with *H. pylori* as a majority of the population has been infected. The overall seropositivity in the country was 66.9% in 1998, but this number has since decreased to 51% in 2015 (J. H. Lee et al.). This decrease is most likely due to the 2013 revisitation of guidelines aiming for *H. pylori* eradication in South Korea (Lee). Furthermore, in 1998 the Korean *H. pylori* study group was founded and contributed greatly to research in the country (I. S. Park et al.) However, this level of seroprevalence is still elevated compared to rates in other developed countries in the west, with their average being 34.7 percent in 2018 (Zamani et al.).

There are several possible person-to-person methods for *H. pylori* infection, the most commonly agreed upon being gastro-oral, fecal-oral, and oral-oral (Stefano et al.). One possible transmission pathway for *H. pylori* in Korea is the spread and survival through contaminated food and water. In developing countries where *H. pylori* incidence rates are higher, contamination of water with fecal matter, saliva, vomitus, or dental plaque are thought to be an important pathway for infection (Brown). *H. pylori* has shown the ability to survive in water, foodstuffs like ready-to-eat foods, vegetables, some fruits, and ground beef and shown the ability to form biofilms (Quaglia and Dambrosio). Similarities between *H. pylori* and the foodborne

pathogen *C. jejuni* also lend to this theory (Wesley). Additionally, it was found that handwashing with soap in Korea after contact with excreta was only 23.5% in 2013 (J.-H. Kim et al.), so contact with food or water with unsanitary hands is a likely contributor to the spread of *H. pylori*. However, while a fecal-oral pathway is observed to be responsible for the majority of person to person infection, the evidence for an oral-oral pathway is varied (Duan et al.). In fact, various methods may alter the apparent appearance of *H. pylori* in the mouth as results from molecular and immunological methods may be impacted by cross-reactions with members of the campylobacter species (Mao et al.). However, some studies like Goud et al. found the presence of *H. pylori* in saliva, which suggests its ability to spread through the cross-contamination of saliva on foods (Sekhar Goud et al.).

The communal eating culture in Korea could be another contributor to high *H. pylori* prevalence. The nature of eating in Korea is centered around sharing, with side dish-sharing behaviors and many instances of saliva cross-contamination (Oh). It is common to see people use the same utensils for serving, grabbing, and eating food. These eating habits have started to be reconsidered with the recent COVID-19 pandemic by highlighting good hygiene practices and encouraging personal plates and individual servings (A.-R. Jung et al.). Epidemiological characteristics of the infection lend to this theory through the high incidence rates within families of children with *H. pylori* gastritis. A study by Roma-Giannikou et al. hypothesized that these close groups gained the infection through a common source, such as the consumption of the same meals during mealtimes (Roma-Giannikou et al.). Sanitary eating and general hygiene practices appear to have an impact on the transmission and infection of *H. pylori* which suggests risks associated with certain practices in communal eating.

High salt consumption and its impact on gastric cancer incidence

Sodium intake in South Korea has historically been high, with levels of consumption in Korean adults being an average of 260% of the daily sodium intake goal in 2011 (Kweon et al.). In 2024 Korea had the 15th highest salt consumption with 12.3 grams of daily salt consumption per capita (World Population Review). Koreans tend to consume foods with high sodium contents like seasoned stews, and pickled vegetables. Kimchi alone contributed to 23.5% of a Korean's average daily sodium intake in 2008. From 2010-2014, alterations of sodium content in foods contributed to more than 80% of a 19.5% decrease in sodium intake (H.-K. Park et al.). This trend persisted to 2019, but salt intake levels were still 1.5 times higher than the recommended level (Han et al.). Salted and fermented foods are prominent in Korean cuisine, especially in banchan, side dishes meant to be eaten with rice. Many fermented foods require salt in their preparation, such as kimchi, fermented soybean products, and fermented fish. In addition to salt's ability to improve taste, it also serves as a preserver and bacteria killer (S.-J. Jung et al.). In most Korean dishes, fermented and salted products such as soy sauce or chili bean paste are used as essential flavorers, often in addition to more salt (Kwon et al.).

High salt consumption is a well known risk factor for several major health problems like cardiovascular disease, chronic kidney disease, osteoporosis, and stomach cancer (Harvard)

Globally, around 1.89 million deaths a year are due to diseases associated with high sodium intake (WHO, “Sodium Reduction”). Excessive salt in the blood is often too much for most kidneys to handle, causing build up in the body. To dilute this sodium, the body produces fluid, which increases blood volume that can cause high blood pressure and harm to the heart and blood vessels (Michigan State University). Salt has also been found to increase incidence rates of gastric cancer when consumed at high levels (Guggenheim and Shah; Wu et al.; D’Elia et al.). Heavily salty foods like pickled foods, salted fish, dried fish, and processed meats have been correlated with high gastric cancer incidence rates (Guggenheim and Shah; D’Elia et al.), with risk increasing by 24%-27%. Salt may cause inflammation and damage to the stomach’s mucosa, which can lead to hyperplasia and raise the risk of cancerous mutations (Furihata et al.; Guggenheim and Shah; Wu et al.; D’Elia et al.). Guggenheim et al. also found a negative correlation between refrigerator use and gastric cancer incidence, which could relate to the use of salt as a preserver.

In preservation, salt provides an important role in preventing bacterial growth. Salt causes many bacteria to undergo osmotic shock, resulting in a loss of water in the cell causing death or limited growth, which can also be impacted by a decrease in oxygen solubility (Henney et al.). However, salt is shown to increase risk for carcinogenesis caused by *H. pylori* infection (Gaddy et al.). Due to the bacterium’s highly adaptable nature, it is likely that *H. pylori* bacteria can sense surrounding salt concentration levels and change gene expression in response (Gancz et al.). In fact, high sodium was found to influence various *H. pylori* strains by encouraging several gene expressions, most notably *cagA* (Gaddy et al.; Loh et al.). The resulting impact is an enhanced ability to affect and alter gastric epithelial cells, thus increasing risk for carcinogenesis (Loh et al.). Furthermore, a high salt diet changes the viscosity of the stomach’s protective mucous barrier, impacting immune homeostasis and increasing susceptibility to *H. pylori* infection (Guggenheim and Shah; Wu et al.; D’Elia et al.). Additionally, high sodium concentrations in the gut are associated with decreases in the amount of bacteria that produce immuno-competent short-chain fatty acids (Hamad et al.). Sodium butyrate, which is a key component of short-chain fatty acids, may decrease inflammation caused by *H. pylori* infection (Huang et al.), implying that low levels of sodium butyrate could increase risk for gastric malignancy.

Salty foods with large amounts of nitrate and nitrite may also contribute to the formation of carcinogenic N-nitroso compounds (NOCs) and nitrosamines (Wu et al.). According to the FDA, over 90% of known nitrosamines are carcinogenic, one of the most notable being NDMA (Atrakchi). NOCs and nitrosamines are often created as a byproduct of food production, when nitrates, nitrites, and various amines react. Areas of the body with an acidic pH like the mouth or stomach can also cause the formation of nitrosamines through the consumption of foods with nitrates and nitrites (Robles). Fruits and vegetables are a common source of nitrates and nitrites, along with processed meats which are common sources of nitrite due to nitrate/nitrite salt’s role in the curing process (Shakil et al.). A study done by Park et al. in 2015 found NDMA most frequently in agricultural food products and meat food products. The study also suggested that

fermentation in foods like kimchi could form NDMA. Additionally, seasonings were found to have the highest concentration of nitrosamines when tested on soybean paste, soy sauce, artificial sauces, and common seasonings (J. Park et al.). A study by Tatematsu et al. on rats observed increased carcinogenesis when salt was in combination with N-methyl-N-nitro-N-nitrosoguanidine (MNNG) and 4-nitroquinoline-1-oxide (NQO). However, groups that were fed salt alone displayed no increased carcinogenesis, highlighting salt's role in enhancing other possibly carcinogenic factors (Tatematsu et al.). As many Korean foods are processed products that use salt as a curer or preservative, it is very possible that NOCs and nitrosamines are being formed in these products, with their carcinogenic effects being enhanced due to a high salt diet. *H. pylori* has also been found to amplify carcinogens from nitrosamines by causing inflammation (Risch).

Despite this research, nitrate rich foods may have some benefits. Research has found that consumption of nitrate rich foods shows a decrease in gastric cancer risk and displays some protective benefits. This is most likely due to the fact that dietary sources of nitrate also contain high amounts of fiber and vitamin C, a proposed protective factor against gastric cancer (Ma et al.). Ma et al. also questioned negative findings about nitrite consumption, as many studies that found positive correlations with gastric cancer could have ignored effects from nitrosamines and simply attributed those impacts to nitrites (Ma et al.). Nevertheless, meta-analyses by Song et al. and Zhang et al. noted a negative correlation between nitrate intake and gastric cancer, and a positive correlation between nitrite and NDMA intake and gastric cancer (Song et al.; Zhang et al.). Thus, the impacts of nitrate and nitrite are somewhat unclear.

Effects of heavy alcohol consumption on gastric cancer

While gastric cancer is a prevalent issue across populations in South Korea, there is a clear disparity between male and female populations, where men experience higher incidence and mortality rates of stomach cancer (Figure 2).

Gastric cancer incidence by sex

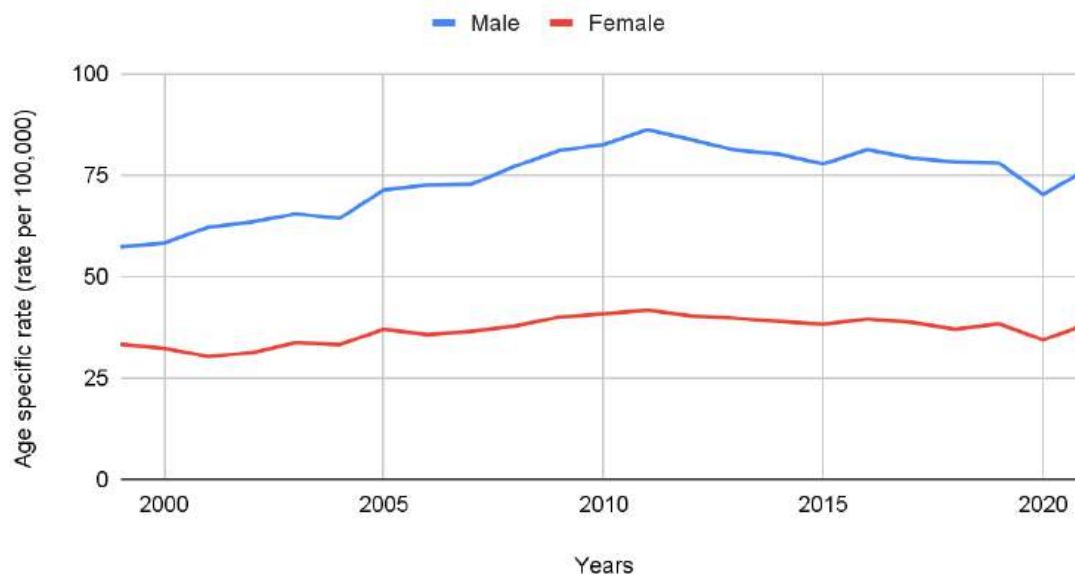


Fig 2: Graph shows gastric cancer incidence by sex in South Korea by year from 1999-2021 (reproduced from https://kosis.kr/statHtml/statHtml.do?orgId=117&tblId=DT_117N_A00023&conn_path=I2&language=en)

In 2023, the crude incidence rate for stomach cancer in South Korea was 63.9 per 100,000 in men and 32.4 per 100,000 in women (K.W. Jung et al.). Additionally, the crude mortality rate was 12.7 per 100,000 in men and 6.7 per 100,000 in women (K.W. Jung et al.). One factor that could contribute to this difference is the heavy alcohol consumption, which is prevalent in South Korean culture. Drinking is often done in social settings with colleagues or friends and acquaintances and by shots instead of sips. It isn't uncommon to drink over several rounds as well, with 55.5 percent of surveyed subjects reporting that they usually take two drinking rounds. Binge-drinking is also more common, where around 70 percent of subjects were found to binge-drink when drinking with friends or colleagues (KO and SOHN). In South Korea, men have historically consumed more alcohol than women. In 1998, the daily alcohol consumption for women was 2.09g and 14.78g for men. While women show a steeper trend in consumption with an increase to 5.79g for women and 23.94g for men in 2016-2018, there is still a large difference (Kim and Kim). The number of male drinkers is also higher, with 84 percent of adult men and 65.8 percent of adult women consuming alcohol in 2023. Furthermore, 28.3 percent of men and 14.5 percent of women 65 years old or younger and 30 percent of men and 4.5 percent of women older than 65 years old were classified as heavy drinkers (S. Lee et al.).

Whether or not alcohol affects gastric cancer risk is still debated, but there is some evidence that suggests that heavy alcohol consumption has some impact. Studies by He et al. and Tramacere et al. found that while light to moderate drinking had little to no effect, heavy drinkers had a significantly higher risk of gastric cancer (He et al.; Tramacere et al.). Heavy drinking is associated with poor nutrition and decrease in food intake, which limits important nutrients that

help prevent cancer (M. H. Kim et al.; Tramacere et al.). Excessive consumption of alcohol can also cause less folic acid to be produced, which has been tied to higher cancer risk (M. H. Kim et al.). Flavorings and additives may also be a contributor (M. H. Kim et al.). Additionally, a Korean study was done which found that former and current drinkers with variations of the ALDH2 genotype had a significantly increased risk for gastric cancer than those who drank little to none (Shin et al.). ALDH2 is an enzyme involved in the metabolism of alcohol by breaking down acetaldehyde in the body. Variants of ALDH2, which are more common in East Asian populations, result in increased acetaldehyde levels in the blood and saliva (Moy et al.). As a toxin, acetaldehyde could increase risk of cancer by interfering and damaging DNA (Salaspuuro; Seitz and Becker; He et al.).

Conclusion

In summary, factors related to diet such as personal and cultural habits that encourage the transmission of *H. pylori*, high salt levels in food, and high rates of alcohol consumption and heavy drinking have some impact on increasing stomach cancer risk in South Korea. By acknowledging these risk factors, efforts can be made to decrease high salt and alcohol consumption, and continue to eradicate *H. pylori* infection, helping to improve lives by decreasing risk for gastric cancer. Infection of *H. pylori* through cross-contamination of saliva in foods requires more research to prove, as various methods for measuring the presence of *H. pylori* yield different results and could be inaccurate. The role of alcohol in increasing gastric cancer risk also has inconsistent evidence and is still contested. Additionally, it must be acknowledged that many factors external to the review such as genetics, heavily processed foods, advances in technology and preventative care, environment, and others have significant contributions to the development of gastric cancer as well and have impacts on conclusions and statistics. In the future, looking into the methods that reduce gastric cancer risk such as supplements like vitamin C and *H. pylori* eradication and screening would continue expanding understanding about preventative measures, and investigating genetic factors like variants of the ALDH2 gene that could increase gastric cancer risk would help inform which individuals are inherently at higher risk.

Works Cited

- Atrakchi, Aisar. Nitrosamines as Impurities in Drugs - Health Risk Assessment and Mitigation Public Workshop.
- Brown, L. M. "Helicobacter Pylori: Epidemiology and Routes of Transmission." *Epidemiologic Reviews*, vol. 22, no. 2, 2000, pp. 283–97. PubMed, <https://doi.org/10.1093/oxfordjournals.epirev.a018040>.
- Buckley, Martin J. M., and Colm A. O'Morain. "Helicobacter Biology - Discovery." *British Medical Bulletin*, vol. 54, no. 1, Jan. 1998, pp. 7–16. Silverchair, <https://doi.org/10.1093/oxfordjournals.bmb.a011681>.
- Correa, P. "Chronic Gastritis as a Cancer Precursor." *Scandinavian Journal of Gastroenterology. Supplement*, vol. 104, 1984, pp. 131–36.
- D'Elia, Lanfranco, et al. "Habitual Salt Intake and Risk of Gastric Cancer: A Meta-Analysis of Prospective Studies." *Clinical Nutrition (Edinburgh, Scotland)*, vol. 31, no. 4, Aug. 2012, pp. 489–98. PubMed, <https://doi.org/10.1016/j.clnu.2012.01.003>.
- Duan, Miao, et al. "Transmission Routes and Patterns of Helicobacter Pylori." *Helicobacter*, vol. 28, no. 1, 2023, p. e12945. Wiley Online Library, <https://doi.org/10.1111/hel.12945>.
- Furihata, C., et al. "Cause and Effect between Concentration-Dependent Tissue Damage and Temporary Cell Proliferation in Rat Stomach Mucosa by NaCl, a Stomach Tumor Promoter." *Carcinogenesis*, vol. 17, no. 3, Mar. 1996, pp. 401–06. PubMed, <https://doi.org/10.1093/carcin/17.3.401>.
- Gaddy, Jennifer A., et al. "High Dietary Salt Intake Exacerbates Helicobacter Pylori-Induced Gastric Carcinogenesis." *Infection and Immunity*, vol. 81, no. 6, June 2013, p. 2258. www.ncbi.nlm.nih.gov, <https://doi.org/10.1128/IAI.01271-12>.
- Gancz, Hanan, et al. "Sodium Chloride Affects Helicobacter Pylori Growth and Gene Expression." *Journal of Bacteriology*, vol. 190, no. 11, June 2008, pp. 4100–05. PubMed Central, <https://doi.org/10.1128/JB.01728-07>.
- Graham, David Y. "Helicobacter Pylori Update: Gastric Cancer, Reliable Therapy, and Possible Benefits." *Gastroenterology*, vol. 148, no. 4, Apr. 2015, pp. 719-731.e3. PubMed, <https://doi.org/10.1053/j.gastro.2015.01.040>.
- Guggenheim, Douglas E., and Manish A. Shah. "Gastric Cancer Epidemiology and Risk Factors." *Journal of Surgical Oncology*, vol. 107, no. 3, Mar. 2013, pp. 230–36. PubMed, <https://doi.org/10.1002/jso.23262>.
- Hamad, Ibrahim, et al. "High-Salt Diet Induces Depletion of Lactic Acid-Producing Bacteria in Murine Gut." *Nutrients*, vol. 14, no. 6, Mar. 2022, p. 1171. PubMed Central, <https://doi.org/10.3390/nu14061171>.
- Han, Ye-Ji, et al. "Sodium Intake Trend and Current Intake Level by Meal Provision Place among the Citizens of Seoul." *Nutrition Research and Practice*, vol. 17, no. 3, June 2023, pp. 516–28. PubMed Central, <https://doi.org/10.4162/nrp.2023.17.3.516>.
- Harvard. "Salt and Sodium." *The Nutrition Source*, 18 July 2013, <https://nutritionsource.hsph.harvard.edu/salt-and-sodium/>.
- He, Zheng, et al. "Association between Alcohol Consumption and the Risk of Gastric Cancer: A

- Meta-Analysis of Prospective Cohort Studies.” *Oncotarget*, vol. 8, no. 48, Sept. 2017, pp. 84459–72. www.oncotarget.com, <https://doi.org/10.18632/oncotarget.20880>.
- Henney, Jane E., et al. “Preservation and Physical Property Roles of Sodium in Foods.” *Strategies to Reduce Sodium Intake in the United States*, National Academies Press (US), 2010. www.ncbi.nlm.nih.gov, <https://www.ncbi.nlm.nih.gov/books/NBK50952/>.
- Huang, Yumei, et al. “Effects of Sodium Butyrate Supplementation on Inflammation, Gut Microbiota, and Short-Chain Fatty Acids in Helicobacter Pylori-Infected Mice.” *Helicobacter*, vol. 26, no. 2, Apr. 2021, p. e12785. PubMed, <https://doi.org/10.1111/hel.12785>.
- Israel, Dawn A., et al. “Helicobacter Pylori Strain-Specific Differences in Genetic Content, Identified by Microarray, Influence Host Inflammatory Responses.” *Journal of Clinical Investigation*, vol. 107, no. 5, Mar. 2001, pp. 611–20.
- Ito, Yoshiyuki, et al. “Sequence Analysis and Clinical Significance of the iceA Gene from Helicobacter Pylori Strains in Japan.” *Journal of Clinical Microbiology*, vol. 38, no. 2, Feb. 2000, pp. 483–88.
- Jung, A.-Reum, et al. “COVID-19 Impact on Hygiene Practices for Food Safety in South Korea.” *Public Health in Practice*, vol. 3, Feb. 2022, p. 100241. PubMed Central, <https://doi.org/10.1016/j.puhip.2022.100241>.
- Jung, Kyu-Won, et al. “Prediction of Cancer Incidence and Mortality in Korea, 2023.” *Cancer Research and Treatment : Official Journal of Korean Cancer Association*, vol. 55, no. 2, pp. 400–07. PubMed Central, <https://doi.org/10.4143/crt.2023.448>.
- Jung, Su-Jin, et al. “Fermented Foods of Korea and Their Functionalities.” *Fermentation*, vol. 8, no. 11, 11, Nov. 2022, p. 645. www.mdpi.com, <https://doi.org/10.3390/fermentation8110645>.
- KGCA. “Nationwide Gastric Cancer Report in Korea.” *Journal of the Korean Gastric Cancer Association*, vol. 2, no. 2, June 2002, pp. 105–14. jgc-online.org, <https://doi.org/10.5230/jkgca.2002.2.2.105>.
- Kim, Jong-Hun, et al. “Burden of Disease Attributable to Inadequate Drinking Water, Sanitation, and Hygiene in Korea.” *Journal of Korean Medical Science*, vol. 33, no. 46, Oct. 2018, p. e288. PubMed Central, <https://doi.org/10.3346/jkms.2018.33.e288>.
- Kim, Mi Hui, et al. “Alcohol Consumption and Gastric Cancer Risk in Korea: A Case-Control Study.” *Nutrition Research and Practice*, vol. 13, no. 5, Oct. 2019, pp. 425–33. PubMed Central, <https://doi.org/10.4162/nrp.2019.13.5.425>.
- Kim, Sang Young, and Hyun Ja Kim. “Trends in Alcohol Consumption for Korean Adults from 1998 to 2018: Korea National Health and Nutritional Examination Survey.” *Nutrients*, vol. 13, no. 2, Feb. 2021, p. 609. PubMed Central, <https://doi.org/10.3390/nu13020609>.
- KO, Seungduk, and Aeree SOHN. “Behaviors and Culture of Drinking among Korean People.” *Iranian Journal of Public Health*, vol. 47, no. Suppl 1, July 2018, pp. 47–56.
- Kweon, Sanghui, et al. “Data Resource Profile: The Korea National Health and Nutrition Examination Survey (KNHANES).” *International Journal of Epidemiology*, vol. 43, no. 1, Feb. 2014, pp. 69–77. Silverchair, <https://doi.org/10.1093/ije/dyt228>.

- Kwon, Dae Young, et al. "Science and Philosophy of Korea Traditional Foods (K-Food)." *Journal of Ethnic Foods*, vol. 10, no. 1, July 2023, p. 26. BioMed Central, <https://doi.org/10.1186/s42779-023-00194-3>.
- Lee, Eunjung, et al. "Stomach Cancer Disparity among Korean Americans by Tumor Characteristics: Comparison with Non-Hispanic Whites, Japanese Americans, South Koreans, and Japanese." *Cancer Epidemiology, Biomarkers & Prevention*, vol. 26, no. 4, Apr. 2017, pp. 587–96. Silverchair, <https://doi.org/10.1158/1055-9965.EPI-16-0573>.
- Lee, Jeong Hoon, et al. "Seroprevalence of Helicobacter Pylori in Korea: A Multicenter, Nationwide Study Conducted in 2015 and 2016." *Helicobacter*, vol. 23, no. 2, Apr. 2018, p. e12463. PubMed Central, <https://doi.org/10.1111/hel.12463>.
- Lee, S. K., et al. "Acculturation and Dietary Practices among Korean Americans." *Journal of the American Dietetic Association*, vol. 99, no. 9, Sept. 1999, pp. 1084–89. PubMed, [https://doi.org/10.1016/S0002-8223\(99\)00258-8](https://doi.org/10.1016/S0002-8223(99)00258-8).
- Lee, Sami, et al. "Current Status of Korean Alcohol Drinking in Accordance with the Korean Alcohol Guidelines for Moderate Drinking Based on Facial Flushing." *Korean Journal of Family Medicine*, vol. 44, no. 3, May 2023, pp. 129–42. PubMed Central, <https://doi.org/10.4082/kjfm.23.0024>.
- Lee, Sun-Young. "Current Progress toward Eradicating Helicobacter Pylori in East Asian Countries: Differences in the 2013 Revised Guidelines between China, Japan, and South Korea." *World Journal of Gastroenterology : WJG*, vol. 20, no. 6, Feb. 2014, pp. 1493–502. PubMed Central, <https://doi.org/10.3748/wjg.v20.i6.1493>.
- Loh, John T., et al. "Regulation of Helicobacter Pylori cagA Expression in Response to Salt." *Cancer Research*, vol. 67, no. 10, May 2007, pp. 4709–15. PubMed, <https://doi.org/10.1158/0008-5472.CAN-06-4746>.
- Ma, Linsha, et al. "Nitrate and Nitrite in Health and Disease." *Aging and Disease*, vol. 9, no. 5, Oct. 2018, pp. 938–45. PubMed Central, <https://doi.org/10.14336/AD.2017.1207>.
- Maeda, S., et al. "High Seropositivity of Anti-CagA Antibody in Helicobacter Pylori-Infected Patients Irrelevant to Peptic Ulcers and Normal Mucosa in Japan." *Digestive Diseases and Sciences*, vol. 42, no. 9, Sept. 1997, pp. 1841–47. PubMed, <https://doi.org/10.1023/a:1018846723379>.
- Mao, Xiaojun, et al. "Colonization of Helicobacter Pylori in the Oral Cavity – an Endless Controversy?" *Critical Reviews in Microbiology*, vol. 47, no. 5, Sept. 2021, pp. 612–29. Taylor and Francis+NEJM, <https://doi.org/10.1080/1040841X.2021.1907740>.
- Michigan State University. "How Does Salt and Sodium Impact the Human Body?" Michigan State University, 19 Dec. 2014, https://www.canr.msu.edu/news/how_does_salt_and_sodium_impact_the_human_body.
- Miftahussurur, Muhammad, et al. "Helicobacter Pylori as an Oncogenic Pathogen, Revisited." *Expert Reviews in Molecular Medicine*, vol. 19, Jan. 2017, p. e4. Cambridge University Press, <https://doi.org/10.1017/erm.2017.4>.
- Moy, Kristin A., et al. "Alcohol and Tobacco Use in Relation to Gastric Cancer: A Prospective

- Study of Men in Shanghai, China.” *Cancer Epidemiology, Biomarkers & Prevention* : A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology, vol. 19, no. 9, Sept. 2010, pp. 2287–97. PubMed Central, <https://doi.org/10.1158/1055-9965.EPI-10-0362>.
- NCI. “Causes of Stomach Cancer - NCI.” National Cancer Institute, 7 Apr. 2023, <https://www.cancer.gov/types/stomach/causes-risk-factors>. nciglobal,ncienterprise.
- NCI. *Helicobacter Pylori (H. Pylori) and Cancer - NCI*. 25 Sept. 2013, <https://www.cancer.gov/about-cancer/causes-prevention/risk/infectious-agents/h-pylori-fact-sheet>. nciglobal,ncienterprise.
- Oh, Se-Young. “Food Sharing Characteristics in Modern Korean Society.” *Journal of the Korean Society of Food Culture*, vol. 20, no. 6, 2005, pp. 683–87.
- Park, Hye-Kyung, et al. “Progress on Sodium Reduction in South Korea.” *BMJ Global Health*, vol. 5, no. 5, May 2020, p. e002028. PubMed Central, <https://doi.org/10.1136/bmjgh-2019-002028>.
- Park, I. S., et al. “Helicobacter Pylori Infection in Korea.” *Yonsei Medical Journal*, vol. 42, no. 4, Aug. 2001, pp. 457–70. PubMed, <https://doi.org/10.3349/ymj.2001.42.4.457>.
- Park, Jong-eun, et al. “Distribution of Seven N-Nitrosamines in Food.” *Toxicological Research*, vol. 31, no. 3, Sept. 2015, pp. 279–88. PubMed Central, <https://doi.org/10.5487/TR.2015.31.3.279>.
- Quaglia, Nicoletta C., and Angela Dambrosio. “Helicobacter Pylori: A Foodborne Pathogen?” *World Journal of Gastroenterology*, vol. 24, no. 31, Aug. 2018, p. 3472. www.ncbi.nlm.nih.gov, <https://doi.org/10.3748/wjg.v24.i31.3472>.
- Risch, Harvey A. “Pancreatic Cancer: Helicobacter Pylori Colonization, N-Nitrosamine Exposures, and ABO Blood Group.” *Molecular Carcinogenesis*, vol. 51, no. 1, Jan. 2012, pp. 109–18. PubMed, <https://doi.org/10.1002/mc.20826>.
- Robles, H. “Nitrosamines.” *Encyclopedia of Toxicology (Third Edition)*, edited by Philip Wexler, Academic Press, 2014, pp. 584–85. ScienceDirect, <https://doi.org/10.1016/B978-0-12-386454-3.00523-6>.
- Roma-Giannikou, Eleftheria, et al. “Intrafamilial Spread of Helicobacter Pylori: A Genetic Analysis.” *Helicobacter*, vol. 8, no. 1, Feb. 2003, pp. 15–20. PubMed, <https://doi.org/10.1046/j.1523-5378.2003.00126.x>.
- Salaspuro, Mikko. “Acetaldehyde and Gastric Cancer.” *Journal of Digestive Diseases*, vol. 12, no. 2, Apr. 2011, pp. 51–59. PubMed, <https://doi.org/10.1111/j.1751-2980.2011.00480.x>.
- Seitz, Helmut K., and Peter Becker. “Alcohol Metabolism and Cancer Risk.” *Alcohol Research & Health*, vol. 30, no. 1, 2007, pp. 38–47.
- Sekhar Goud, E. V. Soma, et al. “Identification of Helicobacter Pylori in Saliva of Patients with and without Gastritis by Polymerase Chain Reaction.” *Journal of Pharmacy & Bioallied Sciences*, vol. 11, no. Suppl 3, Nov. 2019, pp. S523–29. PubMed Central, https://doi.org/10.4103/jpbs.JPBS_260_18.
- Shakil, Mynul Hasan, et al. “Nitrites in Cured Meats, Health Risk Issues, Alternatives to Nitrites:

- A Review.” *Foods*, vol. 11, no. 21, Oct. 2022, p. 3355. PubMed Central, <https://doi.org/10.3390/foods11213355>.
- Shin, Cheol Min, et al. “Association between Alcohol Intake and Risk for Gastric Cancer with Regard to ALDH2 Genotype in the Korean Population.” *International Journal of Epidemiology*, vol. 40, no. 4, Aug. 2011, pp. 1047–55. PubMed, <https://doi.org/10.1093/ije/dyr067>.
- Song, Peng, et al. “Dietary Nitrates, Nitrites, and Nitrosamines Intake and the Risk of Gastric Cancer: A Meta-Analysis.” *Nutrients*, vol. 7, no. 12, Dec. 2015, pp. 9872–95. PubMed Central, <https://doi.org/10.3390/nu7125505>.
- Stefano, Kayali, et al. “Helicobacter Pylori, Transmission Routes and Recurrence of Infection: State of the Art.” *Acta Bio Medica : Atenei Parmensis*, vol. 89, no. Suppl 8, 2018, pp. 72–76. PubMed Central, <https://doi.org/10.23750/abm.v89i8-S.7947>.
- Sung, Hyuna, et al. “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.” *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, May 2021, pp. 209–49. PubMed, <https://doi.org/10.3322/caac.21660>.
- Tatematsu, M., et al. “Effects in Rats of Sodium Chloride on Experimental Gastric Cancers Induced by N-Methyl-N-Nitro-N-Nitrosoguanidine or 4-Nitroquinoline-1-Oxide.” *Journal of the National Cancer Institute*, vol. 55, no. 1, July 1975, pp. 101–06. PubMed, <https://doi.org/10.1093/jnci/55.1.101>.
- Tramacere, I., et al. “A Meta-Analysis on Alcohol Drinking and Gastric Cancer Risk.” *Annals of Oncology*, vol. 23, no. 1, Jan. 2012, pp. 28–36. www.annalsofoncology.org, <https://doi.org/10.1093/annonc/mdr135>.
- Tserentogtokh, Tegshee, et al. “Western-Type Helicobacter Pylori CagA Are the Most Frequent Type in Mongolian Patients.” *Cancers*, vol. 11, no. 5, May 2019, p. 725. PubMed Central, <https://doi.org/10.3390/cancers11050725>.
- WCRF. “Stomach Cancer Statistics.” WCRF International, 2022, <https://www.wcrf.org/cancer-trends/stomach-cancer-statistics/>.
- Wesley, Irene V. “Helicobacter and Arcobacter: Potential Human Foodborne Pathogens?” *Trends in Food Science & Technology*, vol. 8, no. 9, Sept. 1997, pp. 293–99. ScienceDirect, [https://doi.org/10.1016/S0924-2244\(97\)01050-9](https://doi.org/10.1016/S0924-2244(97)01050-9).
- WHO. “Cancer.” World Health Organization, 2019, <https://www.who.int/health-topics/cancer>.
- WHO. “Sodium Reduction.” World Health Organization, 2023, <https://www.who.int/news-room/fact-sheets/detail/salt-reduction>.
- World Population Review. “Salt Consumption by Country 2024.” *World Population Review*, <https://worldpopulationreview.com/country-rankings/salt-consumption-by-country>. Accessed 5 Sept. 2024.
- Wu, Xiaomin, et al. “Effect of Dietary Salt Intake on Risk of Gastric Cancer: A Systematic Review and Meta-Analysis of Case-Control Studies.” *Nutrients*, vol. 14, no. 20, Oct. 2022, p. 4260. PubMed Central, <https://doi.org/10.3390/nu14204260>.
- Zamani, M., et al. “Systematic Review with Meta-Analysis: The Worldwide Prevalence of Helicobacter Pylori Infection.” *Alimentary Pharmacology & Therapeutics*, vol. 47, no. 7,

2018, pp. 868–76. Wiley Online Library, <https://doi.org/10.1111/apt.14561>.

Zhang, Fei-Xiong, et al. “Association Between Nitrite and Nitrate Intake and Risk of Gastric Cancer: A Systematic Review and Meta-Analysis.” *Medical Science Monitor : International Medical Journal of Experimental and Clinical Research*, vol. 25, Mar. 2019, pp. 1788–99. PubMed Central, <https://doi.org/10.12659/MSM.914621>.

Predicting TED Talk Views Using Topic Tags By Ernest Lai

Abstract

TED Talks are globally recognized as an influential way to spread important and diverse ideas. The organization TED has created a repository of all TED Talks on its website and annotated them with topic tags, the essential themes present in the Talk. In this project, I aim to develop an effective and accurate model to predict the number of views on a TED Talk using its associated topic tags. This is important because it can help TED to better select TED Talk topics and speakers to reach a larger audience and impact more people. I developed three models based on machine-learning algorithms to predict the number of views, the first based on Universal Sentence Encoder (USE) vectors, the second utilizing the topic tags associated with each TED Talk, and the third using computer-generated topic tags based on the USE vectors. I found a significant improvement in accuracy from Model 1 to Model 2, and, unexpectedly, another significant increase in accuracy from Model 2 to Model 3. This demonstrates that computer-generated topic tags are more effective than manually-assigned tags at predicting the number of views on a TED Talk.

Introduction

TED Talks are world-renowned presentations covering a dazzling array of subjects and given by diverse, expert speakers with the vision of sharing “ideas worth spreading”.¹⁰⁰ The TED website contains a repository of all TED Talks, dating back to 2006. Each of these TED Talks has been labeled with a group of tags, which are keywords, phrases, and themes that are relevant and important to the Talk. In this project, I seek to build a model that accurately predicts the number of views on a TED Talk using the tags associated with each TED Talk. One main reason that predicting views is important is that it can help optimize content and speakers. This model can identify which TED Talks and/or TED Talk topics generate the most viewership, and using this information TED can cater its Talks and speakers so that they are more likely to attract viewers. This project involves supervised learning and both regression and classification in its algorithms; it combines numerical and language data to output a mean absolute error of the number of views on each TED Talk.

Background

Much work has been done using the same dataset. Some have also looked to predict the number of views on a TED Talk, while others have instead pursued data analysis and TED Talk recommendation. One Kaggle user by the name Dochev created machine-learning models to predict the number of views (Dochev, 2019). They utilized data such as duration, number of languages available, date of filming and publication, related talks, and various other features to build their model. Compared to their work, I instead used the tags associated with each TED Talk to build my model.

¹⁰⁰ “TED Talks.” *Ted Talks*, www.ted.com/about/programs-initiatives/ted-talks. Accessed 8 Sept. 2024.

Other researchers have built similar algorithms to attach tags to documents. For example, Tuarob et al. (2013) created a model to automatically annotate metadata in ONEMercury, a harvester that collects ecological and environmental data from repositories and makes it searchable. They proposed two algorithms for tag recommendation, one based upon term frequency-inverse document frequency (TF-IDF), a measure of how important a word is to a document, and the other based upon Latent Dirichlet Allocation (LDA), a modeling technique that attaches “topics” or sets of words related to the document. While their approach relates to mine in its general goal, my approach is alternatively built using USE vectors and documents with pre-existing tags.

Dataset

I obtained the dataset for this project from the website Kaggle, a dataset repository; this dataset was uploaded to Kaggle by the user *Rounakbanik* (Banik, 2017) and contains data about all the TED Talks from the TED.com website until September 21, 2017. The first file in this dataset captured data on the TED Talk metadata, including the name of the speaker, the title of the talk, the number of views garnered, and the tags for the talk. The second file provided the transcript for each speech, along with the speech’s URL link. After importing these two files into Python, I merged them on the URL column as part of data preprocessing. The final data included 2467 TED talks. This project focuses on the tags and the views for each TED Talk, therefore combining both numerical and language data.

Methods

After completing the data preprocessing, a computational representation of the meaning of each TED talk was needed. I decided to use the Universal Sentence Encoder (USE). USE is a model that encodes any length text into a 512-dimensional vector. Developed by Google, the model is trained on a wide variety of data sources and texts greater than word length. USE can be used for a variety of NLP tasks, including semantic textual similarity (Cer et al., 2018).

I began by using the USE to encode each of the 2467 TED Talks. A baseline model was constructed to predict the number of views on each TED Talk purely based on the USE vectors for each Talk. I split the data into a training set, randomly assigning 20% of the data to the test set, then fit a Random Forest Regressor (RFR) to the training set using the sklearn library. A random forest is an ensemble machine-learning algorithm, meaning it combines multiple decision trees that “vote” to make a prediction. Next, I used the Regressor to predict the number of views for the TED talks in the test set. Finally, I calculated the mean absolute error (MAE) between the predicted and actual views.

Attempting to improve upon the first model, a second model was built but trained instead on the tags manually assigned to each Talk. First, the total number of unique tags across all the 2467 TED Talks was calculated, which was 416. I then created a binary matrix of dimensions 2467 x 416 to represent the tags associated with each Talk, with a 1 representing that the tag *is* associated with that Talk and a 0 if it is *not* associated. The second model was then trained with

this tag matrix instead of the USE vectors as was the case in the first model. I assigned 20% of the data to the test set, created and fit an RFR on the training set, used the RFR to predict the number of views for the test set, and calculated the MAE between the predicted and actual views.

The third and final model aimed to achieve the same results as the second model, while requiring less human effort since the topic tags were manually generated. As a first step, this final model used the USE vectors to predict the binary tag matrix for all 2467 TED Talks. Again, 20% of the data was assigned to the test set. Since this model was predicting the presence or absence of each topic tag, I used a Random Forest *Classifier* rather than a Regressor. After this, I created a Regressor identical to that of the second model, but using the computer-generated tags instead of the manually-assigned tags to predict the number of views. At the end, the MAE between the predicted and actual views was computed once more.

Results and Conclusion

The results from my three models that predict the number of views are presented in Fig 1. As shown by the figure, the model that did not use any topic information was much less accurate than the two models that did. This shows that topic information is informative in predicting the number of views, which was an expected result. However, it was unexpected that the computer-predicted topics would be more informative than the manually tagged topics, as shown by the positively sloped line segment in Fig 1.

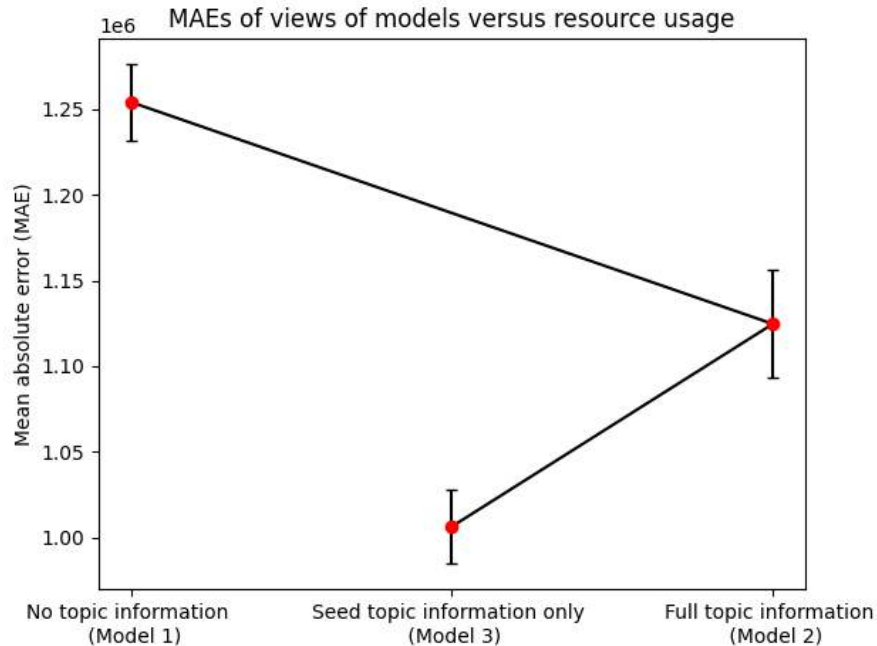


Fig 1: Comparison between the MAEs of the three models, based on and in order of amount of topic information used by the model.

To obtain my results, I created 50 test/train splits, again randomly assigning 20% of the

data to the test set, and explored the corresponding MAEs for the three models. Across the 50 trials, Model 1 returned an average MAE of 1,254,145, Model 2 returned an average MAE of 1,124,651, and Model 3 returned an average MAE of 1,006,352. To strengthen these results, I performed significance testing using the 50 trials. Two paired t-tests were conducted, one between Model 1 and Model 2 and the other between Model 2 and Model 3. I found that the MAE for Model 1 was significantly higher than that of Model 2 ($t(49) = 15.6, p = 1.2 \times 10^{-20}$). I found that the MAE for Model 2 was also significantly higher than that of Model 3 ($t(49) = 12.4, p = 9.8 \times 10^{-17}$). The error bars in Figure 1 also represent 95% confidence intervals on the MAEs of the models that estimate the number of views for each TED Talk.

Discussion

The aim of this project was to create a model effective at predicting the number of views for a given TED Talk. To complete this, I utilized the topic tags that were manually assigned to each TED Talk. I created three different models, increasing in accuracy of predictions. The first model was trained to predict the number of views based on the USE vectors of each TED Talk. The second model was trained to predict the number of views purely based on the manually assigned tags associated with each TED Talk. The third model used the USE vectors of each TED Talk to predict the associated tags, then used these computer-predicted topics to predict the number of views. I found significant differences among the three models based on the amount of topic information given to that model. Full topic information provided a better estimate than no topic information, while computer-generated topic information provided an even better estimate than full topic information. It was unexpected that the third model, which used computer-generated tags, outperformed the second model, which used manually assigned tags. This is important because it means that the tedious task of manually assigning topic tags to TED Talks and other similar mediums can be phased out and replaced by computer-generated topic tags since they provide a better estimate of the number of views on a TED Talk. A limitation of the third model is that it does not truly predict new tags, instead only predicting the binary tag matrix of all 416 existing tags. Future research could include creating a model that generates completely novel tags and uses them to predict the number of views. Additional work could also involve using computer-generated tags to create an improved TED Talk recommendation system.

Works Cited

- Banik, Rounak. "Ted Talks." Kaggle. Last Modified September 25, 2017. Accessed September 7, 2024.
<https://www.kaggle.com/datasets/rounakbanik/ted-talks/data?select=transcripts.csv>.
- Cer, Daniel, Yinfei Yang, Sheng-yi Kong, Nan Hua, Nicole Limtiaco, Rhomni St John, Noah Constant et al. "Universal sentence encoder for English." In Proceedings of the 2018 conference on empirical methods in natural language processing: system demonstrations, pp. 169-174. 2018.
doi.10.18653/v1/D18-2029
- Dochev. "Predicting Ted Talks Views with ML Models." Kaggle. Last Modified November 16, 2019. Accessed September 7, 2024.
<https://www.kaggle.com/code/dochev/predicting-ted-talks-views-with-ml-models>.
- "Ted Talks." TED. Last Modified December 2019. Accessed September 7, 2024.
<https://www.ted.com/about/programs-initiatives/ted-talks>.
- Tuarob, Suppawong, Line C. Pouchard, and C. Lee Giles. "Automatic tag recommendation for metadata annotation using probabilistic topic modeling." In Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries, pp. 239-248. 2013.
doi.10.1145/2467696.2467706

Soft Tones, Strong Identities: Mandarin Language Dynamics in Taipei, Taiwan

By Sindhu Talluri

I. Introduction

This paper explores the evolving linguistic identities in Taiwan and China, focusing on the distinctions between Taiwanese-accented Mandarin and Chinese Mandarin. While both are the same language, the sociopolitical divide between China and Taiwan has resulted in linguistic variations that are associated with sociocultural implications. By placing these linguistic distinctions within the framework of the third wave of sociolinguistic variation studies, this paper aims to analyze how these language differences index broader social meanings and contribute to the construction of identities in both Taiwanese and Mainland Chinese societies.

In recent research, linguists have become increasingly interested in the sociocultural distinctions between Taiwan and Mainland China arising from the impact of language on national and regional identities. Linguists like Penelope Eckert, through her work on third wave sociolinguistic variation, have emphasized the importance of examining linguistic features as indexical of broader social meanings and identities (Eckert 2012). As such, previous studies on accents, such as those conducted by Andrew Lynch on Spanish in Miami, have shown how linguistic variation becomes a key marker of regional social identity (Lynch 2022).

Taiwanese Mandarin (Guoyu), as a distinct variety of Mandarin, has served as a cultural emblem that differentiates Taiwanese identity from that of Mainland China, where Putonghua, the 'standard' version of Mandarin, dominates. While both linguistic varieties share common roots, the separation between the two regions has resulted in noticeable differences in pronunciation, vocabulary, and societal attitudes toward language use, even influenced by politics as with the standardization of Mandarin varieties in China.

Despite the substantial body of work on the linguistic and cultural differences between Taiwan and Mainland China, there has been little focus on the lived experiences of Taiwanese individuals who engage with Mandarin in both local and cross-strait contexts. Additionally, while previous studies have examined how Taiwanese Mandarin is perceived by Mainland Chinese speakers, there remains a gap in the research regarding how Taiwanese people themselves navigate these linguistic differences in everyday life and how these perceptions shift depending on their degree of exposure to Chinese Mandarin speakers (Juan 2011). Much of the existing research has concentrated on the phonological and lexical differences between Guoyu and Putonghua from an external perspective by treating language as a static marker of regional identity. Although it is difficult to approach due to its interdisciplinary nature, dynamic sociolinguistic research that examines how these linguistic distinctions are negotiated and recontextualized in everyday social interactions is critical to bridging Mandarin research into the third wave of linguistic studies. Specifically, the role of identity construction, interactional dynamics, and language attitudes among Taiwanese speakers has yet to be fully explored.

The purpose of this investigation is to address this gap by examining the linguistic perceptions and attitudes of Taiwanese individuals towards both Guoyu and Putonghua. This

study aims to analyze how varying levels of exposure to Chinese Mandarin speakers influence Taiwanese people's perceptions of their own language and cultural identity.

This paper is organized as follows: the next section will cover past research on related subjects, and the following will present the methodology used to gather data from participants, followed by an analysis of the results. The discussion will then contextualize these findings within the framework of third wave sociolinguistics and offer a comparison to existing research on Mandarin variation. Finally, the conclusion will address the implications of this research for future studies on language and identity, as well in the context of cross-strait relations.

In this study, I will be using the terms 'Taiwanese Mandarin' or 'Guoyu' to refer to Taiwanese-accented Mandarin. In previous literature, this Mandarin variety has been referred to as 'Gangtai qiang,' 'Taiwan qiang,' and 'Taiwan Mandarin' (Peng 2018, Juan 2011). Additionally, for the purposes of this paper, I will refer to Standard Mandarin, also known as 'Standard Chinese,' as Chinese Mandarin.

II. Literature Review

Placing this linguistic study in the context of Eckert's third wave of linguistic variation studies allows us to consider the results of this and previous studies as parts of cultural studies. Penelope Eckert indicates that this third wave of linguistic studies is broadly characterized by the tenants of indexicality and enregisterment that allow extension of linguistic implications to society and culture. The former suggests that linguistic features serve as signs that index social meaning and identities. In her paper "Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation," Eckert references Labov's study in New York City that correlated the pronunciation of the postvocalic /r/ with social class (Eckert 2012).

In line with third wave linguistics, Bucholtz et al. find that sociocultural approaches to linguistics allow deeper insight into identity and interaction (2005). For this study, the most notable principles will be those of emergence, indexicality, and enregisterment. The emergence principle indicates that identity is the product, rather than the source, of linguistic and other semiotic practices. Therefore, in third wave research, linguistic practices serve as a social and cultural phenomenon rather than a primarily internal and psychological one. Meanwhile, the indexicality principle allows linguistic features, systems, or implicatures to identify an identity within a group. The following tenant of enregisterment refers to the process through which linguistic forms or practices become associated with specific social meanings and become part of a recognized stylistic repertoire for a community of speakers.

Similarly, Andrew Lynch's work on accents of Spanish in Miami identifies regional identities through linguistic research, in line with third wave principles (Lynch 2022). Lynch's work also demonstrates the enregisterment tenant,. As such, Eckert writes, "The emphasis on stylistic practice in the third wave places speakers not as passive and stable carriers of dialect, but as stylistic agents, tailoring linguistic styles in ongoing and lifelong projects of self-construction and differentiation" (Eckert 2012). The salience of this community may also be

strengthened by communities' further identification of and with these linguistic traits, which would allow them and outsiders to associate these traits with the community's identity.

Contextualizing this linguistic study within the third wave of variation studies is critical when considering the nature of Chinese languages. Chinese languages and literature are inextricably linked to culture, both as its reflection and its product. The *Analects of Confucius*, a collection of sayings and ideas attributed to Confucius, emphasizes the application of traditional virtues in daily society. These concepts are embedded in the language and are expressed through literature, reinforcing social norms and values that persist in contemporary Chinese and Taiwanese family structures and educational systems. Lu and Chen note that "traditional values are rooted in a culture's philosophical, religious, and ideological traditions and are called upon by the social need and context of a culture" (Lu et al. 2011). As such, principles rooted in Confucian thought remain central to modern Chinese social norms and practices.

In both China and Taiwan, Confucian teachings remain significant. Despite rapid modernization and economic development, these values shape ethical standards and social behaviors. The emphasis on education and moral development in both societies reflects the enduring influence of Confucian ethical idealism. In China, this is evident in the rigorous academic expectations placed on students and the moral education integrated into the school curriculum. Similarly, in Taiwan, the emphasis on such morals are explicitly linked to Confucian teachings and literature, very clearly evident when considering the statue of Confucius outside the entrance to National Taiwan Normal University. Thus, combining these beliefs under third wave linguistics connects existing linguistic circumstances to cultural dynamics.

Qing Zhang's study, "Rhotacization and the 'Beijing Smooth Operator': The social meaning of a linguistic variable" provides a similar framework through which to view this linguistic study via the analysis on the Beijing Mandarin accent. Zhang found that Beijing accent linguistic features like rhotacization are not neutral, but carry with them layers of social and cultural connotations (2008). Zhang demonstrates that these meanings are constructed and reinforced through their frequent use in cultural terms and literary representations, making rhotacization a salient emblem of Beijing localness. The association of rhotacization with the 'Beijing Smooth Operator' character type, a stereotypical character known for his smooth-talking, street-smart persona, often depicted in 'Beijing-flavored' literature, exemplifies how linguistic features can reflect and construct social identities (Zhang 2008).

The study discusses how both implicit and explicit metapragmatic acts contribute to the social meaning of rhotacization. Literary critics add rhotacization to characters' speech in their analyses to emphasize their smooth-talking nature, performing explicit acts that anchor the feature to a marker of an authentic Beijing style, which includes attributes like worldly wisdom, urban versatility, and social astuteness (Zhang 2008). Implicitly, the pervasive use of rhotacization in representations of Beijing speech and characters in literature also reinforces its social significance, contributing to the further association of these features with Beijingers. Thus, these linguistic traits are ideologically linked to the social image of Beijingers, solidified as the male 'smooth operator' persona.

Additionally, rhotacization frequently co-occurs with key cultural terms that denote distinctively Beijing concepts. Terms like "jing qiang" (Beijing tune) and "jing diao" (Beijing tone) are used to describe the Beijing accent, both often indicating rhotacization. This consistent co-occurrence helps to solidify the association between rhotacization and the authentic Beijing style of speech (Zhang 2008). The term "jing wei," meaning "Beijing flavor," is another example where rhotacization is used to invoke the local culture and speech, reinforcing the connection between the linguistic feature and local identity.

Zhang's study emphasizes the importance of understanding the historical and social context of linguistic features to fully grasp their meaning potential. The findings highlight that sociolinguistic variables like rhotacization have pre-existing social meanings that can be accessed and recontextualized in new interactions with and between speakers. Taken in the context of this study of Taiwanese Mandarin, Zhang's research provides an example for the scope of sociocultural studies that can be derived from a linguistic investigation.

Juan's study on the Chinese teenagers' perceptions of and attitudes toward Taiwan Mandarin concludes that Taiwan Mandarin is perceived as a non-standard variety of Mandarin Chinese by the majority of mainland Chinese teenagers, yet it garners positive attitudes from them (2011). This dichotomy between non-standardness and positive perception suggests that the traditional association of standardness with social prestige does not hold strong in the Chinese context. Instead, Taiwan Mandarin's perceived desirability and pleasantness stem from cultural influences and interpersonal interactions rather than linguistic purity.

Juan's research indicates that the nature and quality of exposure to Taiwanese popular culture and direct communication with Taiwan Mandarin speakers play pivotal roles in shaping these perceptions (2011). Positive experiences with Taiwanese media and personal interactions with Taiwan Mandarin speakers contribute to favorable perceptions of the variety and its speakers. One such vehicle for exposure is through mediatization of Taiwan. Mainland Chinese teenagers are significantly exposed to Taiwanese popular culture, as many popular Taiwanese singers, actors, and TV shows have a broad fan base in mainland China. This cultural consumption allows mainland teenagers to develop a familiarity with Taiwan Mandarin, often associating it with positive, trendy, and modern attributes. For instance, Taiwanese dramas and pop songs frequently use Taiwan Mandarin, making it an integral part of the entertainment experience. The study indicates that those who are regular consumers of Taiwanese media are more likely to perceive Taiwan Mandarin as pleasant and friendly, associating the language with the positive emotions and enjoyment derived from these media experiences (Juan 2011).

Furthermore, the study reveals that the differences in phonological features between Taiwan Mandarin and Mainland Standard Mandarin are the most salient to the respondents. These differences, such as the lack of rhoticity in Taiwan Mandarin, are frequently mentioned and align with academic discourses on the subject (Juan 2011). However, lexical differences, while academically significant, are less noticed by the respondents, highlighting a gap between linguistic scholarship and public awareness, suggesting a lessened importance on regional attitudes.

Gao's research highlights that the use of full tones in Mandarin, typically associated with non-standard varieties like Taiwanese Mandarin, is increasingly adopted by young female speakers in mainland China to construct a feminine persona, often referred to as 'sajiao' (2020). The study found that female speakers who present themselves as cute and feminine in their vlogs frequently use full tones, a pattern not observed among those who present a more independent and strong-minded persona. This suggests that full tone realization is a strategic linguistic choice to enhance a specific social image, aligning with the broader East Asian cultural trend of emphasizing cuteness, seen in Japanese 'kawaii' and Korean 'aegyo' speech styles.

The study also reveals a clear gender difference in the use of full tones. While female speakers who adopt a cute persona use full tones extensively, male speakers generally do not exhibit the same pattern, regardless of their social identities (Gao 2020). This gender-specific usage indicates that full tones are more closely associated with feminine speech and the social persona of cuteness, rather than being a general feature of cosmopolitan identity, as previously suggested by other studies. The female speakers' use of full tones varied significantly with the topics they discussed, further supporting the idea that this linguistic feature is contextually driven to align with the persona they wish to project (Gao 2020).

Gao situates these findings within the broader sociocultural context of contemporary China, where economic reforms and globalization have introduced new lifestyle influences and ideologies of femininity. The influx of non-mainland Chinese popular culture, particularly from Taiwan and Hong Kong, has played a crucial role in shaping these new speech practices. The full tone realization, originally a feature of Taiwanese-accented Mandarin, has been adopted by young mainland Chinese women to convey a cosmopolitan and youthful image. This adoption reflects the dynamic nature of language ideologies, where linguistic features are recontextualized to serve new social functions and identities.

III. Methodology

This research utilizes a qualitative approach, focusing on semi-structured interviews, to explore the linguistic perceptions and cultural attitudes surrounding Guoyu and Putonghua. The study's participants were categorized into three distinct groups following their interviews based on their level of interaction with individuals from mainland China: (A) Taiwanese individuals with limited interaction, (B) Taiwanese individuals with frequent interaction, and (C) Chinese immigrants living in Taiwan. By selecting participants from various backgrounds and age groups, ranging from their 20s to their 70s, the study captured a wide range of perspectives on the linguistic and cultural distinctions between Taiwan and China.

The interviews were conducted one-on-one in Mandarin and took place at locations including street markets, universities, and professional environments across Taipei. Each interview was audio recorded with the participants' consent.

The questions were designed to contextualize participants' perceptions of linguistic differences, specifically addressing pronunciation, tone, and cultural attitudes between Guoyu and Putonghua. For example, participants were asked to discuss specific topics in order to lead

into conversations regarding if they believed they could identify different Mandarin varieties, whether they perceived cultural or communicative differences between people from Taiwan and China, and how their interactions with Chinese individuals influenced their views on language and culture. The use of looser discussions was to prevent interference from leading questions, and as such, some topics that were used to begin these discussions, included discussing their understanding of the terms ‘Guoyu’ and ‘Putonghua,’ as well as their personal experiences with speakers of each variety. Additionally, participants were encouraged to share their thoughts on the historical and political contexts that might have shaped their language perceptions. This open-ended approach allowed for natural conversation flow and ensured that the participants could express their views freely, minimizing bias and leading responses.

The data was analyzed using thematic analysis, which identified recurring themes in the context of the purpose of this study coupled with ideas from previous literature addressed in the previous section of this paper. These findings were then situated within the framework of third wave sociolinguistics to reveal how linguistic variations reflect and shape broader social identities and cultural dynamics between Taiwan and mainland China.

IV. Results

In this section, I will discuss specific interviews broken down into three groups. Through my interviews, I identified three broad groups based on their interactions and perceptions of Mandarin and culture: Taiwanese with little interaction with Chinese individuals, Taiwanese with more interaction with Chinese individuals, and Chinese immigrants to Taiwan. The first group, predominantly older individuals (60 years and above), perceived no significant differences between Taiwanese and Chinese Mandarin and culture. The second group, consisting of Taiwanese aged 20 to 60 who had interacted with Chinese people through daily interactions, vacations to China, or studying abroad, identified notable differences in pronunciation and attitudes between Chinese and Taiwanese. The third group comprised Chinese immigrants to Taiwan, who had relocated for various reasons, including economic and social motivations.

Group A consisted of Taiwanese people with little daily interaction with Chinese individuals, who had never traveled to China.

Speaker one in Group A is a woman in her late 70s, born and raised in Kaohsiung, fluent in Mandarin and Minnan, working on Dadaocheng’s Dihua Old Street. The woman agreed with the statement that she was able to distinguish between the Chinese and Taiwanese based on how they spoke Mandarin. When asked about the differences in accent between Putonghua and Guoyu, she said that the “Mainlanders speak badly - they pronounce the words incompletely, and sound very harsh in their speech”. However, she did not think there were major cultural differences between Taiwan and China, instead finding that Taiwanese culture was just an offshoot of Chinese culture. When asked further, the woman thought that there were also minimal personality differences between Chinese and Taiwanese people, as they were from the same culture. Thus, this interview still classified her as Group A, both by age range and limited interaction with the Chinese. However, she was able to identify a difference between Chinese

and Taiwanese Mandarin accent. This woman found that her exposure to the Chinese was limited to tourists who stopped by her stall, in a historical area of Taipei, carrying many traditional goods. However, other vendors in the same area found no difference in accent.

Speaker two in Group A is a man on Dihua Old Street, a well-known peanut ice cream vendor. The vendor, in his early seventies, was born and raised in Taiwan, and found that there was no difference between Chinese and Taiwanese culture and personality. He cited that Dihua Old Street attracts significant international tourist traffic, and has many traditional Taiwanese stalls, including the vendor involved in the previous interview. As such, these vendors interact with Chinese people on occasion. Similarly, when asked about whether he could identify the difference between Chinese and Taiwanese people based on their Mandarin accent, he said that he could hear differences in speech, but the differences were not particularly notable. His younger apprentice agreed with him throughout the interview, however cited that there were differences in Chinese and Taiwanese lifestyles, explaining that she had visited China in the past.

Speaker three from Group A is a vendor in his late 70s working at a street market frequented by Taiwanese and international students from surrounding universities, citing having minimal interaction with Chinese people daily, and never having visited China. This person, any time China was referenced, preferred to refer to it as “大陆” (mainland). When asked about differences in attitude and culture, he believed there were none. Additionally, he was confused by questions regarding differences in Mandarin spoken in Taiwan and China, saying that Putonghua and Guoyu were the same language, and responded the same when asked more specifically about accent differences. However, he did say he was able to tell the difference between a Chinese and Taiwanese person based on “their manner of speaking”.

The first interview in Group B was with two women nearby a local university, who interviewed together in a casual format. They were both born in Taiwan, and had lived in Taiwan for most of their lives. They referenced having frequent exposure with people from China, indicating having Chinese friends and colleagues that they connected with on a daily basis. When answering the first question about accent differences, they found that the Chinese Mandarin accent was broadly harsher than the Taiwanese Mandarin accent, citing rhotacization and tone sandhi. One of the women said that the Chinese ‘roll their tongue’ (卷舌) more than the Taiwanese when pronouncing the retroflex ‘r’ sound. compared to the Taiwanese pronunciation of ‘s’ and ‘z’ syllables. When asked about personality or attitude differences, they found that the Chinese sometimes came off as ‘rude’ or more direct as compared to the Taiwanese attitude and personality. However, they did cite that this was a generalization of personality, backtracking to note that this idea was not applicable to every Chinese person, as it was also not applicable to every Taiwanese person. Finally, they found that the society culture was different in China than Taiwan, and thus this perceived rudeness would not be considered rude in China, simply the accepted behavior of Chinese society which was generally more direct.

Speaker two in Group B was with a woman in her 20s, working as a teaching assistant at a local university. She begins by explaining that her grandfather was originally from mainland China but moved to Taiwan early in life, which led him to adopt more Taiwanese cultural traits

and a less “China-centric” mindset. While he retained some of his mainland Chinese accent, the influence of Taiwanese culture is evident in his way of thinking. As a result, the speaker is still able to understand and communicate with him, though she notices a clear distinction between how he speaks and how native Taiwanese people converse. This early exposure to both accents and cultural perspectives gave the speaker an awareness of the subtle differences between the two regions.

The speaker further discusses how, during her middle school years, she and her friends began using mainland Chinese social media platforms and watched various online content from China. This exposure helped her become familiar with mainland Chinese accents and simplified Chinese characters, making it easier for her to understand mainland Chinese speech patterns. However, despite this familiarity, she elucidates the clear differences in communication styles and personality traits between people from Taiwan and mainland China. She observes that mainland Chinese people tend to be more direct in their communication. They are more likely to publicly express their dissatisfaction or opinions without hesitation, such as saying, “I think this is wrong,” or correcting someone’s actions in a straightforward manner. In contrast, Taiwanese people, in her experience, are generally more indirect and would likely address such matters privately rather than in a public setting.

When asked about her frequency of interaction with mainland Chinese people, the speaker notes that she seldom encounters them in her everyday life, although has had interaction with them in the past, simply by living in an area with many international universities. She found that Chinese individuals are more commonly seen in tourist areas or popular travel destinations. The exception is when someone actively seeks out interactions through online platforms or specific apps, where they may have more exposure to people from mainland China.

She also touches on her experience of being recognized as Taiwanese by Chinese people during these interactions. She mentions that mainland Chinese individuals tend to take a more authoritative or directive stance in conversations. This dynamic can make the speaker feel like the interaction is less of an equal exchange and more like following instructions, which hinders the ability to form genuine friendships. According to the interviewee, this contrasts with how Taiwanese people generally approach discussions. She highlights that Taiwan has a greater openness to accepting different cultures, which translates into a higher degree of tolerance and flexibility in dealing with various viewpoints. Taiwanese people, in her view, are less likely to insist on being the dominant voice in a conversation and are more willing to listen to different perspectives before presenting their own opinions.

Furthermore, she attributes some of the contrasts to historical and societal backgrounds. Mainland Chinese people, she notes, may be more skeptical or distrusting of others due to their historical and political context. Linguistically, she highlights that while both groups speak Mandarin, the way they speak differs significantly due to geographic separation and linguistic influences. Taiwanese Mandarin, for example, has been influenced by other languages, such as Minnan and Japanese, leading to a softer, more fluid way of speaking. Taiwanese conversations often lack the prominent “erhua” (儿化) sounds found in northern Chinese speech, and they

incorporate words from other languages, which makes Taiwanese speech sound distinct from Chinese Mandarin.

Speaker three is a Taiwanese woman in her late thirties who lived and worked in Shanghai in her twenties. The speaker describes her experiences navigating cultural and linguistic differences between Taiwan and China, drawing on both personal and professional experiences. She begins by discussing the relative freedom of expression in Taiwan, emphasizing how people in Taiwan feel more comfortable voicing their opinions, particularly on political matters. In contrast, she notes that people in China tended to be more reserved and limited in their political expressions due to the more controlled environment. The speaker attributes this to the fact that news and information in China are often filtered by the government, which in turn shapes a more narrow worldview. In Taiwan, by comparison, people are encouraged to engage in open discussions on a wide range of topics, especially politics. She points out that, even though not all Taiwanese people are passionate about politics, there is still a cultural tendency to openly criticize or support ideas without fear.

The speaker then shifts to a discussion of the linguistic diversity between the regions. While Mandarin, specifically the Beijing dialect, is the official language in China, she emphasizes that China is home to many other dialects, such as Minnan and Hakka, especially in the southern regions. She explains that this linguistic variety also contributes to cultural differences. In northern China, for example, the accent tends to be stronger, and the speaker notes that this correlates with a more assertive personality. On the other hand, southern Chinese people, with their softer dialects, tend to have a more gentle communication style. In Taiwan, Mandarin has evolved over time, integrating elements from other languages like Japanese and Minnan due to Taiwan's unique historical background. This has led to a smoother, less rigid way of speaking, where people often use softening particles like “ㄚ” (-a), “ㄟ” (-ei), and “ㄠ” (-ou) to sound more friendly and approachable. She believes that these linguistic nuances contribute to a more welcoming tone in Taiwanese conversations, compared to the more straightforward approach often found in mainland China.

She goes on to discuss the differences in work culture between Taiwan and China, reflecting on her time living and working in Shanghai. Taiwanese workers, she explains, tend to have a strong sense of responsibility and are often willing to work beyond regular hours to ensure tasks are completed to a high standard. In contrast, she observed that some Chinese workers adhere more strictly to working hours, and once their shift ends, they are less inclined to stay late unless absolutely necessary. The speaker sees this as part of a larger cultural difference, where Taiwanese people are more accustomed to a "責任制" (responsibility system) that drives them to finish what they start, while Chinese workers may prioritize work-life balance differently.

Additionally, she notes that in more competitive cities like Shanghai and Beijing, there is a strong sense of ambition among Chinese workers. Some are eager to learn from Taiwanese professionals with the goal of eventually replacing them in their roles. This sense of competition creates a high-pressure work environment in these first- and second-tier cities, where

advancement opportunities are fiercely pursued. Yet, for those in lower-tier management roles, the speaker observed a more laid-back attitude toward work, where some employees feel that their salary and position limit their motivation to excel or seek promotion.

The first interviewee in Group C was a woman in her late twenties who had moved from southeast China to Taipei three years prior, working in the Taipei 101. She notes that just as there are variations in accents between northern and southern Taiwan, there are also distinct speech patterns in different regions of China. She finds that in Taiwan, the way people speak is generally more gentle and soft, and describes how this tone differs from the faster, harsher way of speaking often found in northern China, where people tend to speak loudly and with a sense of urgency, though she believed this to be more a matter of habit than actual intent to be aggressive. She also highlights the difference in interpersonal interactions, using the idiomatic phrase "平易近人" (approachable and easygoing) to describe how Taiwanese people generally interact with one another in a friendly, welcoming way. This reflects the overall friendliness and warmth that she associates with Taiwanese communication and culture, in contrast to the sometimes more abrupt speech styles in northern China.

The second speaker in Group C is a woman in her early forties who had moved to Taipei from Northern China eight years prior. She highlights the cultural and social differences between Taiwan and mainland China, emphasizing the contrast in politeness and communication styles. In Taiwan, the speaker notes, people are generally polite and quick to express gratitude, creating a more open and welcoming environment. This differs from her experience in mainland China, where help is available but not as readily offered, and communication tends to be more direct and assertive, especially in northern China.

Language also plays a key role in the speaker's adaptation. Although she is from northern China, after years of living in Taiwan, her accent is often mistaken for a southern Chinese one, which aligns more closely with Taiwanese speech. The speaker discusses how these linguistic nuances influence social dynamics, and how being perceived as having a southern accent allowed her to blend in more easily. The speaker further contrasts the personality traits of people from northern China and Taiwan. She describes northern Chinese people as bold and straightforward, while Taiwanese people are more reserved and prefer indirect communication. Despite initial difficulties in making friends in Taiwan, particularly due to her focus on raising children, the speaker eventually built strong relationships with other mothers. She found that, contrary to some stereotypes, people in Taiwan were easy to connect with once she got past initial cultural differences, and were not abrasively political once hearing of her background.

The third interview in Group C is with a Chinese professor who is married to a Taiwanese woman, both living in the United States. The professor shares his background, starting with his journey to the U.S. for higher education in the late 1980s, where he also met his wife, a Taiwanese woman. He reflects on their shared experience of studying abroad, which provided a neutral space for them to appreciate each other's expertise and cultural background.

One of the main themes of the interview is the professor's perception of cultural differences between China and Taiwan, which he notes have sometimes been exaggerated. He

emphasizes that while there are distinctions in language and expressions, these differences are often more about formality in speech rather than deep cultural divides. For example, he explains how certain words and phrases in Taiwan are considered more formal compared to their use in mainland China, where speech tends to be more casual and direct.

He also touches upon Confucian values, pointing out that they have been better preserved in Taiwan than in mainland China due to Taiwan's relatively stable political environment. In contrast, mainland China's various political campaigns have disrupted traditional cultural continuity. He mentions that his wife would often quote the Analects to him, later saying that he would respond by quoting what he remembered from the Little Red Book. Regarding their personal lives, the professor explains that although his wife is from Taiwan, she was never perceived as different when they traveled to China because they share a similar appearance and language.

He believes that Taiwanese Mandarin (Guoyu) is often indistinguishable from standard Mandarin in China unless people intentionally speak with localized accents. However, nearing the end of his interview, he also discusses stereotypes, noting that Taiwanese are often perceived as more "soft-spoken" and polite compared to people from mainland China. He attributes this difference to cultural upbringing, where Taiwanese are trained to be considerate of others in public spaces, reflecting Confucian values of respect and social harmony. Later in the interview, he revisits his statement, elaborating that this soft spoken nature especially applies to the Taiwanese Mandarin accent. He expresses distaste for this accent, and says that "some Taiwanese boys I hear speak like women, the way they speak is high-pitched and lisping their pronunciation."

Toward the end of the interview, he elaborates on the changing perceptions of Taiwan and China, noting that younger generations in Taiwan are increasingly reluctant to identify with China, reflecting a growing sense of Taiwanese identity. He goes on to state that he believes part of the reason for a drastic difference in the Taiwanese and Chinese Mandarin accents is that as the Taiwanese younger generation attempts to distinguish themselves culturally from China, they exaggerate the differences that do exist. As such, he states that the younger Taiwanese generation emphasize the full tone and 'sajiao' pronunciation to distinguish themselves from Chinese Mandarin.

IV. Discussion

The findings from Group A, which consists of older Taiwanese individuals with minimal interaction with Chinese people and no travel experience to China, reveal complex attitudes towards linguistic and cultural distinctions between Taiwan and China. Despite limited exposure, many of these individuals demonstrate an awareness of linguistic differences, particularly in the accents of Putonghua and Guoyu, yet they largely perceive cultural and personality unification between the two groups. This paradox highlights the nuanced nature of linguistic and cultural perceptions in Taiwan, especially among those with limited direct experience with Mainland Chinese people.

One observation from Group A is the ability of some members to distinguish between Chinese and Taiwanese accents despite their minimal exposure to Chinese individuals. For example, the female vendor from Dihua Street noted that “Mainlanders speak badly—they pronounce the words incompletely, and sound very harsh in their speech.” This comment indicates an awareness of the phonological differences that exist between Putonghua and Guoyu, suggesting that even limited interactions, such as brief encounters with tourists, can reinforce perceptions of linguistic distinction. However, the significance of these differences varies among individuals, as seen with the peanut ice cream vendor who acknowledged differences but did not find them particularly notable in terms of culture or identity.

While linguistic distinctions were broadly recognized in the context of cultural and personality differences between Taiwanese and Chinese, Group A largely did not place importance on these distinctions despite recognizing them. The vendors interviewed consistently viewed Taiwanese culture as an offshoot of Chinese culture, emphasizing a shared cultural heritage rather than focusing on distinctions. This perception aligns with the historical narrative that sees Taiwan as culturally linked to China, especially among this generation, despite the political and social differences that have emerged over time. The vendor from Taichung, for instance, did not recognize significant cultural or attitudinal differences, preferring the term Mainland to refer to China, which implies a continued identification with a broader Chinese cultural sphere.

While younger generations in Taiwan may be more inclined toward a convergence in language attitudes due to increased interaction with modern Chinese culture, older generations might feel a closer connection to China for historical reasons. Many older Taiwanese individuals, particularly those from families that immigrated to Taiwan in 1949 after the ROC government relocated, may retain a strong sense of cultural and historical ties to the mainland (“History of Taiwan”). These ties are often rooted in personal or family histories of displacement, loyalty to the Republic of China, and the shared experience of fleeing the Chinese Civil War. For these individuals, Taiwan represents a continuation of Chinese culture rather than a departure from it, leading to a perception of cultural continuity with the mainland. This historical connection may result in a more pronounced sense of shared identity with China, making these older individuals less likely to emphasize linguistic and cultural differences between the two regions. As a result, their perceptions of Taiwan as an offshoot of Chinese culture might reflect a generational perspective shaped by their lived experiences and the political context of their youth.

The limited exposure of Group A to Chinese individuals affects their perceptions. The occasional interactions they have, mostly with tourists, seem insufficient to cultivate a deeper understanding or awareness of the evolving linguistic and cultural distinctions between Taiwan and China. Instead, these limited encounters reinforce a perception of continuity and similarity, which contrasts with findings from Group B, where more frequent interactions with Chinese individuals led to a greater recognition of linguistic and cultural differences. Thus, the level of exposure and interaction appears to be a key factor in shaping how Taiwanese individuals perceive their cultural and linguistic relationship with China.

The findings from Group B, which highlight more frequent interactions with Chinese individuals, can be contrasted with those of Group A, where limited interaction results in a less distinct perception of differences. Group A's limited exposure to mainland Chinese people, primarily through occasional interactions with tourists, seems to minimize the perception of linguistic and cultural distinctions between Taiwanese and Chinese. This suggests that the frequency of exposure not only influences the perception of similarities but also plays a significant role in maintaining or even amplifying perceived differences when interaction is minimal.

In contrast, Group B, composed of Taiwanese individuals who frequently engage with people from Mainland China, demonstrates a more nuanced understanding of the differences between Chinese and Taiwanese Mandarin. Through regular exposure—whether via travel, media, or personal connections—these individuals become more attuned to the linguistic and cultural distinctions between the two regions. This greater exposure sharpens their awareness of how the softer tones of Guoyu and the more direct intonation of Putonghua serve as markers of identity. As they navigate cross-strait interactions, the contrasts between Taiwanese and Mainland Chinese communication styles, social behaviors, and cultural attitudes become more evident, reflecting the sociolinguistic trend that identity is continuously shaped by lived experiences and social contexts.

The interviews with members of Group B reinforce this idea. For example, the two women who regularly interact with Chinese colleagues observed that the “harsher” sound of Chinese accents, with features like rhotacization and the retroflex ‘r,’ contrasts with the smoother tones of Taiwanese Mandarin. This directness in speech mirrors cultural traits they associate with Chinese people, such as being more assertive or, at times, perceived as impolite. However, they were quick to note that such characteristics are not universal and are often a reflection of different societal norms rather than individual personality traits.

Juan's research on language attitudes highlights the importance of both content and context in shaping perceptions, emphasizing that positive experiences through media or personal interactions play a more significant role than just exposure alone (Juan 2011). However, the findings from Group A, which had limited exposure to Chinese culture, suggest that frequency of interaction is equally critical in developing distinct language attitudes. Respondents in Group A, despite being aware of linguistic and cultural differences, often perceived only minimal distinctions between Chinese and Taiwanese Mandarin. This lack of regular, meaningful interaction contributes to a more homogenized view of language and culture, where the distinctions are downplayed or overlooked.

The comparison between Groups A and B demonstrates that while positive experiences are often predicated on a foundation of frequent exposure, in contexts where exposure is limited, even the most meaningful interactions may not be sufficient to foster a deep understanding of linguistic differences. Juan's findings, therefore, take on new significance when viewed through the lens of Groups A and B. They suggest that frequency of exposure is a prerequisite for the formation of language attitudes. Without this foundational exposure, language attitudes may

remain largely unaffected, as seen in Group A. On the other hand, in environments where exposure is frequent, as in Group B, the richness of content and context can fully shape and refine these attitudes, leading to a more divergent understanding of linguistic and cultural distinctions.

The interviews conducted with Group C, Mainland Chinese individuals with significant interaction with Taiwanese speakers, provide insight into how linguistic variation, particularly between Guoyu and Putonghua, reflects and reinforces broader socio-cultural divides. One of the most salient themes emerging from these interviews is the perceived tonal and stylistic differences between Taiwanese and northern Chinese Mandarin. The woman working at Taipei 101 describes Taiwanese Mandarin as softer and more gentle in comparison to the faster, harsher tone she associates with northern Chinese Mandarin. This observation mirrors the broader cultural attitudes often associated with Taiwanese politeness and northern Chinese assertiveness. The softer tone of Taiwanese Mandarin, as noted by several participants, indexes a politeness and interpersonal warmth that are often associated with Taiwanese identity.

Interestingly, this perception of Taiwanese Mandarin as more approachable is associated with broader personality traits and communication styles. For instance, the second speaker draws attention to the interpersonal warmth that defines Taiwanese social interactions. She contrasts this with the direct and assertive communication styles of northern China, where conversations can be more abrupt or transactional. She demonstrates the role of language in social indexing, where specific speech patterns signal broader cultural characteristics like approachability or assertiveness, shaping perceptions of identity on both sides of the Taiwan Strait.

This adaptive process also highlights the interactional dynamics between linguistic identities. The second speaker felt that her southern-accented Mandarin (as perceived by Taiwanese speakers) helped her blend in more easily, suggesting that Taiwan's sociolinguistic landscape is not just shaped by static regional markers but is also responsive to variations within Mandarin itself. This adaptability is consistent with Eckert's view that speakers are not passive carriers of dialects but active agents in their identity construction through linguistic styles. In this case, the perceived fluidity between northern and southern accents enables greater flexibility in social integration, particularly in cross-strait interactions.

The sociopolitical divide between Taiwan and Mainland China has influenced perceptions of language, as demonstrated in the interviews. The professor discusses the intentional exaggeration of linguistic differences by younger generations of Taiwanese speakers, particularly the 'sajiao' pronunciation, which he sees as a deliberate attempt to differentiate from Mainland Chinese Mandarin. His belief of the causes of this linguistic divergence is not just a matter of phonetic variation but reflects broader political and cultural efforts to assert Taiwanese identity in opposition to China.

This observation echoes Zhang's study of rhotacization in Beijing Mandarin, where linguistic features index social and cultural meanings that go beyond their phonological form (Zhang 2008). Just as rhotacization has become a symbol of Beijing localness and urban sophistication, the 'sajiao' tones of younger Taiwanese speakers signal a distinctive Taiwanese

identity, shaped by both political resistance and cultural pride. The professor's distaste for the high-pitched tones he associates with some Taiwanese men points to the gendered nature of these linguistic features, further complicating how language indexes not only national identity but also gender norms and social expectations.

Another finding that emerged from the interviews is the gendered dimension of linguistic variation. The professor's remarks about Taiwanese men speaking in a high-pitched, almost feminine manner are reflective of broader sociolinguistic trends in East Asia, where certain linguistic features become associated with gendered identities. This observation parallels Gao's research on how young female vloggers in China use full tones to construct a "cute" and feminine persona, borrowing from Taiwanese Mandarin (Gao 2020). The professor's discomfort with this feminization of speech suggests that language plays a significant role in reinforcing or subverting traditional gender norms, with certain speech patterns being seen as inappropriate for specific gender roles.

This gendering of linguistic features also ties into broader social perceptions of power and identity. In Taiwan, where softer, more indirect speech is often associated with femininity and politeness, the use of such tones by men can be perceived as subversive, particularly by those from more direct, assertive speech communities like northern China. This reinforces the idea that language is not only a marker of national or regional identity but is also deeply intertwined with gendered expectations and the performance of social roles.

The professor's reflection on the Confucian values preserved in Taiwan also speaks to the ways language embodies historical and cultural continuity. In his view, Taiwanese Mandarin, with its politeness (also perceived as indirectness) and formal expressions, carries with it a sense of cultural preservation that contrasts with the more direct and colloquial language used in Mainland China, where political upheaval disrupted Confucian traditions. This dichotomy aligns with the broader theme of Taiwan as a repository of traditional Chinese culture, where language becomes a vehicle for maintaining social and ethical norms rooted in Confucian philosophy (Lu et al. 2011).

V. Conclusion

This study attempts to elucidate part of the intricate relationship between language, identity, and cultural distinction between Taiwan and China. By examining Taiwanese Mandarin (Guoyu) and Chinese Mandarin (Putonghua) through the lens of third wave variation studies, this research reveals how linguistic features are not merely neutral modes of communication but powerful markers of identity, cultural alignment, and sociopolitical stance. The subtle phonological differences between these varieties of Mandarin serve as vehicles through which Taiwanese individuals assert their identity, often in contrast to their Chinese counterparts. For many Taiwanese, the softer tones of Guoyu are not just a linguistic variation; they symbolize an evolving sense of Taiwanese distinctiveness in the face of political tension and historical ties with China.

The findings highlight that perceptions of language and identity are deeply shaped by the level of exposure to cross-strait interactions. Those who have limited contact with Mainland Chinese individuals tend to perceive fewer differences, while those with more frequent interactions are acutely aware of the linguistic and cultural divergences. This demonstrates the idea that language ideologies and identity are not fixed: they are continuously reshaped through lived experiences, social contexts, and cultural exchanges. It is within these interactions that identity takes form, as language becomes a symbol for broader societal and cultural currents.

However, this study has its limitations, primarily in terms of scope and participant demographics. The research focuses on urban centers like Taipei and does not fully capture the diversity of linguistic experiences across different regions of Taiwan or mainland China. To develop a more nuanced understanding, future research could incorporate a more geographically and socially diverse range of participants. Additionally, while this study has provided rich qualitative data through interviews, future studies might incorporate quantitative measures, such as acoustic analysis, to further investigate the phonological variations and how they are perceived across different demographic groups.

One area for further exploration is the generational divide in linguistic attitudes. As Taiwan continues to assert its political and cultural independence, younger generations, in particular, are developing a linguistic identity that diverges from traditional Chinese roots. A longitudinal study could provide invaluable insights into how these shifts are evolving over time, particularly in light of Taiwan's changing geopolitical stance.

This study also offers important implications beyond the context of Taiwan and China by demonstrating language as a representation of culture and interaction. As linguistic differences emerge and evolve, they become enmeshed in broader questions of belonging, exclusion, and power.

As Wei (2008) argues, "The hegemonic myth of creating a standard language, an ideal speaker, and an invariant community rooted in the twentieth century obscures much of what is really happening in Chinese communities - whether in China, Hong Kong, or Taiwan. This essentialist mentality in defining nation and language is not only scientifically untenable; its political effects militate against recognition and respect for genuine ethnic and linguistic diversity."

This sentiment captures the heart of the matter: language is not, and should not be, viewed as a static, singular entity that rigidly defines who we are. In reality, multilingualism, regional dialects, and linguistic diversity are reflections of the complexity and richness of human experience. To hold onto narrow, essentialist definitions of language and identity is to obscure the true nature of human diversity. When we insist on a singular 'correct' language or identity, we suppress the fluid, multifaceted ways in which people interpret and interact with the world.

As Taiwan continues to assert its distinct identity on the global stage, the role of language in this process will only grow more pronounced. The implications of this study extend beyond Taiwan and China, offering a lens through which to understand how linguistic variation serves as a reflection of and response to shifting socio-political landscapes. Future research can build on

these findings to further explore how language continues to evolve in response to global and local influences, particularly in regions where identity is contested or fluid.

Works Cited

- Bucholtz, Mary, and Kira Hall. "Identity and Interaction: A Sociocultural Linguistic Approach." *Discourse Studies*, vol. 7, no. 4–5, 2005, pp. 585–614. Sage Publications.
- Eckert, Penelope. "Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation." *Annual Review of Anthropology*, vol. 41, 2012, pp. 87–100.
- Gao, Feier. "Full Tone to Sound Feminine: Analyzing the Role of Tonal Variants in Identity Construction." *University of Pennsylvania Working Papers in Linguistics*, vol. 25, no. 2, 2020, Article 6. University of Pennsylvania.
<https://repository.upenn.edu/pwpl/vol25/iss2/6>.
- "History of Taiwan." [Taiwan.gov.tw](https://www.taiwan.gov.tw), Government Portal of the Republic of China (Taiwan), https://www.taiwan.gov.tw/content_3.php#:~:text=Dutch%20and%20Spanish%20settlers%20established,late%201940s%20and%20early%201950s. Accessed 8 Sept. 2024.
- Huang, Kuan-Hsin. *Neutral Tone Syllables in Taiwan Mandarin*. Dissertation, University of Hawai'i, 2016.
- Juan, Sherri Chiung-Hsiang. *Chinese Teenagers' Perceptions of and Attitudes toward Taiwan Mandarin*. 2011. Dissertation, University of Iowa.
- Lu, Xing, and Guo-Ming Chen. "Language Change and Value Orientations in Chinese Culture." *China Media Research*, vol. 7, no. 3, 2011, pp. 56–63. University of Rhode Island.
https://digitalcommons.uri.edu/com_facpubs/4.
- Peng, Chun-Yi. "Mediatized Taiwanese Mandarin: Popular Culture, Masculinity, and Taiwan Studies." *Sinophone and Taiwan Studies*, vol. 2, no. 1, 2021, pp. 1–20. Springer Nature.
- Wei, Jennifer M. *Language Choice and Identity Politics in Taiwan*. Lexington Books, 2008.
- Zhang, Qing. "Rhotacization and the Beijing Smooth Operator: The Social Meaning of a Linguistic Variable." *Journal of Sociolinguistics*, vol. 12, no. 2, 2008, pp. 201–222. Wiley-Blackwell.

Etruscan Burial Methods By Shiv Sibal

Abstract

This research paper will explore how the Etruscans buried their dead and what they were buried with. This will happen through an analysis of the burial chambers, as well as the items in the chambers. Based on the findings, it is clear that the wealthy Etruscans put a lot of effort into their burial chambers to ensure the deceased have a better chance in the afterlife.

Introduction

This research paper will look at how the Etruscans bury their dead. Specifically, it will focus on some of the items in the Etruscan burial chambers, especially those that were created for no other use but for burial. Some Etruscans put vases, mirrors among other items into their burial chambers. Etruscan burial practices can inform us of their perspective on death and the afterlife. I accomplish this by analyzing archaeological artifacts, such as vases and mirrors, as well as using scholarly papers on the subject. In conclusion, this paper argues that wealthier people often put unused vases for the dead to use in the afterlife and as symbols of status. Specifically, I do analysis to argue both men and women used mirrors.

Etruscan Burial Chambers

The Etruscan burial chambers tell us about their way of life. In some ways, they mirrored the houses which they lived in during their lives (Sheldon 1). Some of the tombs had roofs which were similar to different types of roofs used in housing, whether for huts or houses (Sheldon 1). Many tombs had multiple different rooms which were separated by door frames. The tombs often had dining rooms or bedrooms surrounding the primary burial chamber. Some tombs had specific architectural features, such as arches, columns or porches (Sheldon 1).

The burial chambers were underground as they never buried their dead on the surface. The tombs were constructed from bedrock or tufa blocks. The Etruscans were excellent with stone, and had the tools, such as pickaxes, to shape them however they desired. The burial tombs of the Etruscans were typically far away from their population centers, as they wanted to separate the living from the dead. This holds true today as there is a stigma around houses near graveyards (Sheldon 1).

For example, the Tomb of the Reliefs can inform our understanding of Etruscan tombs. It is clear the tomb is for the Matunas family because of an inscription that mentions them by name. The tomb is of a quadrangular shape and is supported by two piers. In the room, there are numerous niches that resemble dining couches called *klinai*. Armor was kept on pegs throughout the tomb (Becker 1). On one of the walls of the tomb was an allusion to the afterlife. The hellhound Cerberus and an anguiped demon had been inscribed. This demonstrates the Greek influence on the Etruscans. Cerberus is the 3 headed dog that guards the underworld in Greek mythology (Cerberus 1).

Based on the grandeur of the burial chambers, as well as the effort put into preparing them, it is clear that the Etruscans strongly believed in the afterlife. There is no way of knowing exactly why the Etruscans put so many items in the chambers. A possible explanation for this is they thought the deceased could use these items in the afterlife. They thought the departed would need to be buried with their items for them to use it (Taylor 1). The depictions of Greek characters of death, such as Cerberus, may serve to protect the deceased in the afterlife and ensure they got to a better place. In order for the deceased to thrive in the afterlife, their body had to be buried with much care (Honeycutt 3). Since they were so heavily influenced by the Greeks, they likely held similar beliefs in the afterlife. The Greeks believed that there were three different afterlives. The Fields of Punishment was the resting place of those who did bad things during their life. The Fields of Asphodel was a place for people who didn't really have any effect on the world. Elysium was a place for those who did great things during their lives (The Getty Museum 1). The Etruscans could have thought that by giving their loved ones a great burial, they will have a better chance at a better afterlife.

The Etruscan burial chambers were expensive to create. They were only for the ultra-wealthy. So, how was the rest of Etruscan society buried? The average Etruscan could not afford to build a massive burial chamber for their loved ones or for themselves. Many Etruscans believed in an afterlife and wanted to give their deceased family members a proper burial. Otherwise, they may be haunted by their disappointed former acquaintances (McLean 4). A lot of Etruscans cremated their dead. This was cheaper and they could store them in an urn. However, cremations were still relatively expensive because the family had to buy enough wood and oil to burn the body. To ensure their loved ones would be pleased with their funeral, they often placed religious objects and images around them. Images of Cerberus or other underworld figures were common (Kozlowski 5).

Vases

Second Paragraph

Vases were a critical part of Etruscan burial practices. They tended to place numerous vases in their chambers. There were many different types of vases. The amphora, hydria, krater, kylix, kantharos, and oinochoe were especially notable (Mignani 1).

The amphora was one of the most used vases of them all. Its main purpose is to store and aid in the transportation of wine, oil, honey, and corn seed. It had many different shapes which dictated what its main purpose was. Generally, it had a narrow neck, a wide belly, and a small base. They were typically cylindrical or globular in shape (The Editors of Encyclopaedia 1).

An example of a famous amphora is Exekias Vatican Amphora (Fig.1). It is 61 cm tall, which is around 24 inches, or two feet. It has two main functions: storage and display. It was used occasionally to store resources, but because of its intricate design its functional use was limited. The top and bottom of the amphora was painted black, but the body contained the scene. There is quite a lot going on in the separate scenes. Ajax and Achilles play a boardgame, and it appears Achilles is winning. Dioskouroi, otherwise known as the brothers Castor and Pollux, are

going back to their parents Leda and Tyndareos. A slave is shown giving Castor some clothes while his brother plays with the family dog. The amphora has golden colored handles and a similar colored rim. The artist and creator, Exekias, used a black figure style. This made the figures in the scenes extremely visible and allowed the amphora to last for so many years (Attic Black 3).

Hydriae, similar to amphoras, were common vases in Etruscan and Greek society. They get their name from the Greek word for water and very much stay true to their name because they are used for fetching water. In wealthy families, enslaved peoples typically filled and carried the hydriae back to their masters' houses. For everyone else, however, women were usually the ones to collect water as it fell into their duties as a homemaker. Hydriae were usually made from bronze or terracotta. Occasionally, the handles of the hydriae were covered with a silver inlay. They typically had a narrow top and bottom and a very wide body, with handles on each side (Hemingway 2).

The Meidias hydria is one of the most famous vases ever created (Fig. 2). It was built around 420 BC by an Athenian potter named Meidias. It wasn't meant to be merely a household object. It was filled with images of mythological figures and contains much symbolism. There was nothing special about its shape. The only way you would be able to differentiate it from a normal hydria was by looking at all the images. The front contained scenes of the kidnapping of the Leukipides. The back had numerous images of heroes and kings and the side showed Kastor's chariot and the flight of Persuasion (Villing et. al. 2).

Kraters were used to dilute wine with water. They were massive, so they were likely to be kept in one room, typically the dining room. They were made of pottery or metal. It was common for expensive kraters to be given to temples or other religious places. They have a broad body and a wide mouth, unlike the hydria or amphora. They could either have horizontal handles located around the base, or vertical handles located around its shoulders. There are many different types of kraters, including the bell krater, volume krater and calyx krater, and the column krater. The bell krater, true to its name, was shaped like an inverted bell. The volume krater bore a resemblance to a very large egg. The calyx krater looks like the calyx of a flower. Finally, the column krater has handles rising in a column-like way over its body (The Editors of Encyclopaedia 5).

The artist, Euphronios, created the Euphronios krater (Fig.3). This vase was taken from an Etruscan tomb and illegally sold to a museum. It was in remarkably good condition for its age, as it likely wasn't used that much. The krater tells the story of a scene from the Trojan War. A son of Zeus named Sarpedon had died on the battlefield. Zeus sent two deities to make sure his son's body was not neglected. These deities, very fittingly, were Sleep and Death. Hermes, one of the twelve olympians, guides Sarpedon's soul in the image. This vase tells us that Greeks preferred to die young when they were still beautiful, rather than to succumb to old age. It emphasized the musculature of youth. It was made about 2,500 years ago, yet experts believed it would have looked the same back then as it does now. It was bought by an Etruscan and buried

in a tomb. Grave robbers targeted Etruscan tombs because they knew Etruscans gave elaborate burials (Thompson et. al. 1).

A kylix is a commonly used cup for drinking. Unlike the other vases mentioned above, it was not used to store anything and was usually small. It has horizontal handles and has a similar shape to a dog bowl. The base was stable enough for it to stand on its own. The exterior of the kylix was covered in all sorts of designs (The Editors of Encyclopaedia 2). Archaeologists found many kylix's broken in parts. This led them to believe that they were thrown at walls in certain games. The disc-like shape of a kylix lends itself to being thrown.

The Kachrylion kylix is one of the most famous Greek vases (Fig.4). It is special because it has a rare coral red-slip on the inside, adding to its intricate design. The inside of the kylix has an image of Eros with a lotus tendril in his hand. Eros is the god of love. People think that Eros is flying over the sea, referencing his mother Aphrodite as she was born from the foam of the sea. Others believe it is a reference to the sky and the earth and the difference between the divine and mortal realms. Similar to the interior of the kylix, the exterior paints a story. It has one of the earliest depictions of Theseus's journey from Troizen to Athens. He confronts the minotaur, defeats Sinis, wrestles Kerkyon, and captures the Bull of Marathon. The kylix displays the motif of unity through all its images (The Kachrylion Kylix 1).

The kantharos was also a type of drinking vessel. It is very similar to a goblet and has large handles. It has a wide top and is quite deep, but is very light, making it convenient to drink from. They would be most commonly used at dinner tables, and were logistically a much better cup to drink from than a kylix. They are quite similar in shape to many modern day drinking vessels (The Editors of Encyclopaedia).

Some of the Etruscan kantharos were single handed, rather than double like one of the vases in the MET (Fig.5). Therefore, people could use one hand to drink, rather than two. This particular vase was made between 550-525 BCE during the Archaic period. It isn't a large vase, as it is 21.8 cm tall including the handle and 13.3 cm excluding it. This makes it conducive for everyday use. It was made from a type of terracotta called *bucchero pesante*. This piece has a relatively small base and a very wide top, which gives the user a lot of space to drink. Its neck is extremely narrow, as it does not serve a functional purpose. The handle of the drinking vessel is decorated with a female head, as well as other basic patterns. It is a darker color - blackish silver- which accumulated a lot of rust overtime (Etruscan 2).

Oinochoe were drinking vessels for wine. Oinochoe could be broken down into oinos and cheo. In Greek, these words mean wine and I pour, respectfully. It has one handle, which makes it different from all the other vases listed here. It isn't as wide as it is tall. There are many different types of this vase. The differences came from the handles, the mouths and the bodies of the vase. Smaller versions of these vases are found in children's graves, indicating that there really was no drinking age during the age of the Etruscans (Mannack 3).

One terracotta oinochoe has 3 main figures on it (Fig.6). Pompe, Eros and Dionysus. Dionysus makes sense as he is the Greek god of Wine and this is a vessel mainly used to drink wine. Pompe and Eros were finishing up a ritual meant to honor Dionysus. It may have been a

celebration of marriage. Pompe is holding a wreath and she looks in the direction of Dionysus. Eros is the winged figure in the vase, and he is visually smaller than the two other figures. He is about to move, as he adjusts his slippers. The bottom of the scene has a relatively solid golden background, while the top frame is separated by a vine and leaf pattern (Terracotta 2).

Many rich Etruscans put bucchero pottery in their tombs. This pottery has a distinct black color because of the way it is made. There is very little oxygen in the firing chamber, giving the vases a darker color. These vases were extremely expensive to make and were thus only used by the ultra-wealthy. The reason they were put in tombs is that they were also used in dining settings. The rich Etruscans wanted their family members to use the same pottery they used when they were alive. Furthermore, by burying such expensive vases, they are able to reinforce their superior social standing to the rest of society (Becker 4).

The pottery typically depicted scenes of Greek mythological events. For example, one piece included Odysseus's encounter with the cyclops Polyphemus (Fig. 7). The pottery contained patterns on the top and the bottom. These patterns typically weren't complex, like olive branches or winding swirls. Some vases were made for the sole purpose of burial. We could tell this because the handles of the vases would not have had any wear and tear. On the other hand, some vases were heavily used and were therefore likely put in the burial chambers because the deceased owned it. The Etruscans liked putting pottery in their burial chambers because they thought the dead could use them in the afterlife. They believed that the dead still had to eat or drink, just like their living comrades. How would they be able to drink or eat if they didn't have access to these vases in their burial chambers?

Third Paragraph

Mirrors

Mirrors give us significant insight into the world of the Etruscans. The Etruscans made their mirrors out of bronze or silver. Approximately 3,000 mirrors have survived the test of time. The handles of the mirrors were made from ivory, bone and wood. Many of these mirrors have a concave section that contains the engravings. At some point in the 4th Century BC, the Etruscans started making the mirrors and handles in one piece, rather than two. This demonstrates a significant scientific innovation that took place (Puma et. al. 1067). The mirrors were light and meant to be held in one hand. The sheer number of mirrors found in Etruscan burial sites reinforced the negative stereotype perpetuated by their conquerors, the Romans, that they were narcissists (Cartwright 2).

There are three main groups of Etruscan mirrors. They were differentiated based on their design, workmanship and their features. The three types are Kranzspiegel, Dioskouroi, and Lasa mirrors. Kranzspiegel mirrors tend to have an intricate frame on the side of the mirror that doesn't reflect. People can tell it is this type of mirror based on the inscribed labels. The other two types are not of as great quality as the first. They are smaller and more prone to breaking. They often don't have a frame and it is difficult to tell what is going on in their engravings (Puma et. al. 1041).

Since many mirrors were buried with women, many scholars assumed they were solely used by them. However, a lot of adult women were not buried with mirrors. This suggests mirrors were not completely linked with women and not required for a female burial (Bonfante 1). Furthermore, the engravings on the mirrors should appeal to women, yet some didn't. They often contained scenes of gory violence and battle. This doesn't make sense, because a feminine object has such masculine decorations. This might mean that Etruscan men, as well as women, used the mirrors. The excavations were very uncoordinated and messy, which may have led to the destruction of mirrors in male burial chambers. Furthermore, the excavators likely had their own preconceived notions about the mirrors and this may have influenced their findings.

Etruscan mirrors contained scenes of marriage, union, birth, prophecy and union. They were sometimes given to brides on their wedding day (Bonfante 1). Based on their inscriptions, they could also symbolize the union of two families (Cartwright 1). Many mirrors contained heroes, like Achilles, which may have been an inspiration to men fighting in battle (Puma et. al. 1041).

For example, one of these mirrors tells us about Atlas and Heracles (Fig. 8). The figure on the right is Atlas and on the left is Heracles. The perimeter of the mirror is covered with a long vine. Atlas appears to be holding a massive boulder, while Heracles holds a club and wears a Nemean cape. There is a spear to the right of Atlas, which is interesting because he wouldn't be able to grab it. Since he symbolizes resilience, the spear may mean he is ready to fight the gods again. Next to the spear is a plant that is around half the height of the Titan. Both men are in excellent physical condition, although Heracles is slightly more fit. Heracles appears younger than the Titan, with a cap on his head. Atlas looks much older as he has a beard. This scene depicts the mythological Heracles and the golden apples of Hesperides. In order to get the apples, Heracles tricks Atlas into holding up the sky. Atlas is known as the Titan who was the main general of Kronos during the Titan-God war. The Gods won and punished Atlas by forcing him to hold up the sky eternally. He can't easily escape his punishment because no one else would want to face such a harsh punishment. In the story with Heracles, he appears in a much nicer light. He and Heracles help each other out. Heracles kills the Hesperion Drakon which agonized the Titan while Atlas helped him get the apples. Furthermore, Heracles helped Atlas with the burden of carrying the heavens by building two massive pillars (Atlas 1).

Even though some mirrors may appeal more to men, others definitely had a greater appeal for women. For example, a mirror depicting the birth or care of a child would interest women more than men. In one particular mirror (Fig 9), Hera feeds Heracles when he was a baby. She is surrounded by other women as she takes care of the baby. The three women all wear long dresses. Hera is the one holding the baby. Next to Hera is a tall plant, similar to the one next to Atlas. The perimeter of the mirror is covered with integral shaped figures, as well as plant patterns. The interesting part about this is Heracles is not Hera's son. He is the product of Zeus's cheating (Britanica 2). Therefore, it would make sense if Hera fostered some resentment toward the boy, even if it wasn't his fault. This depiction could be telling us a lot about the values and expectations of women in Etruscan society. They may have been expected to take care of their husband's children, regardless if they were their own.

In contrast, a mirror with an engraving of Hercules defeating the Hydra would not appeal as much to women, even though it contained a female figure (Fig.10). In it, Menrva, the Etruscan equivalent of Minerva or Athena to the Greeks, leads Hercules to slaughter the hydra. Hercules is depicted holding his club and wearing the lion skin, as per usual. Menrva is shown wearing wings, which may be an illusion to her heritage as her father, Tinia, was the Etruscan god of the sky. The hydra looks like 3 connected snakes. The perimeter of the mirror is covered with a vine and leaves. The base of it has two separate geometric patterns. Cutting off the heads of creatures is not something women were expected to experience at this time. Therefore, it would make a lot more sense that this was a man's mirror, rather than a woman's.

Some mirrors have inscriptions on them that mean, for the tomb. This means that the mirrors weren't meant to be used by anyone living and therefore were in better condition. This term is *śuthina*. It is not only on mirrors, but also on vases, jewelry and weapons. In some funerary rituals, the mirrors were intentionally damaged. It is difficult to say why the Etruscans did that. The placement of the mirrors in the tomb was also a significant factor in burial. Because so many tombs were excavated carelessly, archaeologists lost out on crucial information in this regard. Regardless, the mirrors were placed in a somewhat consistent pattern, which may point to another funeral ritual. If a mirror was ever placed in a burial chamber, the reflecting side would always be visible. The Lasa was the most common mirror in these gravesites (Puma et. al. 1041).

Many Etruscans believed that mirrors represent the boundary between the living world and the world of the dead. Mirrors served as a type of telephone that allowed the living to connect with the dead, making it a necessity for burial sites. Mirrors still have a deep association with funerals. For example, some Jewish families cover all of the mirrors in their house when someone dies. Etruscan mirrors often celebrate the lives of the dead. It emphasizes the happiest moments of the deceased lives, such as their marriage, children and accomplishments. They serve as reminders to the dead of what happened in their lifetime (Bonfante 1).

Conclusion

Previously, scholars thought the Etruscans put vases and mirrors in their burial chambers as a sign of wealth, and perhaps this is true to some extent. Based on the archaeological evidence, this paper demonstrates that the Etruscans thoroughly filled their burial chambers with goods primarily for religious reasons. Figures of death, such as Cerberus, the three-headed dog of the underworld were common. Figures associated with death were necessary for a proper burial because they were considered beneficial to the afterlife of the deceased. A proper burial could be the difference between ending up in Elysium or the Fields of Asphodel. Furthermore, the Etruscans believed the deceased could use the objects they were buried with in the afterlife. Therefore, the Etruscans buried the deceased favorite belongings so they could have them in their eternal afterlife. The many dinner cups found in the chambers as well as mirrors indicate the desire for favorable belongings to be buried with the deceased.

Specifically, a lot of the mirrors found in the chambers illustrate critical moments of a person's life, like marriage, the birth of children, or another significant feat. The mirrors may

have been buried with the deceased to remind them of important moments during their own lifetime. Scholars initially believed only women were buried with mirrors, but it is likely this was not the case. I argue that both men and women were buried with mirrors. Some of the depictions on the mirrors, like the one with Hercules slaying the Hydra or Atlas holding up the heavens, were likely not associated with women. They were, on the other hand, associated with men as they contained gory scenes of violence.

While in the past scholarship has applied modern gender norms to artifacts, this research demonstrates that these assumptions are not necessarily valid. Hopefully, this paper inspires future reinterpretations of the objects in tombs and the funerary practices of the Etruscans.



Figure 1: The Exekias gaming amphora (Mangieri, A. F., 2024, Fig. 14.)



Figure 2: The Meidias Hydria (Alexandra Villing and Amy C.Smith, 24 November 2023, <https://www.britishmuseum.org/blog/historys-most-famous-pot-meidias-hydria>)



Figure 3: The Euphronios Krater (Dr.Erin Thompson, Dr.Steven Zucker, and Dr.Senta German, July 13, 2017, <https://smarthistory.org/euphronios-krater/>)

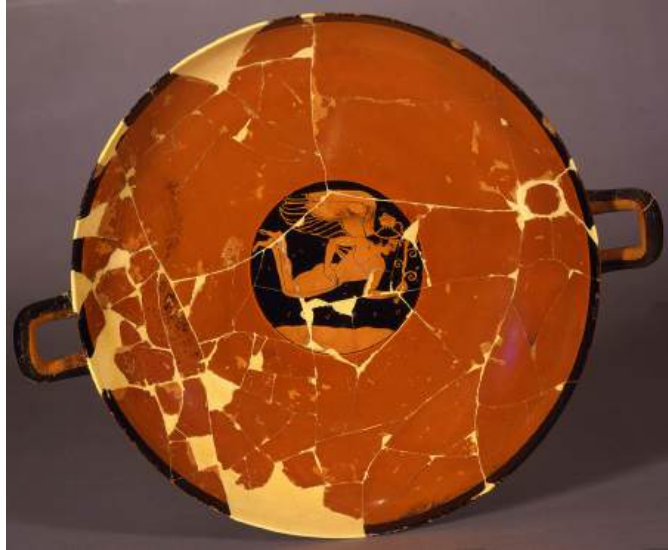


Figure 4: Kachrylion Kylix (Museo Archeologico Nazionale, Florence (inv. 91456), <https://www.metmuseum.org/about-the-met/collection-areas/greek-and-roman-art/kachrylion-kylix-ancient-greek-vasse>)



Figure 5: Terracotta Stemmed Kyathos(Richter Gisela M. A. 1940. P11, fig.35, New York)



Figure 6: Terracotta oinochoe(Hemingway, Sean. 2021. Pp.34, fig 20, New York)

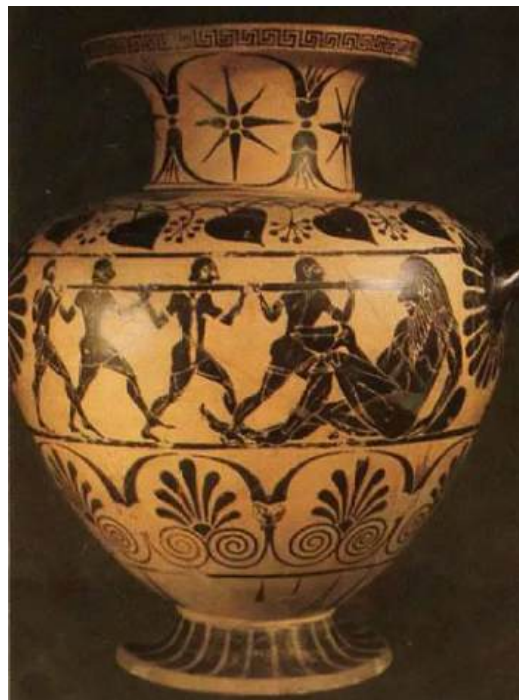


Figure 7: Etruscan Vase(“All in the Head: The Cyclops is Part of Odysseus.” July 29, 2018.)



Figure 8: Hercules and Atlas Mirror(Copeland, Mel. 24 July, 2009)



Figure 9: Mirror of Hera Feeding Heracles(De Puma, Richard Daniel, 2013, p.1047, Fig.58.7)



Figure 10: Mirror showing Hercules and the Hydra(27 Aug.2019,)

Works Cited

- “An Engraving on an Etruscan Mirror Depicting Menrva Leading Hercules to His Second Labour the Destruction of the Many Headed Hydra Stock Photo.” Alamy, 27 Aug. 2019, www.alamy.com/an-engraving-on-an-etruscan-mirror-depicting-menrva-leading-hercules-to-his-second-labour-the-destruction-of-the-many-headed-hydra-image268809888.html.
- “Atlas.” ATLAS - Greek Titan God, Bearer of the Heavens, www.theoi.com/Titan/TitanAtlas.html. Accessed 17 Aug. 2024.
- “Attic Black-Figure.” Exekias, www.carc.ox.ac.uk/carc/resources/Introduction-to-Greek-Pottery/Keypieces/blackfigure/exekias. Accessed 16 Aug. 2024.
- Becker, Jeffrey A. “Bucchero.” Smarthistory Bucchero Comments, 8 Aug. 2015, smarthistory.org/bucchero/.
- Becker, Jeffrey A. “Tomb of the Reliefs.” Smarthistory Tomb of the Reliefs Comments, 8 Aug. 2015, smarthistory.org/tomb-of-the-reliefs/.
- Bonfante, Larissa. “Etruscan Mirrors and the Grave”. *L’écriture Et l’espace De La Mort. Épigraphe Et nécropoles à l’époque préromaine*, edited by Marie-Laurence Haack, Publications de l’École française de Rome, 2015, <https://doi.org/10.4000/books.efr.2741>.
- Britannica, The Editors of Encyclopaedia. "Alcmene". Encyclopedia Britannica, 11 Jul. 2019, <https://www.britannica.com/topic/Alcmene>. Accessed 14 August 2024.
- Britannica, The Editors of Encyclopaedia. "amphora". Encyclopedia Britannica, 1 Aug. 2018, <https://www.britannica.com/art/amphora-pottery>. Accessed 17 August 2024.
- Britannica, The Editors of Encyclopaedia. "Cerberus". Encyclopedia Britannica, 21 Jun. 2024, <https://www.britannica.com/topic/Cerberus>. Accessed 17 August 2024.
- Britannica, The Editors of Encyclopaedia. "kantharos". Encyclopedia Britannica, 25 May. 2018, <https://www.britannica.com/art/kantharos>. Accessed 12 August 2024.
- Britannica, The Editors of Encyclopaedia. "krater". Encyclopedia Britannica, 4 Jan. 2022, <https://www.britannica.com/topic/krater>. Accessed 12 August 2024.
- Britannica, The Editors of Encyclopaedia. "kylix". Encyclopedia Britannica, 25 May. 2018, <https://www.britannica.com/art/kylix>. Accessed 17 August 2024.
- Cartwright, Mark. “Etruscan Bronze Mirrors.” World History Encyclopedia, <https://www.worldhistory.org/#organization>, 24 Jan. 2017, www.worldhistory.org/article/1009/etruscan-bronze-mirrors/.
- Copeland, Mel. “Etruscan Phrases.” Index.Html, Maravot.Com, 24 July 2009, www.maravot.com/translation_shortscripts_d.html.
- De Puma, Richard Daniel. “Mirrors In Art and Society.” *The Etruscan World*, Routledge, London, 2013, pp. 1041–1067.
- “Etruscan Tombs.” Etruscan Corner, 10 Nov. 2023, etruscancorner.com/en/etruscan-tombs-and-necropolis/etruscan-tombs/.
- Hemingway, Sean, and Colette Hemingway. “Greek Hydria (Water Jars) and Their Artistic Decoration: Essay: The Metropolitan Museum of Art: Heilbrunn Timeline of Art History.” *The Met’s Heilbrunn Timeline of Art History*, July 2007, www.metmuseum.org/toah/hd/gkhy/hd_gkhy.htm.

- Hemingway, Sean. *How to Read Greek Sculpture*, The Metropolitan Museum of Art, New York, 2021, p. 34.
- Honeycutt, Jessica. "Funerary Practices throughout Civilizations." *Ancient Art*, 24 Apr. 2015, ancientart.as.ua.edu/funerary-practices-throughout-civilizations/.
- Kozlowski, Jaime. "Obsession with Death." *The Etruscans: Obsession with Death*, www.domspe.org/etruscans/obsession.html. Accessed 17 Aug. 2024.
- Mangieri, A. F. (2024). Heroics of Dress: Exekias and Ornament in Greek Vase Painting. *American Journal of Archaeology*, 128(1), 59-88.
- Mannack, Thomas. "Introduction to Greek Pottery." *Pouring Vessels*, www.carc.ox.ac.uk/carc/resources/Introduction-to-Greek-Pottery/Shapes/Pouring-vessels. Accessed 17 Aug. 2024.
- McLean, John. "Western Civilization." Lumen, courses.lumenlearning.com/atd-herkimer-westerncivilization/chapter/etruscan-religion/#:~:text=The%20Afterlife,-Etruscan%20beliefs%20concerning&text=The%20Etruscans%20shared%20in%20general,treatment%20of%20the%20deceased%E2%80%99s%20remains. Accessed 17 Aug. 2024.
- Mignani, Elena. "Etruscan Vases Shapes and Uses." *Etruscan Corner*, 10 Nov. 2023, etruscancorner.com/en/infographics/etruscan-vases-shapes-and-uses/.
- Richter, Gisela M. A. *Handbook of the Etruscan Collection*, by Gisela M.A. Richter. The Marchbanks Press, 1940.
- Sententiaeantiquae. "All in the Head: The Cyclops Is Part of Odysseus." *SENTENTIAE ANTIQUAE*, 29 July 2018, sententiaeantiquae.com/2018/07/29/all-in-the-head-the-cyclops-is-part-of-odysseus/.
- Sheldon, Natasha. "Etruscan Tombs." *History and Archaeology Online*, 8 Jan. 2022, historyandarchaeologyonline.com/etruscan-tombs/.
- Taylor, Laurel. "The Etruscans, an Introduction (Article)." Khan Academy, Khan Academy, www.khanacademy.org/humanities/ap-art-history/ancient-mediterranean-ap/ap-ancient-etruria/a/the-etruscans-an-introduction#:~:text=Fortunately%2C%20though%2C%20the%20Etruscans%20cared,use%20in%20the%20next%20world. Accessed 17 Aug. 2024.
- "Terracotta Oinochoe (Jug): Greek, Attic: Classical." The Metropolitan Museum of Art, www.metmuseum.org/art/collection/search/251935. Accessed 12 Aug. 2024.
- "Terracotta Stemmed Kyathos (Single-Handled Cup): Etruscan: Archaic." The Metropolitan Museum of Art, www.metmuseum.org/art/collection/search/246224. Accessed 13 Aug. 2024.
- "The Getty Museum." Getty Museum, www.getty.edu/art/exhibitions/ancient_underworld/inner.html. Accessed 12 Aug. 2024.
- "The Kachrylion Kylix: A Major Loan from the National Archaeological Museum, Florence." The Metropolitan Museum of Art, 29 July 2022, www.metmuseum.org/about-the-met/collection-areas/greek-and-roman-art/kachrylion-kylix-ancient-greek-vase.
- Thompson, Erin, et al. "Euphronios, Sarpedon Krater." *Smarthistory Euphronios Sarpedon Krater Comments*, 13 July 2017, smarthistory.org/euphronios-krater/.
- Villing, Alexandra, and Amy C Smith. "History's Most Famous Pot: The Meidias Hydria | British Museum." *History's Most Famous Pot: The Meidias Hydria*, 24 Nov. 2023, www.britishmuseum.org/blog/historys-most-famous-pot-meidias-hydria.

Scream, Typecasting, and African Americans: A Content Analysis of Typecasting in the Scream film Series By Hayden Yoest

Introduction

My research aims to answer the question “How extensively has the practice of typecasting African Americans been employed in the *Scream* film series?” Movies are cultural products, which are often reflective of current ideologies. With specific respect to Hollywood (which refers to the mainstream American film production companies), American culture is dominated by its influential ideologies. Hollywood’s ideologies, however, largely promote rhetoric that reinforces negative beliefs and stereotypes about racial minorities. Typecasting refers to the subjection of actors to stereotypical roles based on qualities such as race or gender; for the purpose of this study, typecasting will be analyzed with just a racial lens. Typecasting is a widespread practice in Hollywood and has grown in popularity among major film studios largely because of economic reasons. Yet the racial bias, that many experts believe to exist in Hollywood, largely allows for white actors to continue playing diverse roles while actors of color are limited to stereotypes.

Media is important to study, as media influences both how African Americans shape their identities and how others (audiences) perceive them. In order to explore how Hollywood portrays African Americans, I will perform a content analysis of the current Scream film series. In order to fully analyze the series, I will be using a reading of racial representation and theories to identify typecasting.

I aim to examine the extent to which the common-place practice of typecasting is employed in the *Scream* film series, and if so how the degree of typecasting varies from each movie. I hypothesize that typecasting will be used to a moderate degree in the first few films before gradually decreasing in prevalence with each subsequent film released. I additionally hypothesize that the decrease will occur around 2015, which would coincide with the rise of #OscarsSoWhite (specified later).

Literature Review

In 2019 Hollywood surpassed its previous box office record with an estimated revenue of \$42.5 billion. The year additionally saw 814 Hollywood movies begin production, an increase from previous years (Malik et al., 2022). Though Hollywood has continued to grow, it has unproductively used its unmatched influence by continuing to underrepresent African Americans. According to UCLA’s 2020 Hollywood diversity report, white and minority population shares in the US are changing. It is projected that by 2050, minorities will make-up 53 percent of the total population, with whites at 47 percent (Wolf, 2020). This is a stark shift from the current US population make-up of roughly 60 percent whites and 40 percent minorities. The expected demographic shift means that industry change will become increasingly important from here on out. The report also found evidence that suggests “America’s increasingly diverse audiences prefer diverse television content” (Wolf, 2020). Yet data from a recent report by

Wendy Lee, a business reporter and graduate from University of California, Berkeley, suggests that Hollywood is backsliding on diversity. The report found that white people represented 78% of lead roles in top movies in 2022, an increase from 72.4% in 2019 (Lee, 2023). In fact, only 12.2% of all speaking roles in Hollywood have gone to African Americans (Suhling, 2017). This number shows a blatant discrepancy when taking into account that African Americans make-up 13.6 percent of the US population (United States Census Bureau, 2022). This discrepancy is palpable evidence for the underrepresentation of African Americans on-screen. Moreover, these studies in unison advance the notion that America is becoming more diverse while Hollywood is becoming less.

Diversity is important in Hollywood, as Hollywood has perfected its ability to generate images in peoples minds, specifically about race. A study by Rawan Elbaba (2019), suggests that minority teens are particularly susceptible to Hollywood's influence. Through 144 interviews with teens about on screen representation, Elbaba found that for many teens, minority portrayal in Hollywood media affects not only how others see them, but also affects how they see themselves. When asked about on screen representation, common responses were “not seeing yourself represented in elements of pop culture can affect mental health” and that “[the teens] said they often look to trends in pop culture when forming their own identities” (Elbaba, 2019). Many of the teens noted their criticism of typecasting, with one stating, “Hollywood is overcompensating for their lack of diversity by depicting exaggerated and stereotypical characters” (Elbaba, 2019). Through these interviews, it is clear that typecasting is both apparent and unappetizing to young audiences. This is corroborated by Wolf’s previously mentioned statement suggesting that Audiences prefer diversity. It is additionally evident that typecasting has negative impacts upon audiences, generating false perceptions of different races. It is of concern that these false perceptions can manifest within not only audiences, but also within an individual - potentially causing prejudice about oneself. This impact can in turn lead to degraded mental health (as described by Elbaba) and false perceptions becoming reality. This impact gives additional importance to the study of representation in Hollywood.

Though it is unclear as to why exactly typecasting is employed by Hollywood when it is harmful and unpopular among audiences, a report by Chantal Suhling provides insight on why typecasting is popular among film studios. Suhling (2017) found that because 94% of Hollywood executives are white, many perspectives are missing from the decision making process. She additionally found that among executives, a commonly stated rationale for the racial underrepresentation in Hollywood is a “lack of economic indicators” (Suhling, 2017). Simply put, financial indicators (used to determine a film's profitability) are used in regard to African Americans in lead roles. These “indicators” are ultimately used to promote rhetoric suggesting that films with African American leads will not perform well at the box office. Therefore, to keep profits high, studios are more likely to go with well established actors who are overwhelmingly white. This problem is exacerbated by the fact that the few films with African American leads receive significantly less in production budget on average (Malik et al., 2022). This greatly limits the potential for films with African American leads, as production budget

often correlates with a film's box office success. Because of this, a cycle of films with black leads financially underperforming is created.

Suhlings findings additionally align with Stuart Hall's theory on *Representation and Discourse*, which focuses on the broader context of the media. Hall theorizes that films are media, which is used to construct culture. The end result of this is that media is reflective of the ideologies of those in charge of the media (Hall, 1997). The diagram below offers a visualization of key relationships according to Hall's theory. These areas are either directly or indirectly controlled by Hollywood executives:

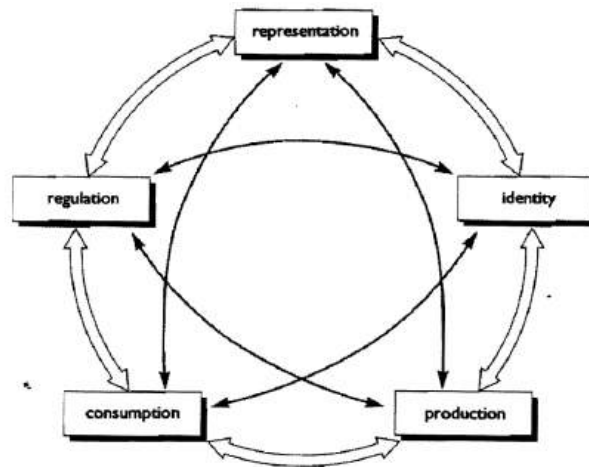


Fig. 1. "Illustration of Stuart Hall's Theory." American Psychological Association, 1997 (p. 1).

The Annenberg school at the University of Southern California, a leader in the research of Hollywood inclusion, further confirms Suhling's claim that executives play an important role in the inclusion and representation of African Americans. Annenberg's *Inclusion List* analyzed the racial makeup of the off-screen workers in Hollywood's 100 top grossing movies between 2019 and 2022. The study found that an overwhelming 86% of casting directors were white, which leaves the remaining 14% of casting directors to be underrepresented people of color. The *Inclusion List* also collected data on its top 100 most inclusive Hollywood films in the same timeframe. The data showed that this time only 70% of casting directors were white, with the remaining 30% of casting directors to be racially underrepresented men and women (Smith, 2023). This data suggests that there exists a correlation between the race of the casting director and the degree of racial representation in the movie.

Though underrepresentation, economic indicators, and typecasting are still commonplace in Hollywood, executives may soon be changing policies in an attempt to save-face with audiences. The Oscars have historically showcased just how much Hollywood underrepresents minorities. The OscarsSoWhite hashtag has dominated social media over recent years. OscarsSoWhite is commonly thought to have first taken hold around 2015, and is a public backlash against underrepresentation in Hollywood. According to Suhling, OscarSoWhite has already had some effect on hiring policies being implemented to be more inclusive of all races

(Suhling, 2017). This suggests that the public backlash has been, to an extent, effective at making improvements towards racial inclusion.

Gap

Research into Hollywood's cinematic representation of African Americans remains limited in a number of ways. Firstly, existing research pertaining to this area is almost entirely focused on quantitative comparisons of data about cast and executive demographics. Thus, the preexisting research fails to account for the qualitative and dynamic portrayal of African Americans in film. Secondly, there is a lack of quantitative data relating to the typecasting of African Americans (which itself remains largely unresearched). Further, no studies using a racial lens have been conducted on the *Scream* series. This study fills all of these gaps in the surrounding literature by employing a mixed methods approach on the *Scream* films.

Methods

There are four main steps to my method: 1) select the films I will analyze, 2) sort each character into a racial category, 3) implement the method to gain qualitative observations, and 4) quantify data to gain a new understanding.

I will be conducting an analysis of all current *Scream* films (1996-2023) using a lens focused on identifying the presence of the common Hollywood practice of Typecasting, with respect to African Americans. A mixed-methods study will be conducted: this approach will allow for both a quantitative and qualitative analysis of African American typecasting in the *Scream* series. My research design will be non-experimental correlational research, as I will be exploring the possible relationship between African American characters in *Scream* and Typecasting.

I selected the *Scream* films because the franchise is an icon of Hollywood's thriller genre, which palpably makes use of character tropes. My research design will be able to explore if *Scream's* use of tropes falls into Typecasting. I will not be looking at other people's opinions or reviews of the movies.

My design for data collection will be inspired by a method created by Musa Malik and colleagues, who created a clearly defined, standardized, and scalable metric for measuring cinematic representation among racial minorities (Malik et al., 2022). The process of collecting the data - which includes characters, dialogue, and the characters position within the larger plot - will be conducted by watching each *Scream* movie multiple times. Additionally, my analysis will be guided by the inductive open coding process as described by Bloomberg and Volpe (2015). Notes assessing character development through dialogue, key words, phrases, or sentences, will be recorded as I watch the movies. I will assign descriptive labels (codes) to the noted discourse to more specifically categorize the discourse; the codes will be formed by the data itself rather than being predetermined (Bloomberg and Volpe, 2015). Through multiple viewings of each movie, the codes will likely evolve throughout the analysis process. I will then compile the labels

and their corresponding dialogue within a spreadsheet to gain a quantitative understanding of my data.

In preparation of collecting the movie data, I will obtain the actor/actress name corresponding to their on-screen character. To do this I will retrieve each movie's Internet Movie Database identifier (IMDb). I will then research background information on the movie's cast, in order to assign each actor (and subsequently character) into one of the following racial categories: White, Black, Mixed (white-black), and other.

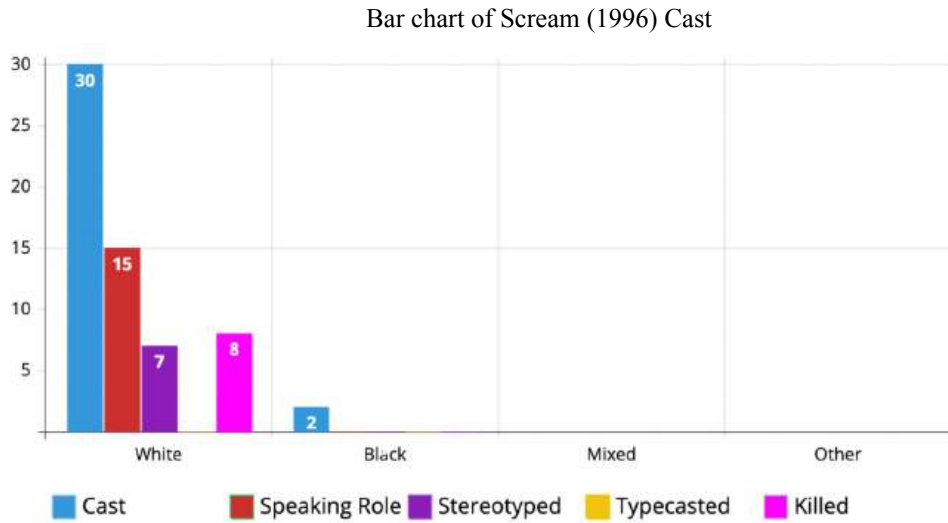
The open coding research process best suits my project because - while very extensive and time consuming - it accounts for all kinds of details as opposed to the priori coding process (using predefined codes), which ultimately leaves some chunks of data to be uncodable. This research design is additionally suitable because the mixed-methods will convert the qualitative data into measurable quantitative data, thus making any trends easily identifiable. An additional advantage of utilizing the open coding research process is that the researcher is forced out of predisposed biases about the topic; this is because the data guides the process of creating the codes rather than the researcher themselves (Bloomberg and Volpe, 2015).

A limitation of my project is that, due to my limited programming abilities, I will be unable to use a character recognition software tool for detection of social characteristics in visual media content. The computer vision software has the ability to quickly process large amounts of film data by identifying predefined characteristics using facial recognition, as opposed to myself personally watching and annotating the films at a much slower pace.

Findings and Analysis

I will be assessing the portrayal of Black characters in relation to prevalent stereotypes in order to inform my new understanding of the nuanced dynamics of racial typecasting in the *Scream* film series. I will assess the portrayal of Black characters with intense scrutiny; I will additionally be examining the portrayal of white, mixed, and other races to create a baseline for comparison to develop a new understanding. It is important to note that if a character is typecasted, then they are also stereotyped.

The bar chart below shows the distribution of roles to each racial category of actors:



The table below is a representation of the relative frequency of the same data above. Each column is totaled to equal 1.00, then divided according to the data from the bar chart to show how prevalent each criteria was within each racial category. The criteria were formed by the codes I created while analyzing my results:

Table 1: <i>Scream</i> (1996) Cast						
Race	Total	Speaking Role	Not Stereotyped or Typecasted	Stereotyped	Typecasted	Killed
White	0.9375	1.0000	0.9200	1.0000	---	1.0000
Black	0.0625	---	0.0800	---	---	---
Mixed	---	---	---	---	---	---
Other	---	---	---	---	---	---

Scream (1996) has a palpable lack of Black cast members. Only two of the 32 credited cast members are black (IMDb, 1990). According to my findings (illustrated in the table and bar chart above) both of the characters portrayed by African Americans lacked speaking roles with more than a few lines, meaning *Scream* (1996) allocated zero speaking roles to African Americans. This revelation is shocking when compared to the previously mentioned 12.2 percent

of speaking roles in Hollywood that go to African Americans (Suhling, 2017). The concern is heightened because the figure of 12.2 percent is, as previously discussed, already unrepresentative of African Americans who make up 13.6 percent of the US population (United States Census Bureau, 2022). The film takes underrepresentation a step further when considering the inclusion of other minority races, as it contains no mixed or other race cast members. Neither black character was stereotyped or typecast. However, it is important to acknowledge a probable nuance to the lack of typecasting, as the lack is most likely due to the fact that their characters were too small to have any personality or character and thus no opportunities for typecasting.

Bar Chart of Scream 2 (1997) Cast

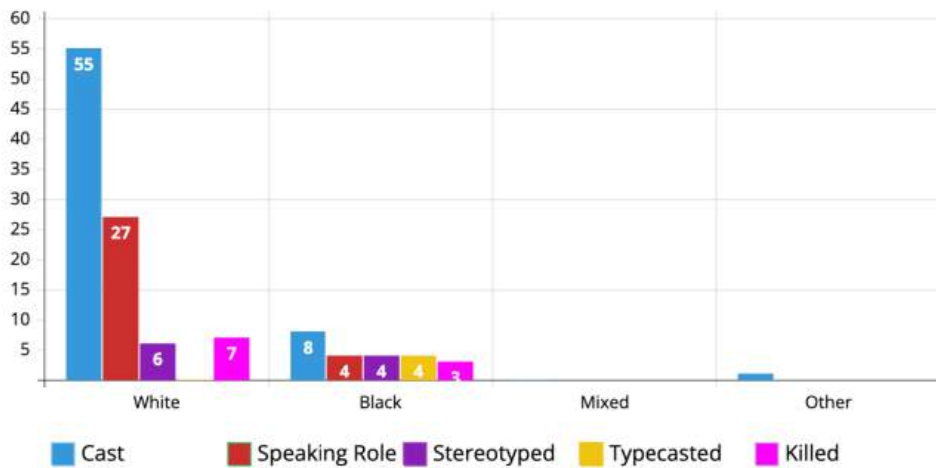


Table 2: Scream 2 (1997) Cast

Race	Total	Speaking Role	Not Stereotyped or Typecasted	Stereotyped	Typecasted	Killed
White	0.8594	0.8710	0.9074	0.6000	---	0.7000
Black	0.1250	0.1290	0.0741	0.4000	1.0000	0.3000
Mixed	---	---	---	---	---	---
Other	0.0156	---	0.0185	---	---	---

The data above shows that *Scream 2* (1997) improves upon the original film's shortcomings in racial representation. Eight of the 64 credited cast members were black, however only four of the eight had speaking roles with more than a few lines. Improvements

were made in the racial representation of African Americans from the prior movie and closely mirrors the 12.2 percent national average of black cast members with speaking roles.

The credited black cast members with speaking roles are Jada Pinket Smith as ‘Maureen’, Omar Epps as ‘Phil’, Elise Neal as ‘Hallie’, and Duane Martin as ‘Joel’. The film opens on Maureen and her boyfriend, Phil, as they stand in line for a movie. While waiting in line, Maureen and Phil engage in conversation with one another. During this conversation, Maureen is overly flamboyant, Phil makes overtly sexual remarks, and the two make excessive use of profane language, more so than any other character. At the start of the conversation, Phil says, “Nobody wants to pay \$7.50 to see some Sandra Bullock shit, unless she's naked” (Craven, 1997, 0:00:37). This first interaction marks Phil as a sexual, profane character. Maureen then says, “[Stab] is a dumbass white movie about some dumbass white girls getting their white asses cut the fuck up” (Craven, 1997, 0:00:54). This tone is carried throughout the rest of the conversation.



Fig. 2. “Maureen and Phil in Line.” *Scream 2*, 1997. Author’s screenshot

As the two near the front of the line, Maureen makes several remarks noting the horror genre's lack of black representation. She states, “All I’m sayin’ is the horror genre is historical for excluding the African American element” (Craven, 1997, 0:01:05). This line was interesting, as it was commenting on the absence of Blacks in many horror films. The line was surprising because so far both *Scream* movies have been particularly guilty of this criticism. However, Maureen then picks up a Ghostface mask and makes a remark about it being white. This paints her as an anti-white radical, essentially invalidating her earlier remarks about the lack of Blacks. These findings emphasize that had this study not employed the open coding research process, this data would most likely not have been accounted for; because this data was a surprise to me, I would likely not have created a predetermined criteria, or codes, that could have encompassed Maureen’s dialogue above (this would have been the case if I opted for the priori coding process, which calls for codes being developed before examining any data [Bloomberg and Volpe, 2015]).

As the couple makes their way towards their seats, they engage in excessive public displays of affection: they make-out and Phil gropes Maureen. Later on, Phil remarks, “I don’t

know about the plot, but I got a stiff one” (Craven, 1997, 0:03:15). Maureen then verbally displays her displeasement with a character in the movie they are watching; she shouts, “Bitch hang up the phone and start 69-ing his ass. Damn” (Craven, 1997, 0:03:58). This action is obnoxious and continues Maureen’s flamboyant portrayal. Phil then reaches for Maureen's thigh and says, “Scary movies are great for foreplay” (Craven, 1997, 0:06:00).



Fig. 3. “Maureen and Phil Kissing During Movie.” Scream 2, 1997. Author’s Screenshot



Fig. 4. “Maureen Shouting During Movie.” Scream 2, 1997. Author’s screenshot

Later on, Phil gets up and goes to the bathroom. He enters a stall and hears moaning in the stall next to his. He is visually pleased, and presses his ear against the stall to listen, doing so in a repulsive manner. The moaning stops after a few seconds, and Phil is stabbed in the head through the other side of the stall. The killer then heads to Maureen, and kills her in her seat. Maureen and Phil exhibited obvious signs of stereotyping. Furthermore, the stereotyping leaned into racial categorization and thus typecasting.



Fig. 5. "Phil Listening to the Next Stall Over." *Scream 2*, 1997. Author's screenshot



Fig. 6. "Phil Stabbed by Ghostface." *Scream 2*, 1997. Author's screenshot

The movie then transitions to a new set of characters, where we are introduced to two new black characters with speaking roles: Hallie and Joel. Hallie's character is presented in a way that makes it clear to the audience that she is the 'white-washed' black girl. She is the only person of color in a sorority composed of mostly rich-white girls. She speaks with an unnaturally soft voice, keeps her hair meticulously straight, and wears chic clothes similar to her sorority counterparts. Hallie is pleasant to everyone she meets, except for Joel, the only other black character she interacts with. During their first interaction, Hallie is visually displeased with Joel and makes several brazen remarks towards Joel. Later on Hallie is killed.



Fig. 7. “Hallie Shouting at Joel.” *Scream 2*, 1997. Author’s screenshot

Joel is depicted as incompetent. He is repeatedly shown being unable to perform simple tasks at his job where he works as a cameraman. While the incompetent stereotype itself is not indicative of typecasting, he is the only one of his peers, all of whom are white, who is depicted as incompetent. He is also harshly berated by his boss and is temporarily fired at one point. These events suggest that he is typecasted.

Overall, *Scream 2* had only eight black characters, half of which were without speaking roles. In addition to a lack of representation, all black characters with speaking roles in the movie were typecasted. It is also worthy of acknowledging that Joel is the only black character with a speaking role to survive, although he was originally slated to be killed but was spared only after actor Duane Martin convinced director Wes Craven to spare Joel.

Bar Chart of *Scream 3* (2000) Cast

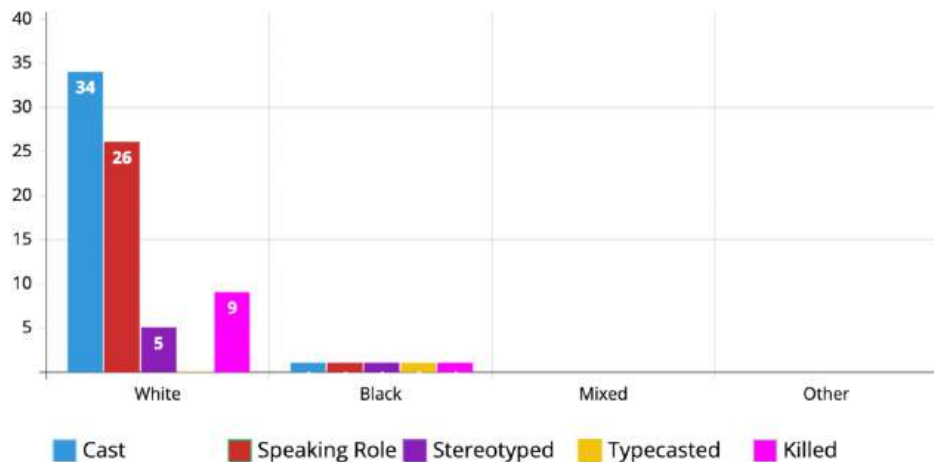


Table 3: Scream 3 (2000) Cast						
Race	Total	Speaking Role	Not Stereotyped or Typecasted	Stereotyped	Typecasted	Killed
White	0.9714	0.9630	1.0000	0.8333	---	0.9000
Black	0.0286	0.0370	---	0.1667	1.0000	0.1000
Mixed	---	---	---	---	---	---
Other	---	---	---	---	---	---

Table 3 shows that the *Scream* series continues its remarkably egregious trend of racial underrepresentation with *Scream 3* (2000). Just 2.8 percent of the credited cast is black. The film's sole minority actor is Deon Richmond, who plays 'Tyson Fox'. Tyson Fox was a sizably large role. Throughout the movie, Tyson served as the sassy, displeased side character. When entering the home of a wealthy man with the other protagonists, Tyson, in a derisive tone, says, "Could the dude have more money?" (Craven, 2000, 1:16:38). Tyson's derisiveness is evident throughout later scenes. The lead protagonist then proposes that the group splits up to find an alleged hidden room within the house. Tyson expressed his exasperation by stating, "Woah woah woah, just one damn minute. There's a psycho killer on the loose and you wanna go traipsing around this gigantic mansion. Have you ever actually seen the Stab movies?" (Craven, 2000, 1:17:03). Once again, this was said with immense insolence. Tyson's objection is entirely disregarded by the rest of the group, and the group ultimately decides to split up. Ghostface enters the house and proceeds to kill five of the eight protagonists in the group. Tyson is the first one targeted by Ghostface, who grabs Tyson and fatally pushed him off the second-floor balcony.



Fig. 8. "Tyson Pushed off Balcony by Ghostface." Scream 3, 2000. Author's screenshot

Of the five deaths that occurred at the house, Tyson’s death was the most pitiable. The other protagonists fought back against Ghostface. Though to no avail, their resistance was admirable. Tyson, however, saw Ghostface coming and did nothing to save himself.

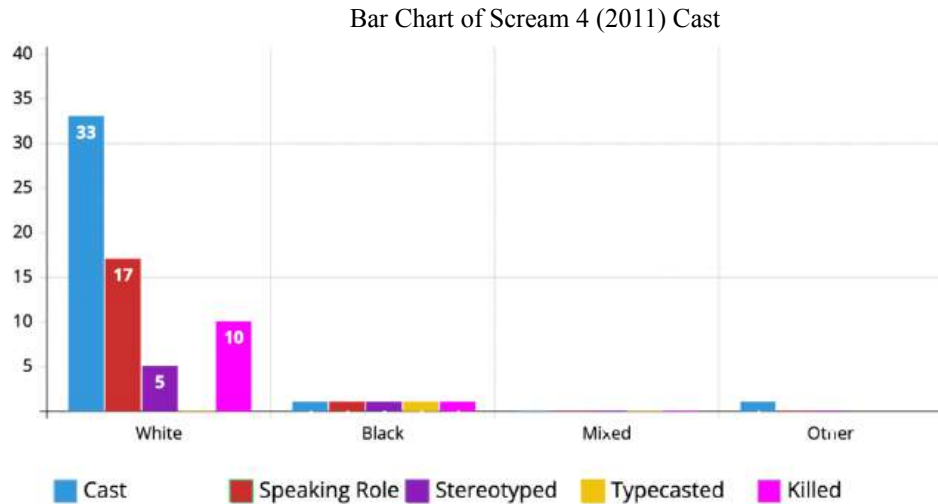


Table 4: Scream 4 (2011) Cast						
Race	Total	Speaking Role	Not Stereotyped or Typecasted	Stereotyped	Typecasted	Killed
White	0.9429	0.9444	0.9655	0.8333	---	0.9091
Black	0.0286	0.0556	---	0.1667	1.0000	0.0909
Mixed	---	---	---	---	---	---
Other	0.0286	---	0.0345	---	---	---

Scream 4 (2011) closely mirrors *Scream 3*’s underrepresentation. The bar chart above shows that out of the 35 credited cast members, there is only one who is black. The only black member of the cast is Anthony Anderson who plays ‘Deputy Perkins’. The role is minor but comes with several spoken lines. When the audience is first introduced to Deputy Perkins, he is stationed outside of the main protagonist’s house at night with another deputy. The two are there to protect the main protagonist because she is being targeted by Ghostface. The camera cuts to the two deputies conversing, when Deputy Perkins sees suspicious movement. He goes to investigate and sees Ghostface only seconds before he is stabbed in the head with a knife. The other deputy then fights Ghostface before being stabbed to death. Similarly to Tyson’s death in

Scream 3, Deputy Perkins is killed with little to no resistance. This theme essentially portrays the black characters as weak and ineffective at fighting and thus denotes typecasting.



Fig. 9. “Deputy Perkins Killed by Ghostface.” *Scream 4*, 2011. Author’s screenshot

Bar Chart of *Scream* (2022) Cast

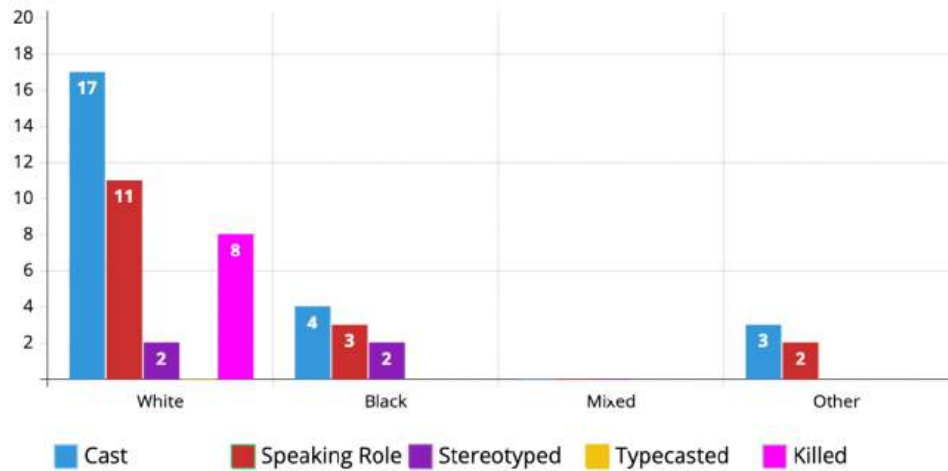


Table 5: Scream (2022) Cast

Race	Total	Speaking Role	Not Stereotyped or Typecasted	Stereotyped	Typecasted	Killed
White	0.7083	0.6875	0.7500	0.5000	---	1.0000
Black	0.1667	0.1875	0.1000	0.5000	---	---
Mixed	---	---	---	---	---	---
Other	0.1250	0.1250	0.1500	---	---	---

The findings above show that the series’s fifth installment, *Scream* (2022), marks the first movie in the series to sufficiently represent African Americans. Of the 24 credited cast members, four were black, which is equivalent to 16.7 percent. Three of the black cast members had speaking roles, which equates to 12.5 percent. Hence, this is the first movie of the series to meet the 12.2 percent national average of black cast members with speaking roles. Additionally, the cast consists of no mixed race cast members and three other race cast members, two of which have speaking roles.

The film’s black cast with speaking roles are Mason Gooding as ‘Chad Meeks-Martin’, Jasmin Savoy Brown as ‘Mindy Meeks-Martin’, and Reggie Conquest as ‘Deputy Farney’. Chad is one of the main protagonists. He is stereotyped as the dull highschool jock. Throughout the film, Chad made a lot of incoherent choices, was highly egoistic, and was easily provoked into fights. Although Chad was a heavily stereotyped character, there was no typecasting, as the stereotypes did not fall along racial lines. Chad’s sister, Mindy, was also a main protagonist and was stereotyped. Mindy was portrayed as the loud, progressive highschool feminist, angry with the patriarchy and lack of equality for minority communities. Like Chad, Mindy’s stereotyping was not race-based, and therefore the actor was not typecasted. Chad and Mindy were one-dimensional characters, completely void of any emotional complexity. Deputy Farney was a smaller character, only appearing in a few scenes. He was portrayed as a typical police officer and behaved similarly to his white officer counterparts.



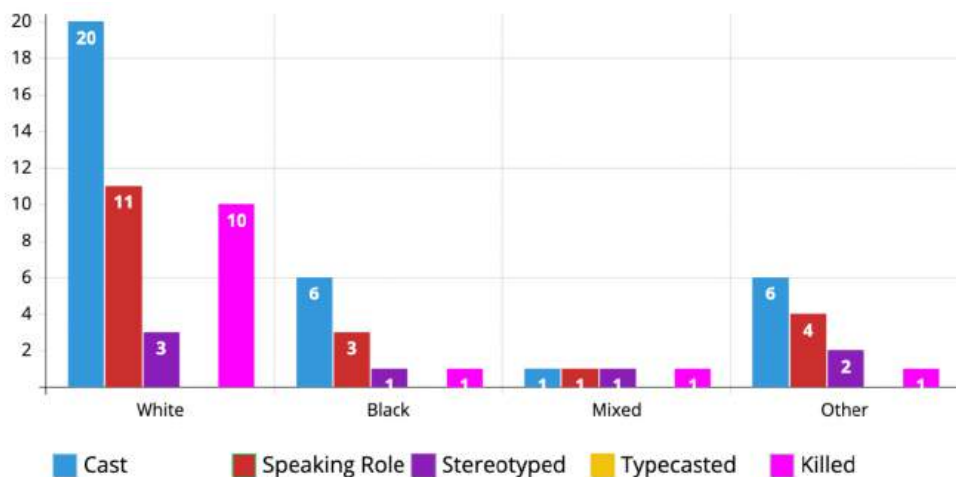
Fig. 10. "Mindy Addressing Friend Group." *Scream*, 2022. Author's screenshot



Fig. 11. "Chad in an Altercation." *Scream*, 2022. Author's screenshot

Scream (2022) was the first film to include actors of other races as characters with speaking roles. The two actors, Melissa Barerra as 'Sam Carpenter' and Jenna Ortega as 'Tara Carpenter' - Sam's sister - are Latina actresses. Their characters are at the center of the film's plot. The film explored the complex dynamic of the sisters' relationship to a degree deeper than what had been done for any other character in the *Scream* franchise. The complexity given to the duo marked a stark contrast from most other characters in the franchise, as most characters had been rooted in stereotypes. The emotional complexity given to the sisters is a probable explanation for why they were not stereotyped or typecasted.

Bar Chart of Scream 6 (2023) Cast



Race	Total	Speaking Role	Not Stereotyped or Typecasted	Stereotyped	Typecasted	Killed
White	0.6061	0.5789	0.7083	0.4286	---	0.7692
Black	0.1818	0.1579	0.2083	0.1429	---	0.0769
Mixed	0.0303	0.0526	---	0.1429	---	0.0769
Other	0.1818	0.2105	0.0833	0.2857	---	0.0769

As seen by the data above, *Scream 6* (2023) is by far the most diverse movie of the series. Six of the 33 credited cast members are black. This equates to an 18.2 percent black cast. Half of the credited black cast members have speaking roles, all of which are substantial. While this means only 9.1 percent of the cast is black with speaking roles - which is below the 12.2 percent national average - African Americans are still adequately represented in the film as they make up nearly half of the major roles. The film also has one mixed race actor with a speaking role, and there are a further six other race actors in the film, four of which have speaking roles.

The black actors with speaking roles are Mason Gooding as ‘Chad Meeks-Martin’, Jasmin Savoy Brown as ‘Mindy Meeks-Martin’, and Thomas Cadrot as ‘Brooks’. *Scream 6* begins to break the stereotypical mold of Chad from the previous film. Chad becomes a nuanced, emotionally complex character. When talking to Tara, his upset friend, Chad shows compassion and empathy for the first time; while attempting to comfort her he says, “You can do a lot better

than date rape Frankie. Tara, you're... I think you're really special... you're only a ghost who has stolen my heart away" (Bettinelli-Olpin & Gillett, 2023, 0:26:53). This is a notable shift from all previous portrayals of Chad as an arrogant, physical douche. He is, for the first time, depicted as empathetic and understanding. During the rest of the scene Chad develops a friendship and romance with Tara. Chad's new-found compassion is built upon later on, this time with Sam. After noticing how upset Sam is after being attacked, Chad reassures her by saying, "Hey, just a reminder, not a single person in this room hates you. We have all been through some fucked-up stuff, and we are all coping with it differently. But, I mean, we moved here together for one specific reason. We are a team" (Bettinelli-Olpin & Gillett, 2023, 0:52:19). This statement also shows a shift from his previous mentality: everyone for themselves. Chad continues to evolve throughout the remainder of the film; the stereotypes from the previous film are shed away.



Fig. 12. "Chad and Tara Connecting." Scream 6, 2023. Author's screenshot

Mindy, unlike her brother, does not experience any expansion of character complexity, and her stereotypes from the previous film are only reinforced. However, her stereotypes remain based solely on generation-z stereotypes, and there is no evidence of typecasting.



Fig. 13. "Mindy Gathering Friend Group." Scream 6, 2023. Author's screenshot

Brooks is a side character in the film. He is in a relationship with Gale Weathers, one of the main protagonists, and serves as her protector when Ghostface enters their home and attempts to kill Gale. He dies while trying to protect Gale. There is no evidence of typecasting or stereotyping.

The film's only mixed race character, Jason Carvey, plays one of several Ghostface killers. In the few scenes he is in, he is depicted as psychotic and sadistic. This depiction coincides with the nature of prior Ghostfaces and is stereotypical.

The four other race actors with speaking roles are Melissa Barrera as 'Sam Carpenter', Jenna Ortega as 'Tara Carpenter', Josh Segarra as 'Danny Brackett', and Devyn Nekoda as Anika Kayoko. Sam and Tara continue to grow in complexity, and a major arc of the movie is centered around their sisterly reconciliation. The two exhibited no signs of stereotyping or typecasting. Danny Brackett, Sam's boyfriend, is stereotyped similarly to Chad. Danny is physical, confrontational, and easily provoked. However, there is not much else to his character, as his character is fairly one-dimensional. Devyn is Mindy's girlfriend, and she has most of the same stereotypes as Mindy. She is depicted as the generation-z activist who lacks morality.

Conclusion

The results captured from the research process confirm that the first four films in the *Scream* film series conspicuously underrepresent African Americans. Additionally, films two and three exhibit blatant signs of typecasting. The series furthers its lack of racial inclusion with mixed and other race actors. There are no mixed, credited actors in the first four films, and there are no other race actors with speaking roles in these films. It is important to acknowledge that the first four movies were released in the period of 1996-2011, before the #OscarsSoWhite phenomenon. Furthermore, typecasting had the effect of portraying African Americans in an immensely negative way. Black characters were repeatedly given similar tropes such as the sassy best friend, the incompetent employee, and the loudmouth. These racial stereotypes carry connotations of flamboyant, obnoxious, profane, and inept. Hallie, portrayed as the white-washed black girl, is the only black character in the first four films portrayed in a mildly positive way, although this is most likely because she possesses few characteristics of being black, aside from the color of her skin.

Films five and six were released after the rise in popularity of #OscarsSoWhite. The films made drastic improvements in racial representation. Black actors were sufficiently represented and were not typecasted. There were additional improvements made in the representation of other race actors, though there are few mixed actors.

I initially thought that the films released prior to #OscarSoWhite would have typecasted the black actors, however, I did not account for the overall absence of black actors with substantial roles. This accounts for why films one and four did not have any typecasted black actors. Therefore, the data suggests that typecasting was moderately prevalent in the films first released, aside from those that omitted black actors from important roles, and then decreased in prevalence from subsequently released films.

It is important to note that the open coding research process selected for this project enabled this study to account for all aspects of the ways in which characters were portrayed. The process became especially helpful in films five and six, where the characters were portrayed in an incredibly nuanced and complex way. The particular advantages of the selected research process were that I was able to account for portrayals I had not predicted, hence the prior process would have omitted data, and through multiple viewings of each film I was able to soundly understand each character.

There is an ethical consideration for this project. The screenshots included from the films are copyrighted. Fortunately, I can use these screenshots under the “fair use” doctrine, since I am using them for educational purposes and am not profiting from them (U.S. Copyright Office, 2022).

Error and Future Research

While documenting the race of all credited cast members in each movie on IMDb, it was unclear whether several actors were black or mixed. However, none of the actors whose race I was unclear of had speaking roles. Therefore, the data should not have been skewed.

Although I attempted an analysis of typecasting in the *Scream* series, there is still further research to be done within the series. By surveying viewers of the films on their perceptions of the black characters, a greater understanding of the impact of typecasting can be gained. Further, there are scenes with other race characters that I neglected discussing in my findings and analysis. However, these interactions and portrayals were accounted for in the overall number of other race actors who were stereotyped or typecasted. Finally, it is important to acknowledge that the *Scream* films are by no means the only films to employ the practice of typecasting, especially in the horror genre. Although the franchise deserves praise for improving representation and eliminating typecasting in recent years, the franchise holistically underrepresents African Americans and portrays them in an unfavorable manner. Nevertheless, the latest two films are able to achieve a depiction of their black characters in a way that is more authentic, multifaceted, and favorable than I have seen in many other recently released films in the horror genre. Extending a critical content analysis to the *Scream* series allows for its faults to be revealed and shows that there is still room for improvement, while additionally showing how complex the issues of racial representation and typecasting in film can be. Continuing to put effort into adequately and appropriately representing African Americans in film is crucial to challenging existing ideologies influenced by prejudice and misconception in America.

Works Cited

- Bloomberg, L. D., & Volpe, M. (2015). *Completing your qualitative dissertation: A road map from beginning to end*. In Google Books. SAGE Publications.
https://books.google.com/books/about/Completing_Your_Qualitative_Dissertation.html?id=Pq9iCgAAQBAJ
- Bettinelli-Olpin, M., & Gillett, T. (Directors). (2022). *Scream* [Streamed]. Paramount Pictures.
- Bettinelli-Olpin, M., & Gillett, T. (Directors). (2023). *Scream 6* [Streamed]. Paramount Pictures.
- Craven, W. (Director). (1996). *Scream* [Streamed]. Dimension Films.
- Craven, W. (Director). (1997). *Scream 2* [Streamed]. Dimension Films.
- Craven, W. (Director). (2000). *Scream 3* [Streamed]. Dimension Films.
- Craven, W. (Director). (2011). *Scream 4* [Streamed]. Dimension Films.
- Elbaba, R. (2019, November 14). Why on-screen representation matters, according to these teens. PBS NewsHour.
<https://www.pbs.org/newshour/arts/why-on-screen-representation-matters-according-to-these-teens>
- Hall, S. (1997). Representation: Cultural representations and signifying practices. American Psychological Association. <https://psycnet.apa.org/record/1997-36930-000>
- IMDb. (1990). Ratings and Reviews for New Movies and TV Shows - IMDb. IMDb.
https://www.imdb.com/?ref_=nv_home
- Lee, W. (2023, March 30). Are Hollywood movies backsliding on diversity? What the latest numbers say. Los Angeles Times.
<https://www.google.com/url?q=https://www.latimes.com/entertainment-arts/business/story/2023-03-30/are-hollywood-movies-backsliding-on-diversity-what-the-latest-numbers-say-ucla-diversity-report&sa=D&source=docs&ust=1699299442753226&usg=AOvVaw0mxvkXYZ2YhUahTx6alGkH>
- Malik, M., Hopp, F., & Weber, R. (2022). Representations of racial minorities in popular movies: A content-analytic synergy of computer vision and network science. *Computational Communication Research*, 4(1).
<https://computationalcommunication.org/ccr/article/view/106>
- Smith, S. (2023). Annenberg inclusion initiative. USC Annenberg School for Communication and Journalism. Usc.edu. <https://annenberg.usc.edu/research/aii>
- Suhling, C. (2017). The underrepresentation of African Americans and the role of casting directors. In *Undergraduate Research School of Communications* (pp. 1–89).
https://scholarworks.gvsu.edu/cgi/viewcontent.cgi?article=1000&context=com_undergrad
- U.S. Copyright Office. (2022, August). U.S. Copyright Office fair use index. U.S. Copyright Office. <https://www.copyright.gov/fair-use/>
- United States Census Bureau. (2022, July 1). U.S. Census Bureau quickfacts: United States. United States Census Bureau.
<https://www.census.gov/quickfacts/fact/table/US/PST045222>
- Wolf, J. (2020, February 6). 2020 Hollywood diversity report: A different story behind the scenes. UCLA. <https://newsroom.ucla.edu/releases/2020-hollywood-diversity-report>

Host Cities' Labor Market Reactions To The Super Bowl By Jason Luo

Abstract

The Super Bowl, one of America's major sporting events, is the pinnacle of football showcasing. However, the game not only impacts sport fanatics, but also host cities and their economies. As a result of an abnormal influx of visitors, there are multiple guidelines relating to adequate existing infrastructure for a successful venue. Other than the overall view of the city, its economy is also shifted to deal with more people. Local businesses for example, offer more job openings in preparation for handling larger customer amounts. The purpose of this study is to examine the degree of change in employment for past host cities, as well as the longevity of the changes. Change in employment will be measured using these metrics: population, labor force, number of employed and unemployed people, and the unemployment rate. The population of a host city will be measured two years prior to the Super Bowl, the year of the Super Bowl, and two years after the Super Bowl. The other three metrics will be measured six months before the Super Bowl, the month of the Super Bowl, and six months after the Super Bowl. Thirteen host cities, starting in 2010 up to 2022, will be analyzed.

Introduction

Despite its colossal television numbers, the Super Bowl also attracts many in-person visitors to the year's hosting city. With the influx of visitors, many economists have proved the clear benefits for the city as a result. These benefits range from greater city exposure, economic gain, and community connection. Because of these vast benefits, cities lunge at the opportunity at hosting the coveted event. In the past, many cities were invited by the NFL to make a bid, with only a few becoming "finalists". These finalists were asked to submit a presentation to the NFL's 32 team owners, who selected the winning city. Since 2018, the NFL switched to a smaller-scale competition, as the NFL started to first contact certain venues to put together proposals. The owners would still have the final say. Selected cities often have newly renovated or created stadiums, optimal weather conditions, necessary infrastructure, and enough hospitality and entertainment areas.

An additional result of the city's influx of visitors is the influx of workers to supply the increase in demand in products and services for all industries. As an example, Super Bowl LVI reportedly brought in between 2,200 to 4,700 new jobs to the LA area. The service industry is traditionally the primary benefiter, with more visitors requiring more restaurant and hotel workers, as well as more taxi and rideshare drivers. With the labor market being excessively affected by the Super Bowl, this study seeks to describe the host city's labor industry's reaction to hosting the Super Bowl. The study uses four major metrics to measure observe the impact on the labor market. A collection of 13 Super Bowls (2010-2022) and their host cities are analyzed. The objective is to determine how these four major metrics have responded in each of these host cities and to create an overall trend for each metric for all cities. Other studies have often referenced the economic benefits of the Super Bowl. However, these benefits are traditionally

limited to the business opportunities, overall GDP increases, and benefits in the hospitality, entertainment, and tourism industries. However, a city’s labor market’s short-term and long-term response to hosting the Super Bowl is normally not reported and is vague in terms of its longevity.

This is unfortunate, given the importance of a city’s labor market. The findings of this study provide evidence of the impact of the Super Bowl on a host city’s labor market both in the short and long-run. This could help city officials, employers, and employees understand potential benefits their city and they themselves could receive from hosting the Super Bowl. Additionally, future hosts can predict the benefit for their labor market based on similarities between their city and the cities discussed in this study.

Super Bowl Impact on Host Cities’ Population Sizes

Figure 1 presents the population % change over the five years of each Super Bowl host city in this study. Out of the thirteen cities, nine experienced a population increase and four experienced a decrease. Of the four cities that experienced a decrease, three are the last three cities of the study. This means that a portion of their period is affected by COVID-19’s population decreasing effect (As many as 73% of US counties experienced a population decline in 2021). Despite the apparent correlation between Super Bowls resulting in population growth in host cities, it is notable that population growth is a result of numerous variables, such as city size, the state of the economy, birth rate, death rate, immigration, and emigration. Still, it is conclusive that the Super Bowl Committee specifically targets growing cities to host their Super Bowls.

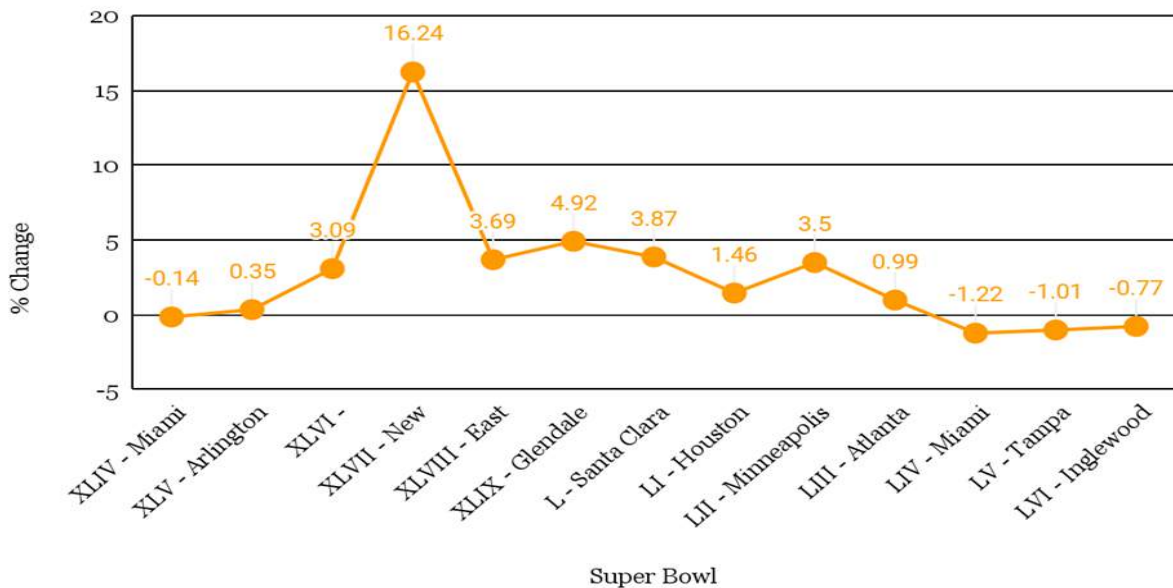


Fig 1: Population % Change Over 5 Years VS Super Bowl Host City

Super Bowl Impact on Host Cities' Labor Forces

Figure 2 presents the combined labor force % change of each month over the year (August to August) of each Super Bowl host city of this study. Out of the thirteen cities, nine experienced a labor force increase and four experienced a decrease. The stand-out data point is Super Bowl LIV, hosted by Miami. The period is from August 2019 to August 2020, meaning that the Miami labor force experienced the most detrimental economic effects of COVID-19. COVID's impact would have negated the Super Bowl's impact, thus we observe the unique data point. Of the other host cities, the range of % change is -1.14% to 2.95%. To observe the shorter-term impacts of the Super Bowl on the labor force, the time periods must be changed.

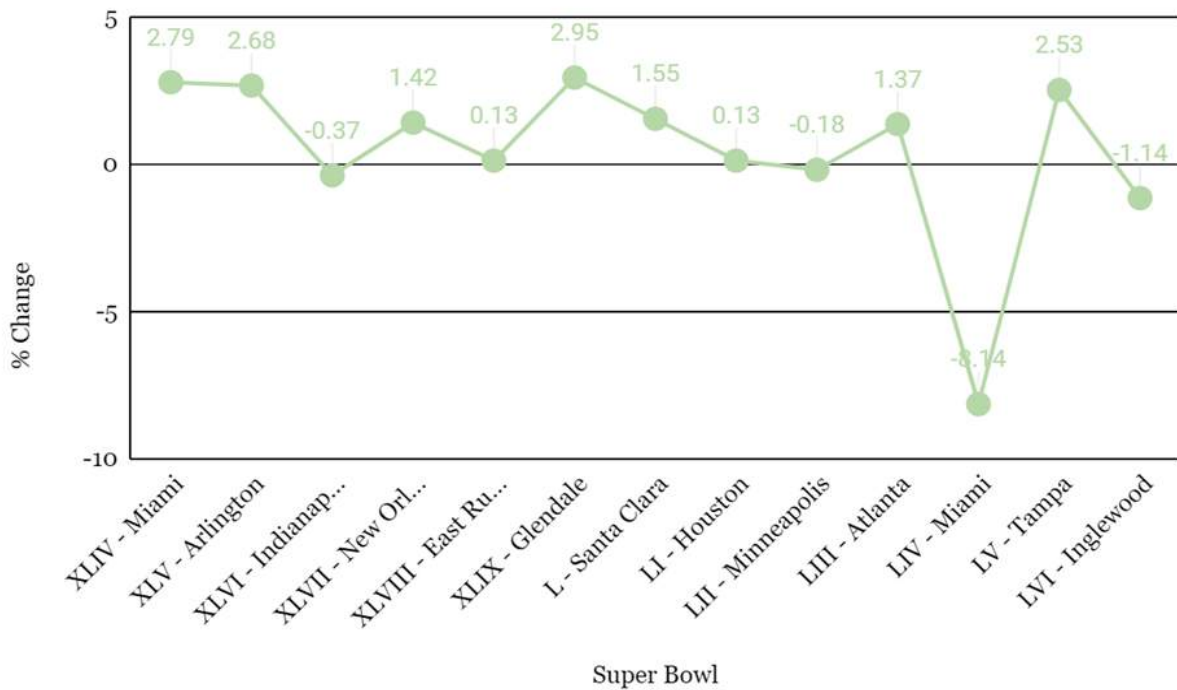


Fig 2: Labor Force % Change Over One Year VS Super Bowl Host City

Figure 3 presents the labor force % change between August (Month-6) and Super Bowl February for the past thirteen Super Bowl host cities. Of the thirteen cities, three display a decrease in labor force. The other ten cities display slight increases, which suggests that the Super Bowl month exhibits a greater labor market than six months before. This does not, however, clearly define when the labor market actually started to grow.

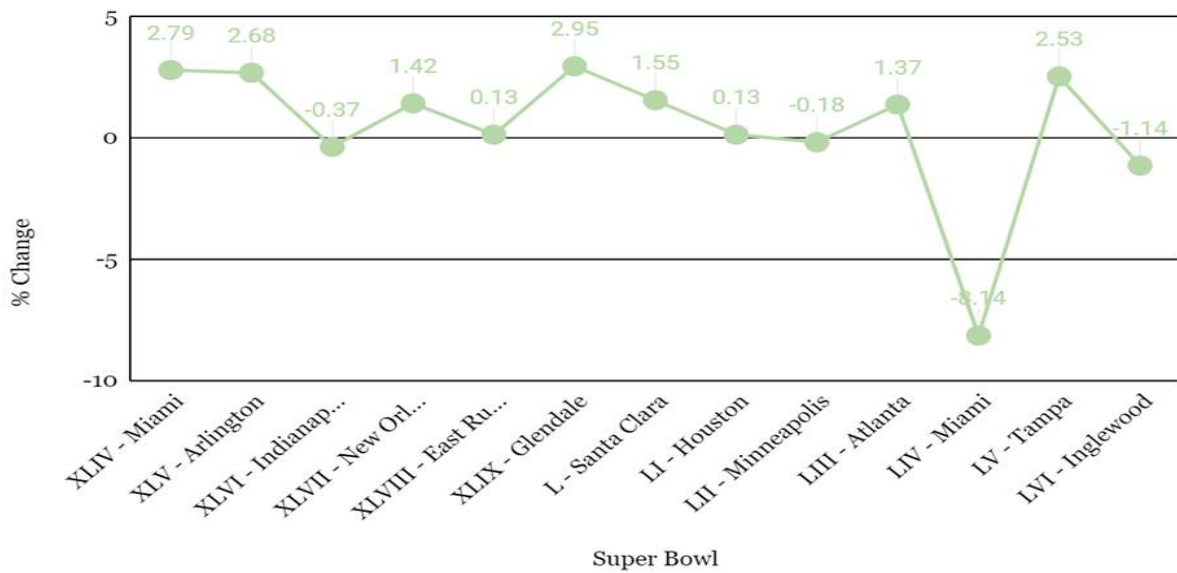


Fig 3: Labor Force % Change Pre-Six Months VS Super Bowl Host City

Figure 4 presents the labor force % change between November (Month-3) and Super Bowl February for the past thirteen Super Bowl host cities. Compared to the findings of Figure 3, around half of the host cities three months before the Super Bowl experience smaller changes. This means that for certain host cities, labor forces were already increasing in the six- through four-month time period before the Super Bowl. However, the other half's labor market experienced greater changes in the three- through one-month period before the Super Bowl. This is also supported by the decreases in % change, which could mean that the Super Bowl does not increase labor force for certain cities.

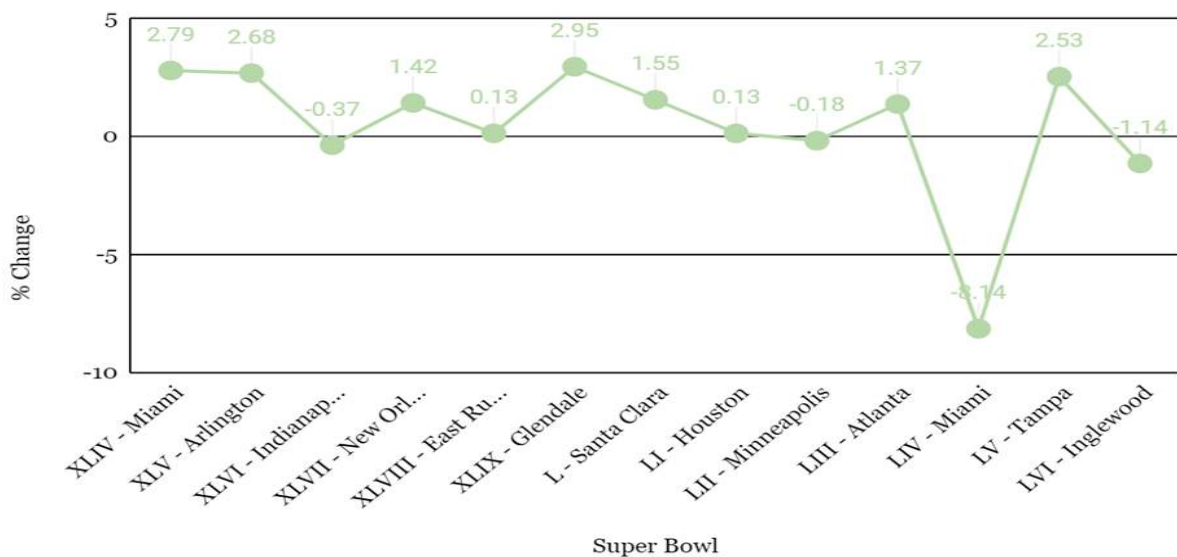


Fig 4: Labor Force % Change Pre-Three Months VS Super Bowl Host City

Figure 5 presents the labor force % change between January (Month-1) and Super Bowl February for the past thirteen Super Bowl host cities. From January to February, three of the last thirteen host cities experienced a decrease in labor force. The range for the increases is 0.28% (Indianapolis) to 0.97% (East Rutherford). From Figure 2, the range of the increases over the entire year is 0.13% to 2.95%. This means that there is an abnormal surge in labor force between January and February, which can be attributed to the Super Bowl. The Super Bowl may influence the labor force earlier, however the greatest surge is in this period.

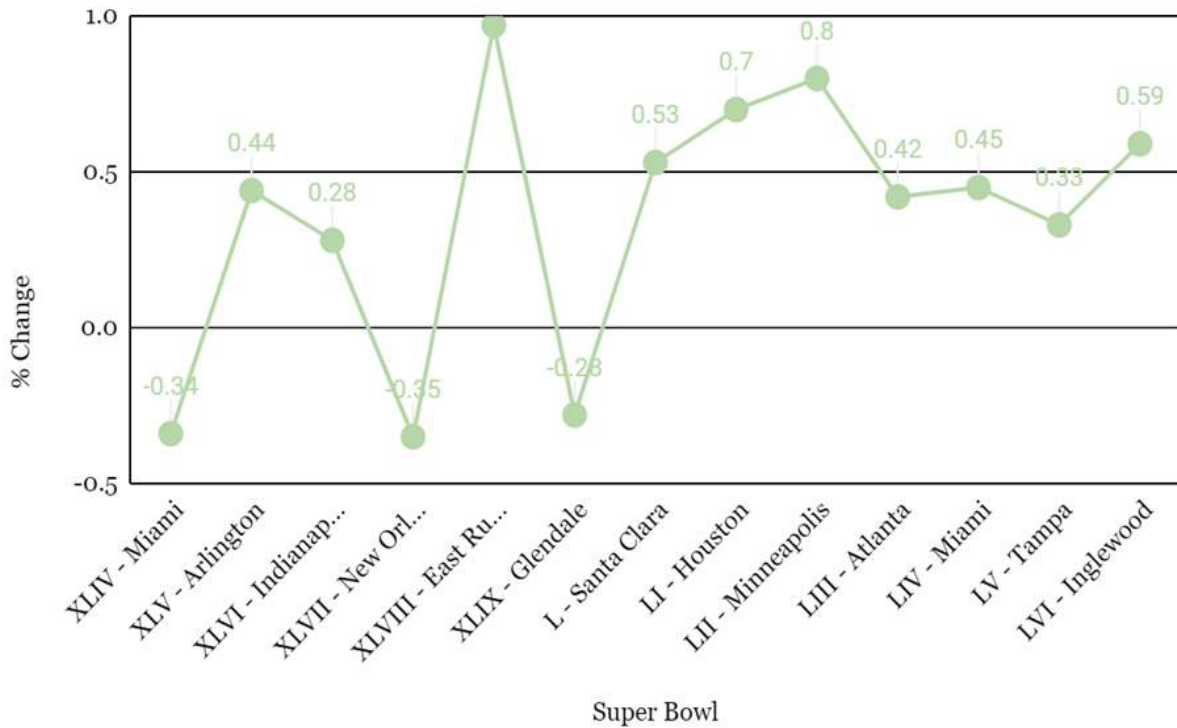


Fig 5: Labor Force % Change Pre-One Month VS Super Bowl Host City

Figure 6 presents the labor force % change between Super Bowl February and March (Month+1) for the past thirteen Super Bowls. Out of the thirteen cities, seven experienced a decrease in labor force, one experienced no change, and five experienced an increase. This means that in the one month period after the Super Bowl, certain cities' labor markets lost people. This suggests that the Super Bowl's observable labor force changes may be temporary. It also means that the greatest labor force changes as a result of the Super Bowl may fall between January and March.

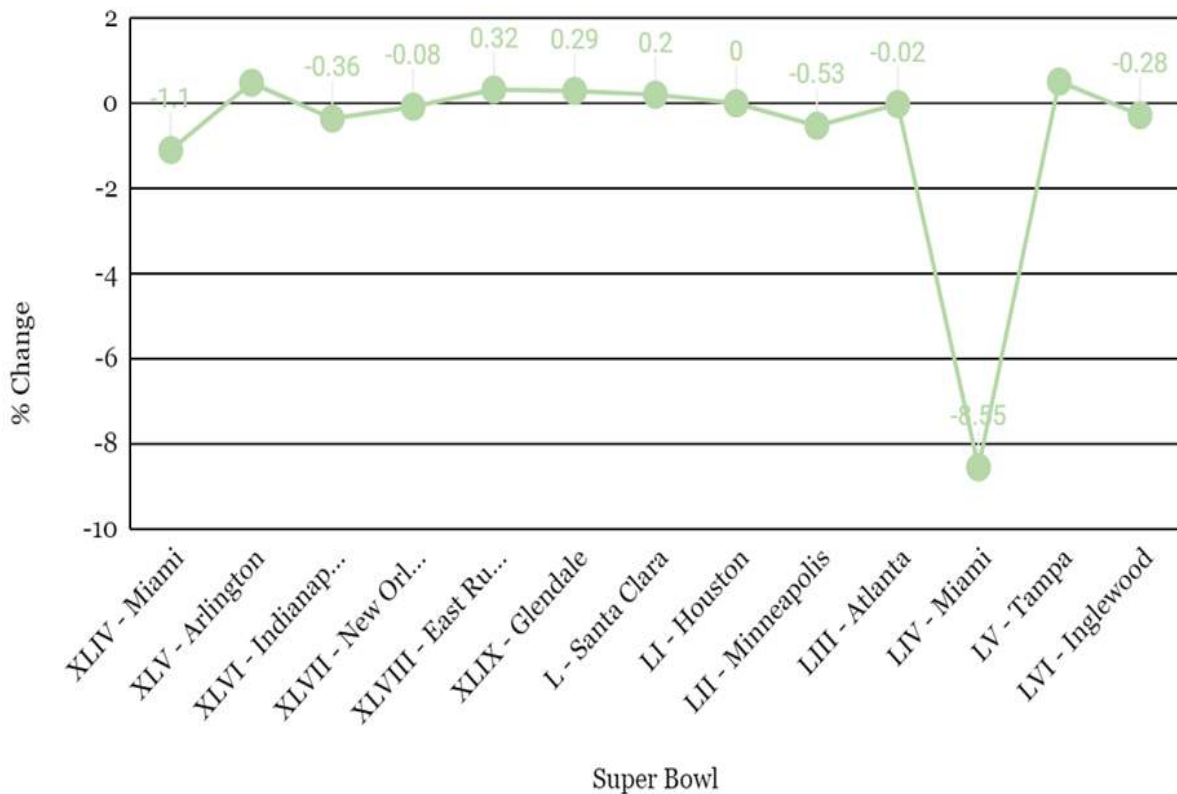


Fig 6: Labor Force % Change Post-One Month VS Super Bowl Host City

Figure 7 presents the labor force % change, combining each month's change in the three month period following the Super Bowl for the last thirteen host cities. Out of the thirteen cities, six experienced a labor force decrease and seven experienced an increase. Since Miami's period was at the height of COVID, we can count it as an outlier. XLIV in Miami and XLVI in Indianapolis both experienced decreases in Figure 6, whereas they experienced an increase in this longer time period. This suggests that their markets substantially increased following the initial decline, hinting that the Super Bowl impact no longer affected the local labor force. The majority of other cities, including East Rutherford, Santa Clara, Houston, Minneapolis, Atlanta, and Inglewood, display a decrease in % change in this three month time period compared to the one month following the Super Bowl. This means that the combined % changes of April (Month+2) and May (Month+3) were usually negative. Thus, in the majority of host cities, the local labor force returned to pre-Super Bowl conditions in no more than three months.

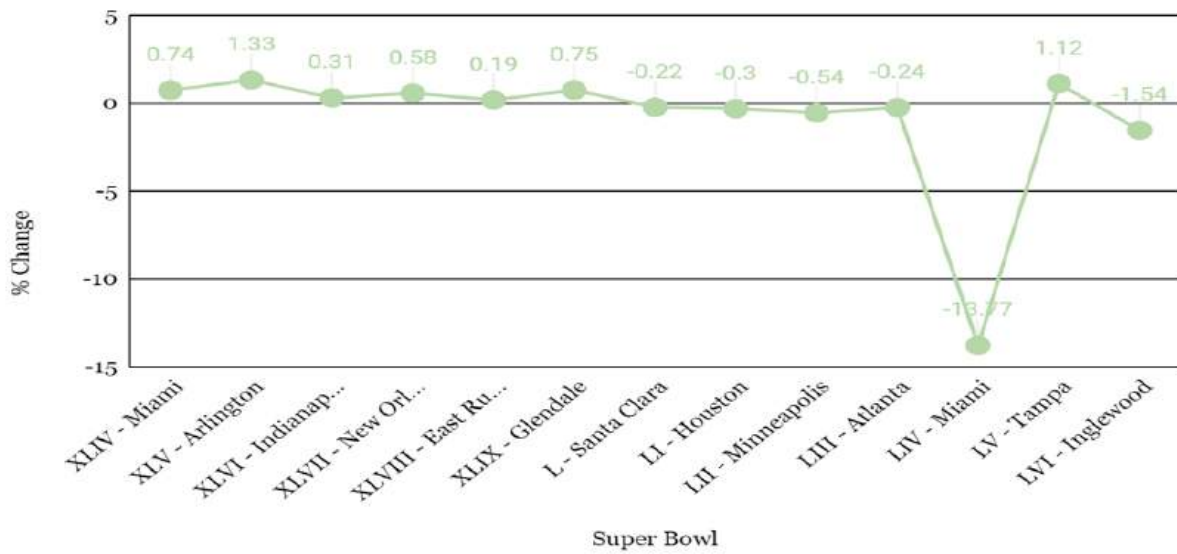


Fig 7: Labor Force % Change Post-Three Months VS Super Bowl Host City

Figure 8 presents the labor force % change between Super Bowl February and August (Month+6) in each of the last thirteen host cities. Out of the thirteen host cities, two experienced a decrease in labor force while the other eleven experienced an increase. When comparing this time period's data to Figure 7, every host city except Inglewood had a greater % change in this time period. When comparing this data to the six months prior to the Super Bowl, the range of % change is smaller following the Super Bowl, however more cities experienced an increase in labor force. This means that the observable labor force increases caused by the Super Bowl had completely subsided in six months and that Super Bowls are normally hosted in cities with a growing labor force.

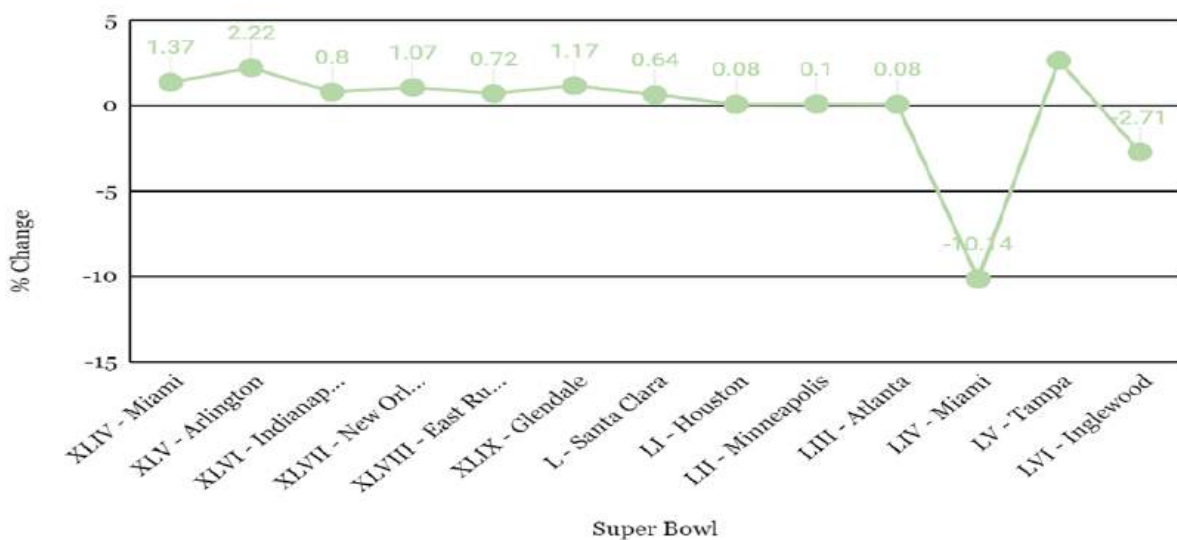


Fig 8: Labor Force % Change Post-Six Months VS Super Bowl Host City

Super Bowl Impact on Host Cities' Number of Employed and Unemployed People

Figure 9 presents the number of employed people % change over a one-year span from August to August. Out of the thirteen host cities, only one experienced an employment decrease and the other twelve experienced an increase. Miami's timespan includes the height of COVID, during which numerous people lost their jobs. This means that the decrease was more of a result of the pandemic rather than the Super Bowl. The increase in the number of employed people may support the long-term expansionist theme of the Super Bowl, however the time frame is not select enough to conclude the short-term employment impact.

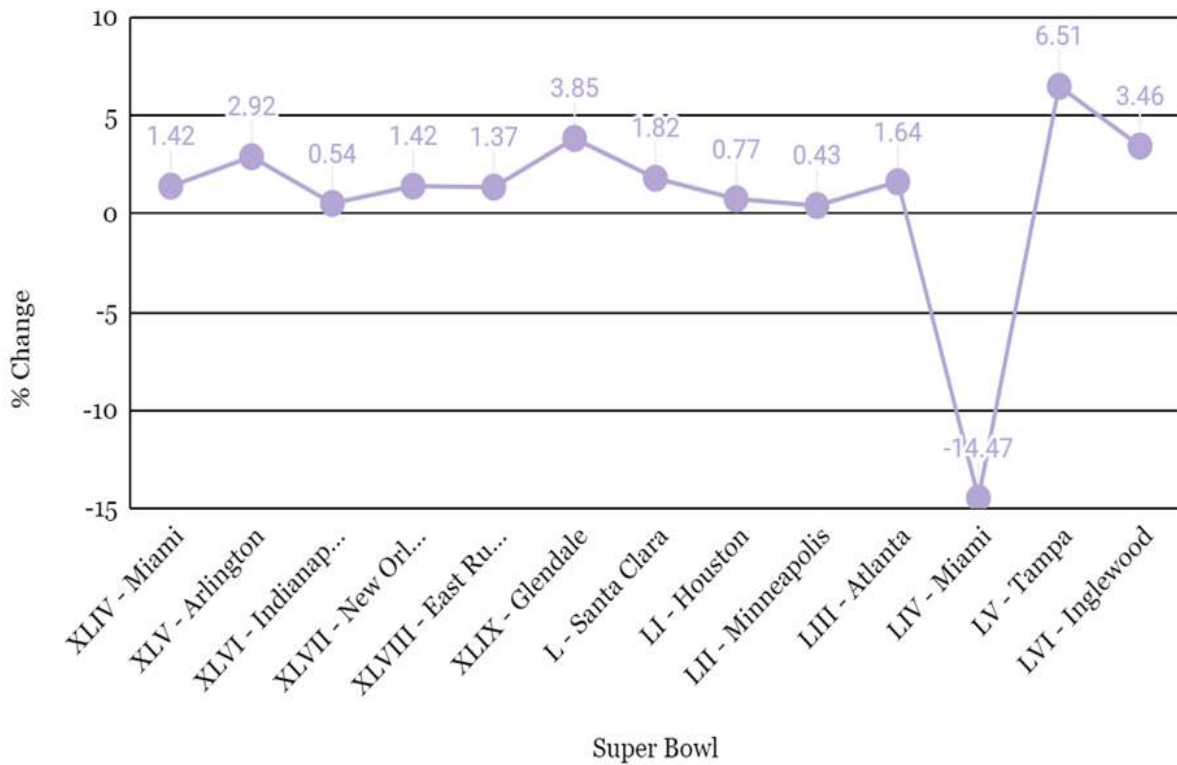


Fig 9: Number of Employed People % Change Over One Year VS Super Bowl Host City

Figure 10 presents the number of unemployed people % change over a one-year span from August to August. Out of the thirteen host cities, nine experienced an unemployment decrease and four experienced an increase. Miami, Tampa, and Inglewood can be considered outliers, as they experienced either the height of COVID-19 or the economy's rebound after the initial downturn caused by the pandemic. The majority of cities experiencing a decrease in unemployed people supports the expansionist theme of the Super Bowl.

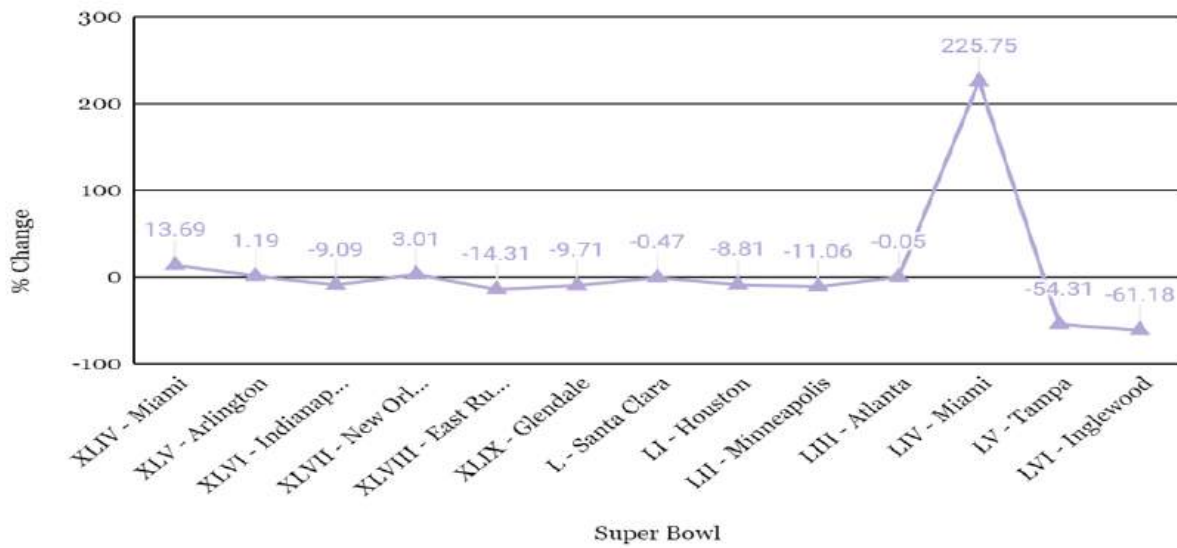


Fig 10: Number of Unemployed People % Change Over One Year VS Super Bowl Host City

Figure 11 presents the number of employed people % change before the Super Bowl. The timespans included are January (Month-1) to February (Super Bowl), November (Month-3) to February (Super Bowl), and August (Month-6) to February (Super Bowl). The overall trend is that the Month-6 -> Super Bowl line has the greatest change, then the Month-3 -> Super Bowl line, and then the Month-1 -> Super Bowl line. Interestingly, there is a noticeable gap between the Month-6 and Month-3 line, which suggests that the Super Bowl is already employing individuals in the period between August and November. Still, the greatest gap is between Month-3 and Month-1, where it is evident that the Super Bowl is rapidly employing individuals.

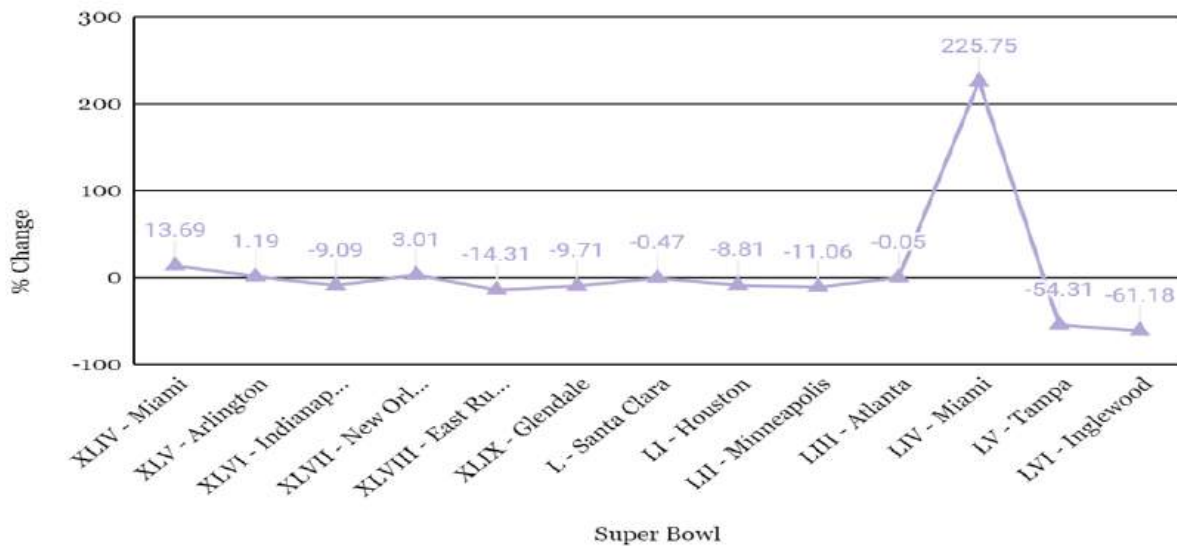


Fig 11: Number of Employed People Before the Super Bowl % Change VS Super Bowl Host City

Figure 12 presents the number of unemployed people % change before the Super Bowl. The timespans included are January (Month-1) to February (Super Bowl), November (Month-3) to February (Super Bowl), and August (Month-6) to February (Super Bowl). The overall trend is Month-1 -> Super Bowl with the greatest % change, then Month-3 -> Super Bowl, and then Month-6 -> Super Bowl. The gap between Month-6 and Month-3 is also evident here, which suggests that unemployment between August and November decreases as a result of the Super Bowl. Interestingly, both Month-3 and Month-1 are mainly in the positives, meaning that in certain cases, unemployment may either increase during the Super Bowl, or workers are immediately laid off after the Super Bowl.

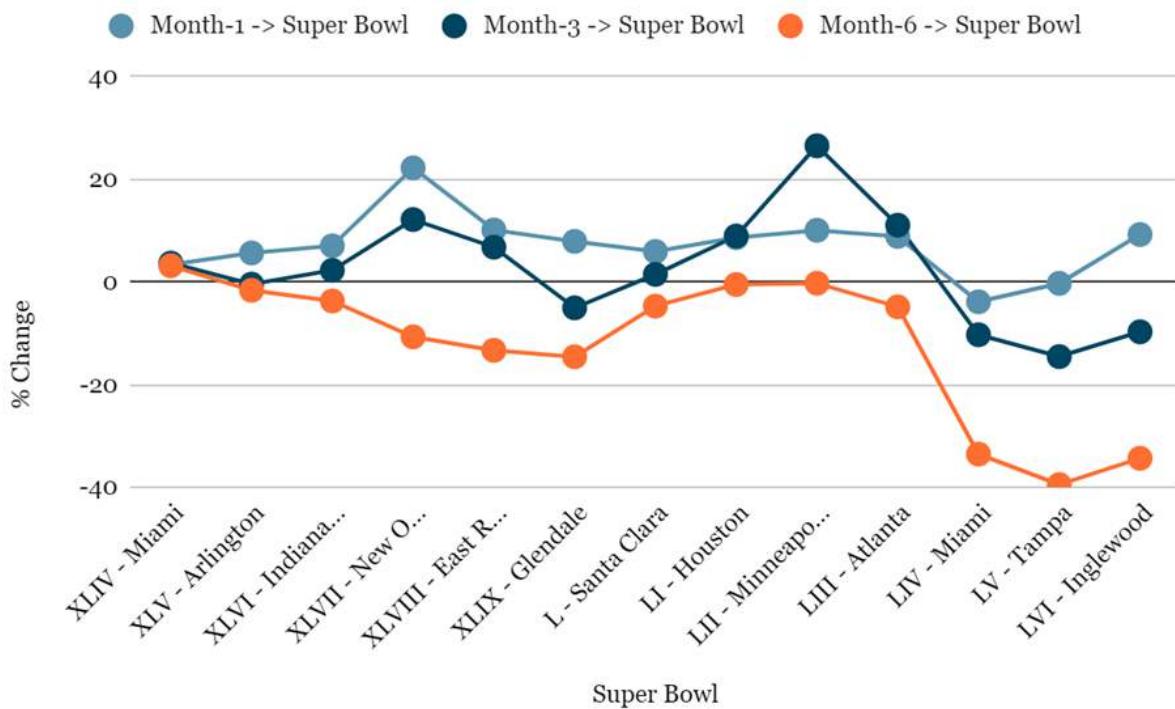


Fig 12: Number of Unemployed People Before the Super Bowl % Change VS Super Bowl Host City

Figure 13 presents the % change in number of employed people after the Super Bowl for the past thirteen Super Bowls and their host cities. The timespans included are February (Super Bowl) to March (Month+1), February (Super Bowl) to May (Month+3), and February (Super Bowl) to August (Month+6). Super Bowl LIV in Miami is a major outlier given COVID-19's impact on employment. Including it in our data would disrupt the overall trend, thus Figure 14 was created with the same data, just excluding Super Bowl LIV in Miami. With the scale change, it is easier to observe that Super Bowl -> Month+6 has the greatest increase, then Super Bowl -> Month+3, and lastly Super Bowl -> Month+1. This means that even after the Super Bowl, employment continues to increase. In fact, the greater the amount of time that has elapsed, the greater the number of employed people is.

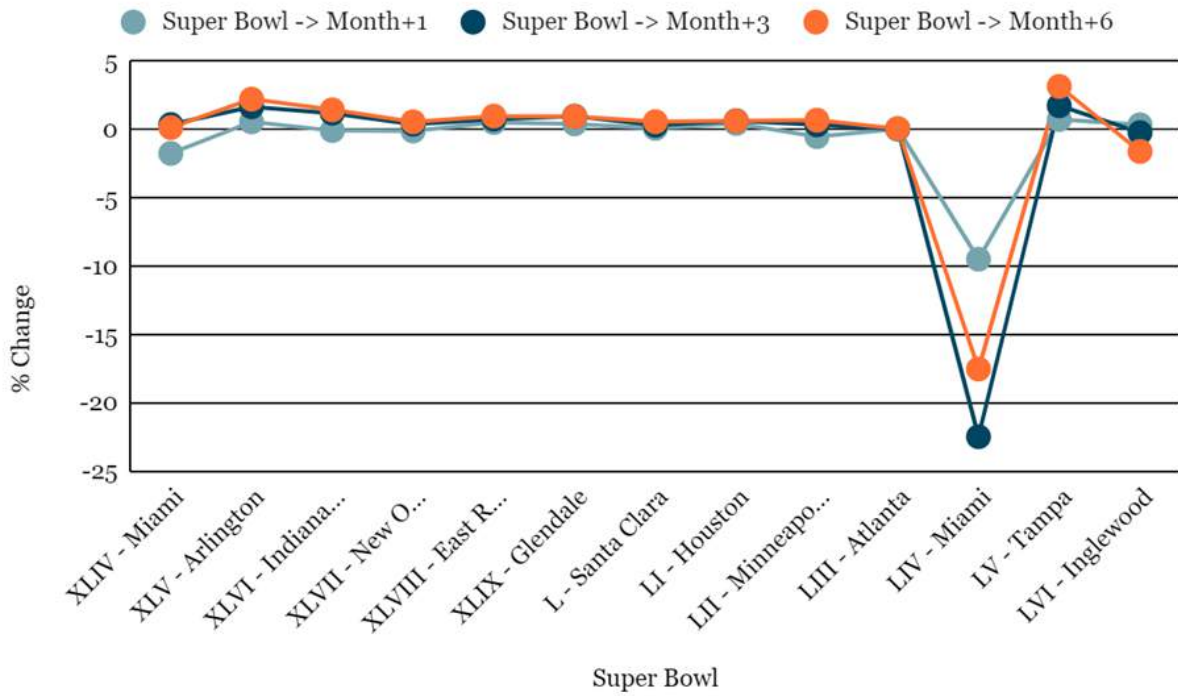


Fig 13: Number of Employed People After the Super Bowl % Change VS Super Bowl Host City

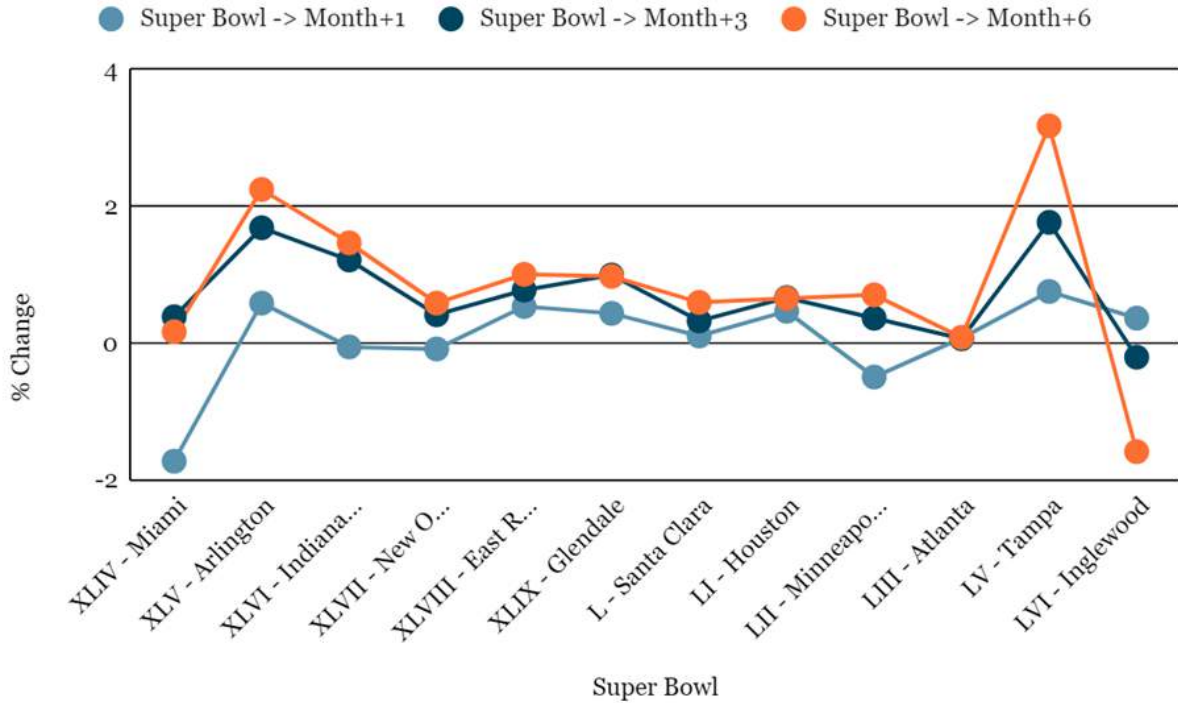


Fig 14: Number of Employed People After the Super Bowl % Change VS Super Bowl Host City (Excluding Miami LIV)

Figure 15 presents the % change in number of unemployed people after the Super Bowl for the past thirteen Super Bowls and their host cities. The timespans included are February (Super Bowl) to March (Month+1), February (Super Bowl) to May (Month+3), and February (Super Bowl) to August (Month+6). Again, Super Bowl LIV in Miami is a major outlier given COVID-19's unemployment increase. Thus, Figure 16 presents the same data, excluding Miami. Given the much smaller scale, it is observable that the overall trend is Month+6 with the greatest % in change, then Month+1, then Month+3. This means that after the Super Bowl, the number of unemployed people continues decreasing between March (Month+1) and May (Month+3). This means that in the short-term, the Super Bowl may be the result of low unemployment in a region for around three months.

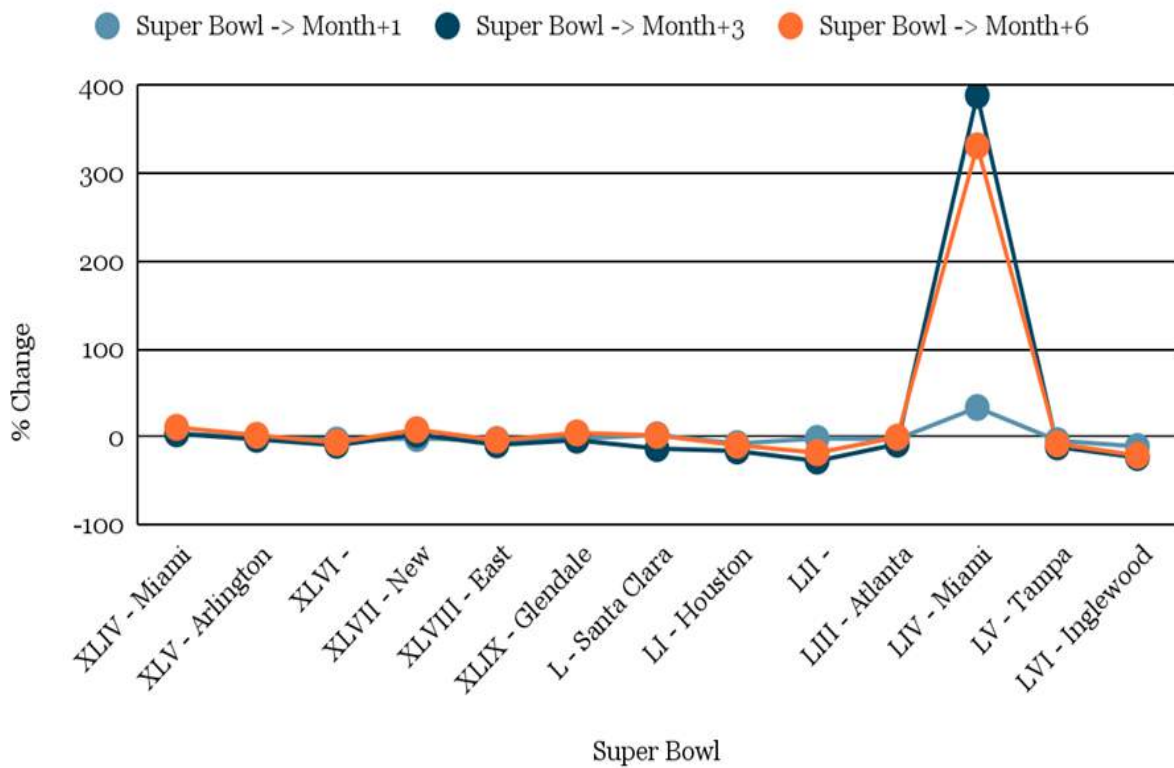


Fig 15: Number of Unemployed People After the Super Bowl % Change VS Super Bowl Host City

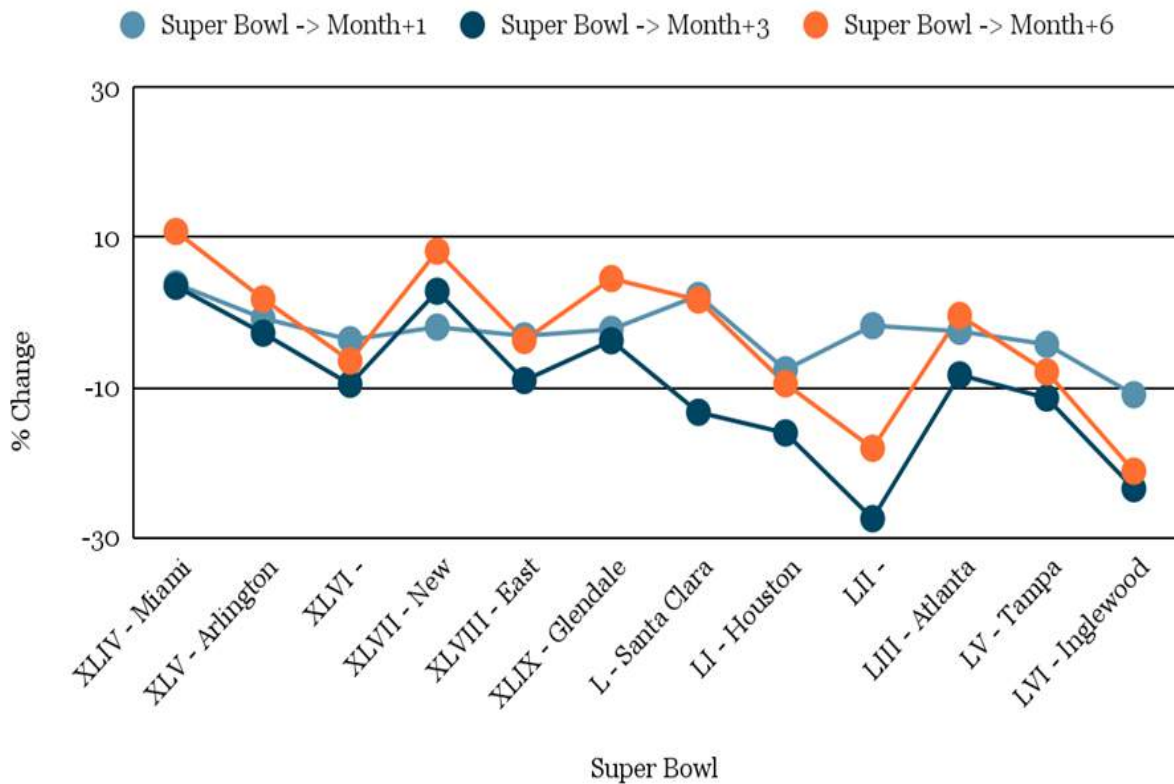


Fig 16: Number of Unemployed People After the Super Bowl % Change VS Super Bowl Host City (Excluding Miami LIV)

Super Bowl Impact on Host Cities' Unemployment Rate

Figure 17 presents the % change in the unemployment rate before the Super Bowl for the past thirteen host cities. The three changes are the difference between the unemployment rate during the Super Bowl and August (Month-6), the difference between the Super Bowl and November (Month-3), and the difference between the Super Bowl and January (Month-1). The difference between the Super Bowl and November (Month-3) is the greatest, then the difference between the Super Bowl and January (Month-1), and finally the difference between the Super Bowl and August (Month-6). This means that three months out from the Super Bowl, there is actually the lowest number of unemployed individuals in comparison to the labor force. A possible explanation for this is that workers are already employed three months before the Super Bowl and individuals who join the market in the period between November and February (Super Bowl) are not hired, raising the unemployment rate.

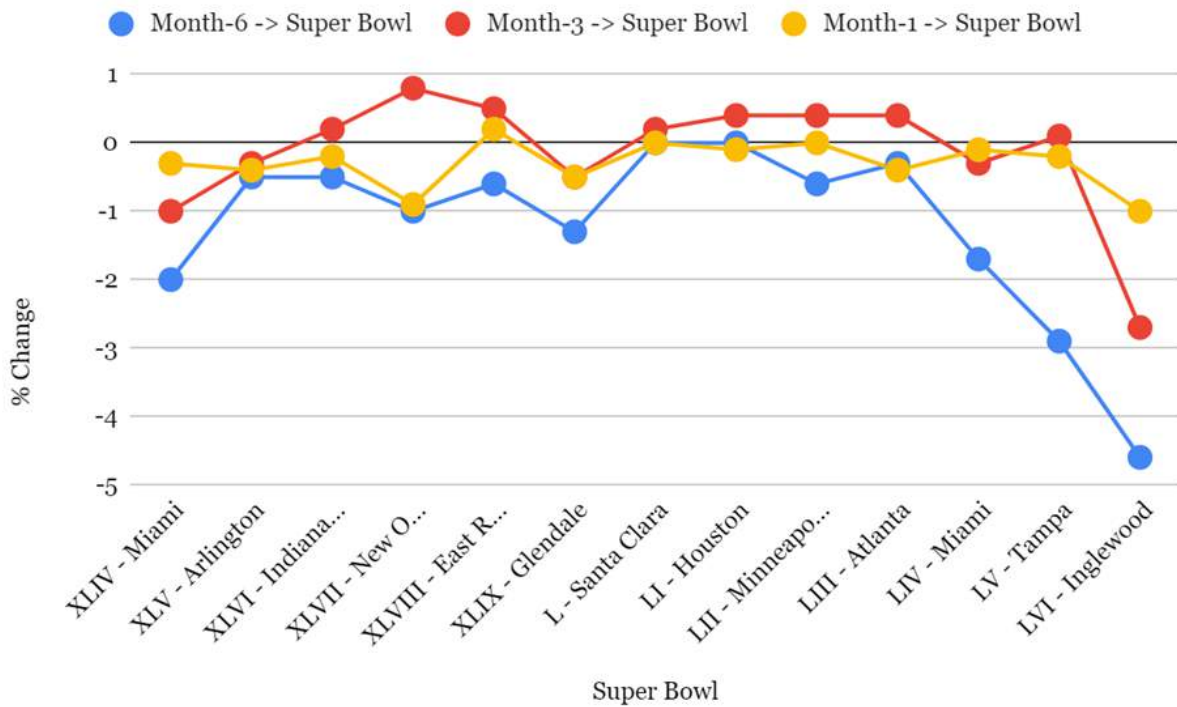


Fig 17: Host City Unemployment Rate Before the Super Bowl % Change VS Super Bowl Host City

Figure 18 presents the % change in the unemployment rate after the Super Bowl for the past thirteen host cities. The three changes are the difference between March (Month+1) and February (Super Bowl), the difference between May (Month+3) and February (Super Bowl), and the difference between August (Month+6) and February (Super Bowl). When observing the figure, it is apparent that Super Bowl LIV in Miami is a major outlier. This is due to COVID-19 majorly impacting the economy following Super Bowl LIV, skewing the unemployment rate. Thus, Figure 19 was created, plotting the same data except Super Bowl LIV in Miami. The greatest % change is between Month+1 and Super Bowl, then Month+6 and Super Bowl, and finally Month+3 and Super Bowl. This means that the unemployment rate is the highest one month following the Super Bowl in the short-run time period (6 months). This means that between the Super Bowl and March (Month+1), many Super Bowl workers, likely part-timers, no longer work for the NFL, resulting in a higher unemployment rate. This is supported by the findings in Figure 16. However, since the difference between Month+3 and the Super Bowl is the lowest, the unemployment rate immensely decreases between March (Month+1) and May (Month+3).

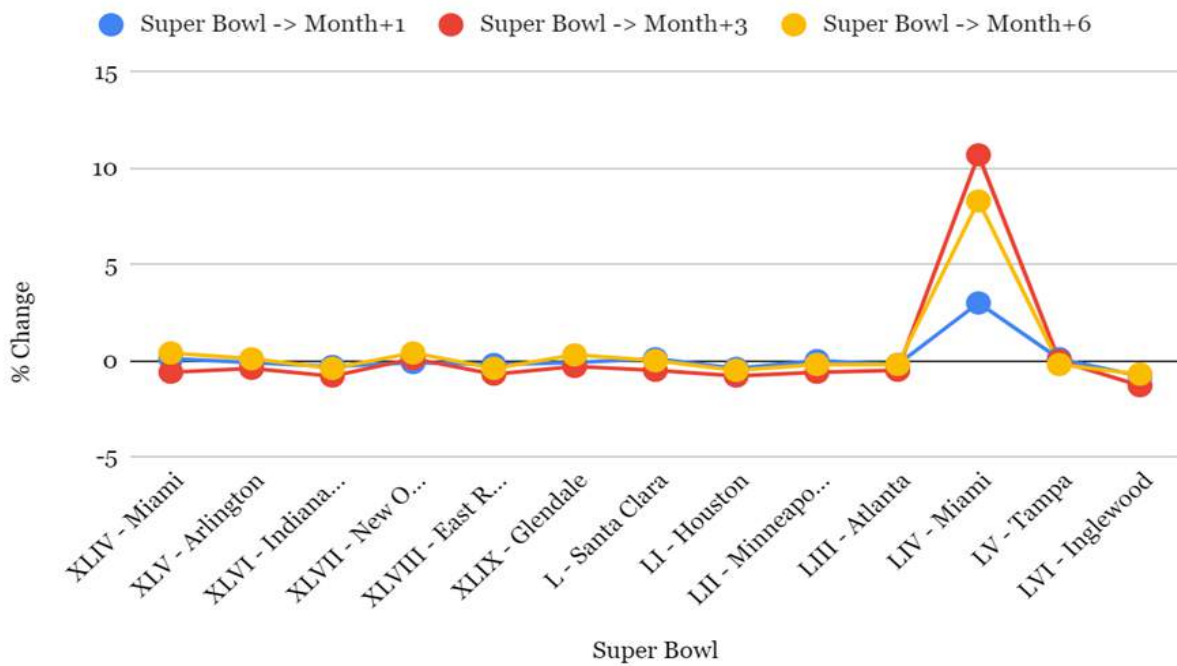


Fig 18: Host City Unemployment Rate After the Super Bowl % Change VS Super Bowl Host City

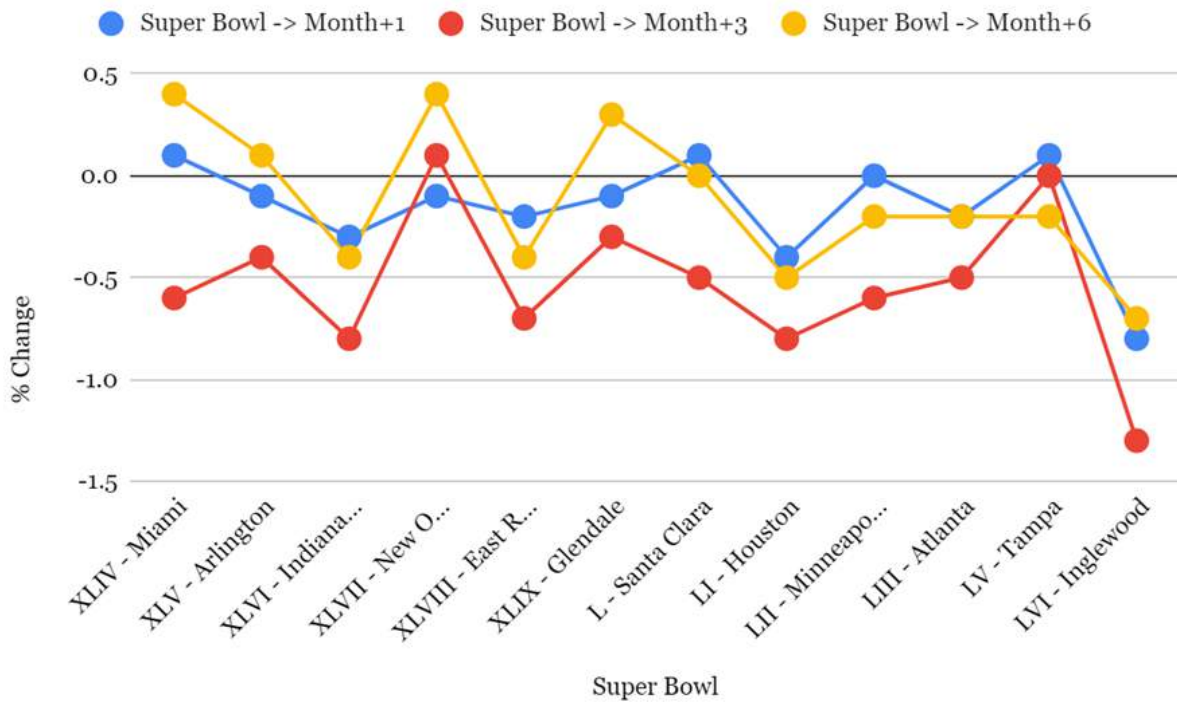


Fig 19: Host City Unemployment Rate After the Super Bowl % Change VS Super Bowl Host City (Excluding Miami)

Conclusion

Although this analysis is relatively sound, there are a few nuances given the complexity of the Super Bowl and the economy. First, the period is confined: five years for population and one year for labor force, number of employed and unemployed people, and unemployment rate. Despite this more useful timespan, it is still possible for the metrics to change as a result of the Super Bowl before or after the time period.

Second, it is possible that other external events impact the metrics under consideration. It is likely that the Super Bowl is the main factor for the metrics' change during the time period, however other events can still skew data points to a degree. One major event that likely skewed results for the recent Super Bowls is COVID-19, which resulted in both a strong downturn and a strong rebound in the labor market.

Third, despite the diverse selection of the thirteen analyzed cities, it is not plausible to generalize the trend and guarantee it for all host cities. Even in cities of similar infrastructure, population, and economies, other factors can still alter the metrics. Lastly, it is a well-known issue in the sports economics community that numbers, such as host city economic impact, are usually inflated for sporting events, including the Super Bowl. This may be for a variety of reasons, such as creating a more beneficial image for the NFL, not considering the displacement effect, and not measuring certain financial losses.

The Super Bowl has a major impact on the nation, the host state, and the host city. It is directly related to the observed changes in population, labor force, number of employed and unemployed people, and unemployment rate.

In terms of population, it is apparent that the Super Bowl happens in growing cities; however due to the large number of other variables that control population size, it is unlikely that the Super Bowl is the main causation for the observed increases.

Regarding the labor force data, prior to the Super Bowl, local labor markets experienced the greatest increase in the one-month period before the Super Bowl. Nonetheless, there is consistent growth as early as six months before the event. Following the Super Bowl, the labor force in most cities returned to "normal" conditions in no more than three months.

Regarding the number of employed and unemployed people, there are observable increases in employment and decreases in unemployment over the one-year time period. Prior to the Super Bowl, the greatest surge in employed people is between November (Month-3) and January (Month-1) while the greatest decrease in unemployed people is between January (Month-1) and the Super Bowl. After the Super Bowl, the number of employed people continued increasing in a region, even for as long as six months. The number of unemployed people increased one month after the Super Bowl, continued to decrease in the following two months, and eventually increased back to "non-Super Bowl" conditions.

Lastly, the lowest unemployment rate prior to the Super Bowl is three months prior and the unemployment rate is at its highest one month following the Super Bowl. With these data sets and analysis of thirteen host cities' labor market reactions to the Super Bowl, it is plausible to create a model to predict labor market reactions in future host cities based on similar city factors.

These factors include but are not limited to, stadium size, proximity to major city, average temperature, and popularity of the city's team affiliation as well as the popularity of the teams playing in the Super Bowl.

Works Cited

- Popke, Michael. "Super Bowl LIII by the Numbers." Sports Destination Management. Accessed December 31, 2023.
<https://www.sportsdestinations.com/management/economics/super-bowl-liii-numbers-15863>.
- College, Lerner. "The Business Side of the Super Bowl." Lerner, February 10, 2023.
<https://lerner.udel.edu/seeing-opportunity/the-business-side-of-the-super-bowl/>.
- Corbett, Peter. "Report: Super Bowl Lifted Valley Economy by \$720 Million." The Arizona Republic, June 24, 2015.
<https://www.azcentral.com/story/news/local/glendale/2015/06/23/super-bowl-valley-economic-impact/29193319/>.
- Darnell, Tim. "Super Bowl to Generate about \$190 Million for Metro Atlanta." Atlanta, GA Patch, January 29, 2019.
<https://patch.com/georgia/atlanta/super-bowl-generate-190-million-spending-atl>.
- Economic impact of super - pr newswire. Accessed December 31, 2023.
https://mma.prnewswire.com/media/516490/Economic_Impact.pdf.
- Heid, Jason. "How Dallas Won the Bid for Super Bowl XLV." D Magazine, October 17, 2022.
<https://www.dmagazine.com/publications/d-magazine/2010/super-bowl/how-dallas-won-the-bid-for-super-bowl-xlv/>
- Homefacts.com. "Get Reliable Neighborhood Information Instantly." Homefacts. Accessed December 31, 2023.
<https://www.homefacts.com/>.
- "The Labor Force." The Labor Force | Workforce & Career Information User's Guide. Accessed December 31, 2023.
<https://www.nhes.nh.gov/elmi/products/ug-labor-force.htm#:~:text=People%20are%20considered%20employed%20if,business%20are%20also%20considered%20employed>.
- Marshall, Beth. "8 Stats to Know about Super Bowl LI's Economic Effect on the Greater Houston Area." Community Impact, May 25, 2017.
<https://communityimpact.com/news/2017/05/25/8-stats-know-super-bowl-lis-economic-effect-greater-houston-area/>.
- Schrotenboer, Brent. "Lessons Learned from Mass Transit Fiasco at Super Bowl." USA Today, February 3, 2014.
<https://www.usatoday.com/story/sports/nfl/super/2014/02/03/super-bowl-xlvi-mass-transit-nfl-metlife-stadium/5184675/>.
- Sports, Front Office, and Nexstar Media Wire. "Super Bowl LVI: New Report Details Possible Economic Impact on L.A." KTLA, October 26, 2021.
<https://ktla.com/sports/super-bowl-lvi-new-report-details-possible-economic-impact-on-l-a/>.
- U.S. Bureau of Labor Statistics, December 22, 2023. <https://www.bls.gov/>.
- "Welcome to Fred, Your Trusted Source for Economic Data since 1991." FRED. Accessed December 31, 2023. <https://fred.stlouisfed.org/>.
- Wilson, Lea. "Houston Super Bowl Host Committee Announces Economic Impact of Super Bowl LI." KPRC, May 25, 2017.
<https://www.click2houston.com/news/2017/05/25/houston-super-bowl-host-committee-announces-economic-impact-of-super-bowl-li/>.
- "World Population." World Population 2023 : Live Updates. Accessed December 31, 2023.
<https://www.populationu.com/>.

Evaluation of Injury Risk Factors in Youth Soccer By Situ Wang

Abstract

This study investigates injury patterns among youth soccer players aged 16-21, focusing on injury types, affected body areas, and positional risks. Conducted between February and May 2024, the survey analyzed injury data to assess frequency, type, and distribution across different player positions and age groups. The results reveal that the knees, ankles, and thighs are the most commonly injured areas, with muscle-related injuries being the most frequent and severe. Forwards have the highest injury rate, while goalkeepers are at the lowest risk. Notably, female players, despite their lower overall number, experience a higher proportion of injuries compared to males. The findings, consistent with previous studies, highlight the critical need for targeted injury prevention strategies, including strength training for vulnerable areas, increased awareness of injury risks, and adaptive coaching practices to mitigate injury risks, particularly for older youth players and female athletes.

Introduction

Soccer's global appeal is unmatched, captivating the attention of over 5 billion fans worldwide [1]. With an estimated 275 million players [3] and nearly 130,000 professionals across more than 4,400 clubs, including men, women and youth [2]. In the United States alone, more than 3 million youth soccer players are registered through various organizations such as the United States Youth Soccer and American Youth Soccer Organization. In Canada, there are approximately 500,000 youth players. It is estimated that there may be more than 50 million youth soccer players active in the sport worldwide. It has become the world's most popular sport [11]. Beyond its popularity, soccer fosters excellent cardiopulmonary health and helps players build strong, well-conditioned physiques [10].

Injuries are inevitable in this highly competitive, high-physical contact sport that involves so many youth. From 1990 to 2014, the rate of soccer-related injuries doubled among U.S. youth ages 7 to 17 [6]. Injuries in soccer are a major concern at all levels, from youth leagues with players aged 10 to 18 to professional teams, which may include young talents as early as 15 [5]. Among professional players, the injury rate is about 6.6 per 1,000 hours of play, with match play significantly higher at 23.8 injuries per 1,000 hours, compared to 3.4 during training [9]. For competitive youth players, injury rates during matches range from 9.5 to 48.7 per 1,000 hours for males and 12.5 to 30.3 for females, with generally lower rates during training [10].

Soccer is a highly confrontational contact sport with a high rate of injury. Injuries caused by soccer exert a significant burden on socioeconomic and healthcare systems. However, there exists a lack of methods to prevent injury specifically for youth soccer players. Therefore, we investigated the relationship between multiple factors that influence youth soccer injuries. This paper analyzes demographics such as the number of players injured, the number of injuries and body parts injured, as well as age and gender. A general evaluation is presented to help prevent and mitigate the risks associated with youth soccer injuries.

Results

Between February 2024 and May 2024, we collected data from three soccer clubs U16 and above through an online survey platform. A total of 368 surveys were distributed and 46 valid responses were collected (valid response rate of 12.5%), 18 (39.1%) of them were female and 28 (60.9%) were male, aged between 16-21 years old (Figure 1A). The survey covered the factors that contributed to the injuries these players experienced while participating in soccer. Eight factors were included such as age, gender, field position, location of injury, type of injury, and severity. The research shows that 16 years old is the age group with the highest number of injuries and that players in the forward position are at the highest risk of injury. And there is a trend of higher number of injuries in higher age groups than lower age groups. Consequently, there is a need to strengthen the measures of injury prevention for young soccer players in high-age groups to avoid the high risk of causing big injuries to them.

Among these 46 participants, 41.3% were goalkeepers, 26.1% were defenders, 15.2% were midfielders and 17.4% were forwards (Figure 1B). The youth soccer clubs usually have two game seasons, one is the summer season which is from May to October, and the other is the winter season from November to April. Each team will have around 9 months of training time. Other times will be the team tryout, preseason, postseason, or break. The training frequency in my survey means the training times during the regular game season. Most of the participants will train either 1-2 or 3-4 times per week. People who train 3-4 times are 2 percent more than people who train 1-2 times. Only 10.9% of people train five or more times per week (Figure 1C). 76.1% of all the participants received education on injury prevention. (Figure 1D).

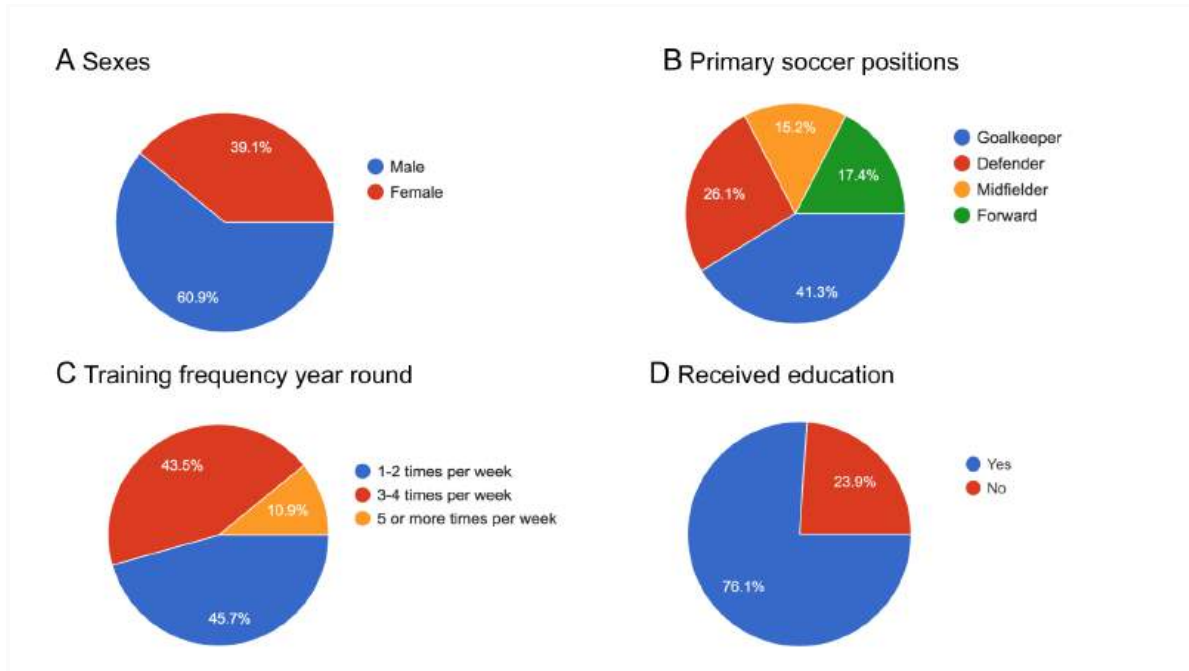


Figure 1. Overall information of 46 youth soccer participants in the survey. (A) Sex of participants. (B) Primary soccer positions on the field. (C) Training frequency year-round. (D) Received injury prevention education.

The Ontario Health Insurance Plan was utilized as a reference to categorize the human body into 14 potential injury locations, including the head/face, shoulders, arms, elbows, hands/wrists, back, chest/abdomen, waist, hips/groin, knees, lower legs, ankles, and feet [13]. Our data from Figure 2A reveal that lower body injuries are most common among participants, with 34.8% experiencing knee injuries, 30.4% suffering ankle injuries, and 28.3% dealing with thigh injuries. The combined categorization shows that there were a total of 67 injuries to the lower extremities (55.37%), 32 to the upper extremities (26.45%), 14 to the trunk (11.57%), and 8 to the head (6.61%) (Figure 3 A). These data indicate that the lower extremity is the area with the highest probability of injury, and in particular, the knee and ankle joint areas are at the highest risk of injury, ranking in the top two places in terms of the number of injuries. Upper body injuries are less frequent, and the results surprisingly show that nobody who was surveyed has ever been through a back injury. The shoulder, waist, and chest also showed a very low injury rate. But there are exceptions, which are the hands with 26% of people and the head with 17% of people. Overall the joints in the lower body are the locations easiest to get injured

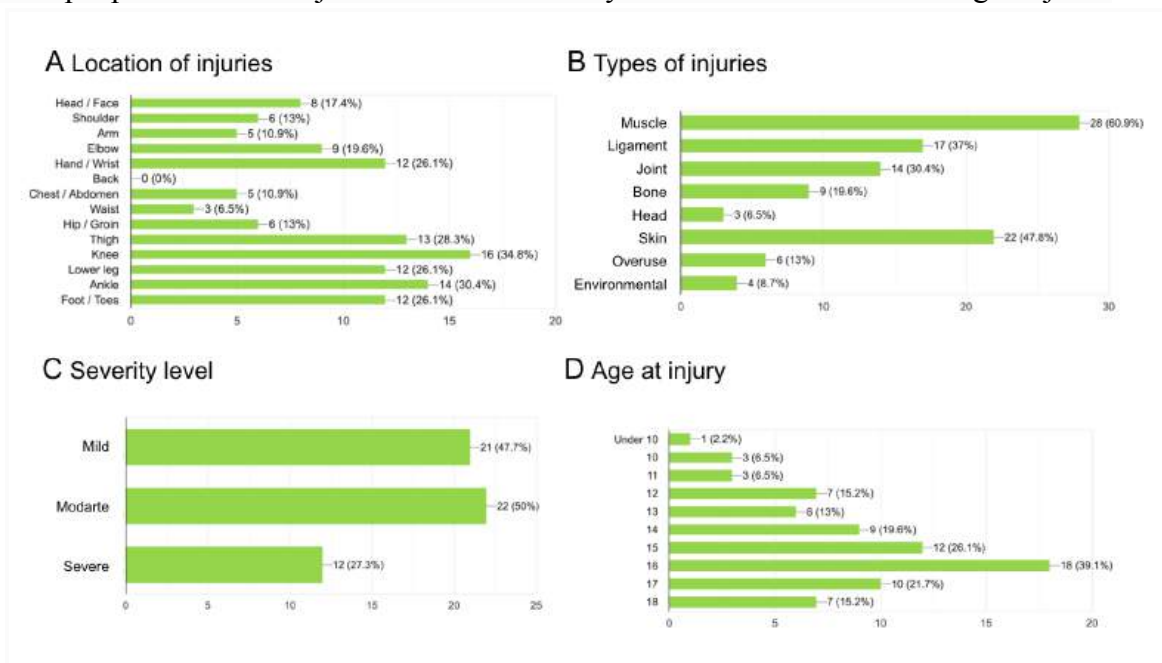


Figure 2. Information on injuries of 46 youth soccer participants who were surveyed. (A) Location of injuries. The survey includes multiple choices in case of multiple injuries. (B) Types of injuries. The survey includes multiple choices in case of multiple injuries. For muscle injuries, the injuries include strains, contusions, bruises, etc. Ligament injuries include sprains, tears etc. Head injuries include concussions and facial injuries. Skin injuries include abrasions and lacerations. Overuse injuries include tendonitis and shin splints. Environmental injuries will include head or cold-related illness. (C) Severity level. Mild requires less than a week off from training. Moderate severity requires 1-4 weeks off. Severe requires more than 5 weeks off. (D) Age at injury

All reported injuries were divided into various types, including muscle injuries (strains, contusions, abrasions), ligament injuries (sprains, tears), joint injuries, bone injuries, head injuries (concussions, facial injuries), skin injuries (abrasions, cuts), overuse injuries (such as

tendinitis, ankle pain), and environmental injuries (like heat-related and cold-related illnesses). The data from Figure 2 B and Figure 3 showed that muscle injuries were the most common, with 28 (60.9%) participants experiencing muscle injuries, accounting for 27.2% of the total injuries; the second most common injury was skin injuries with 22 (47.8%) accounting for 21.4% of the total injuries; and the third most common injury was ligament injuries with 17 (37%) accounting for 16.5% of the total injuries. In contrast, head injuries were relatively rare, with only 8 (17.4%) experiencing injuries, accounting for 2.9% of the total number of injuries.

Figure 2 C categorizes injuries into three severity levels: mild (requiring under one week to recover), moderate (requiring 1- 4 weeks), and severe (requiring over 5 weeks and possibly surgery). The data shows that 50% of participants needed 1- 4 weeks to recover, 47.7% took less than a week, and 26% required 5 weeks or more for full recovery. Most participants had mild and moderate injuries.

Figure 2 D illustrates injury rates across different ages. The data indicates that injury rates are relatively low at ages 10 and 11, but they begin to rise steadily from age 12, peaking significantly at age 16, 18 (39.1%) of participants had been injured. After 16, the injury rate decreases rapidly. Since the survey targeted participants aged 16 to 21, Some of them were not yet 17 or 18 years old. So the data for ages under 16 provides valuable insights into injury trends among younger players.

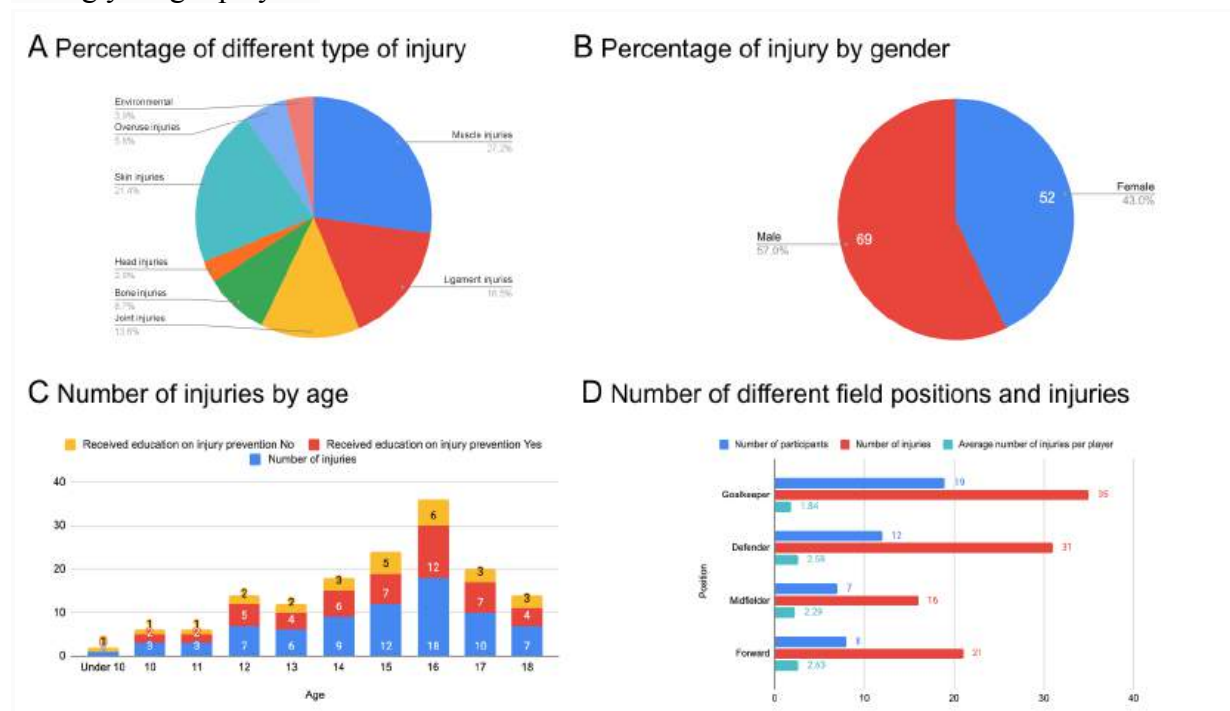


Figure 3. Bar graph and Pie chart mix information (A)Percentage of different types of injury. Percentage of all injuries by type of injury, based on survey data. (B) Percentage of injury by gender. Based on survey data, the percentage of injuries to players of different genders to all injuries. (C) Number of injured players by age and whether or not they received injury prevention education. (D)Number of players at different field positions and number of injuries. And the average number of injuries per player at that position.

Figure 4 analyzes the injuries and illnesses of players in different positions by age group. It was found that the number of injury occurrences for players in all positions showed an increasing trend with age, with players around 15-18 years of age being at the greatest risk of injury.

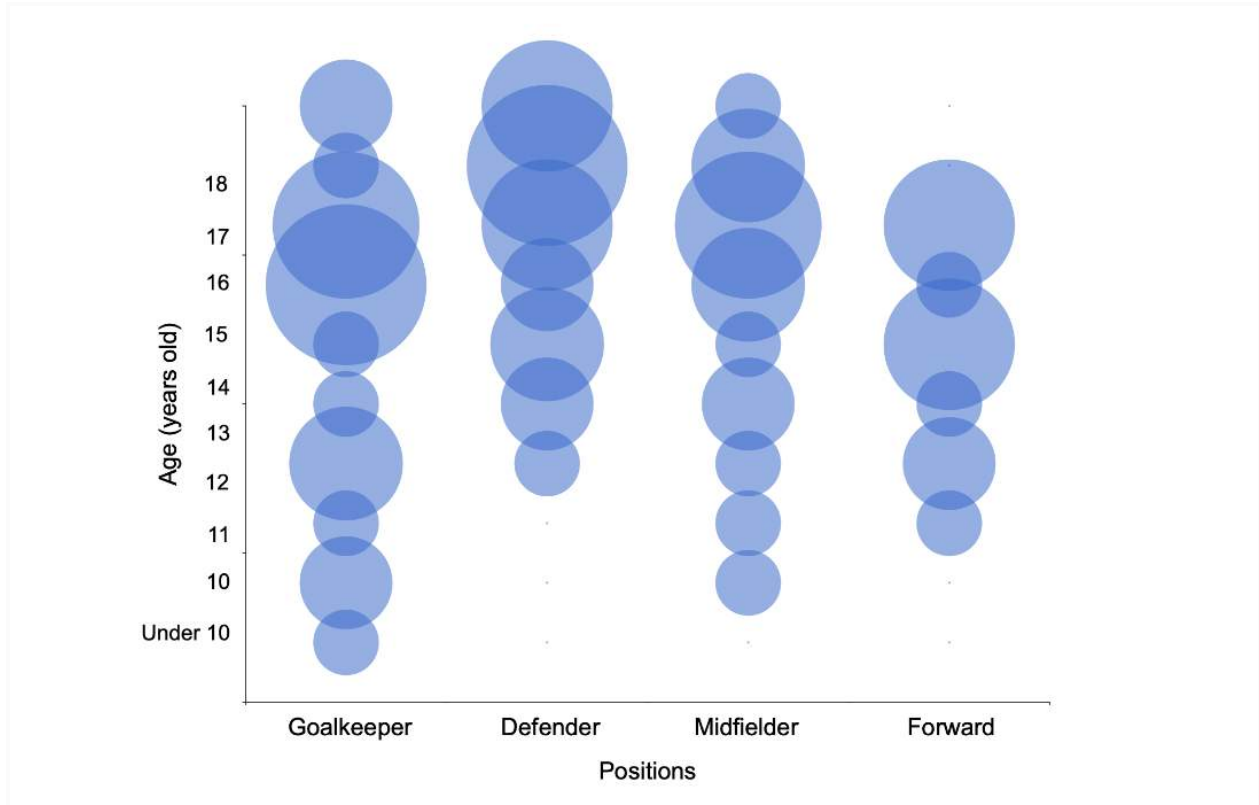


Figure 4. Distribution of players participating in the survey by age group at different positions

Discussion

Based on the data and graphs, we were able to conclude that the lower extremities are the areas with the highest probability of injury. The knees, ankles and thighs are the most commonly injured parts of the body for youth soccer players. This is consistent with the results of a previous study [8][10]. Additionally, the injury type chart reveals that muscle-related injuries are the most frequent, affecting 27.2% of players, followed by skin-related injuries at 21.4%. Among soccer players, muscle injuries are generally more severe and require longer recovery times than skin injuries. Therefore muscle injuries in the lower extremities are the most significant concern for youth soccer players and need to be paid the most attention.

The survey data showed that 16 had the highest number of injuries, with 39.1% of all players experiencing an injury at age 16. The trend in the number of injuries also showed a year-over-year increase in age (Figure 2 D, Figure 4).

After comparing the risk of injury for players at different positions on the field and found that the results were not consistent with our initial hypothesis. The data indicates that forwards have the highest injury rate, averaging 2.63 injuries per player. Defenders follow closely with an

average of 2.58 injuries per player. Interestingly, goalkeepers are at the lowest risk, with only 1.84 injuries per player (Figure 3 D).

In this survey, 39.1 % of participants are female, which is relatively less than males but 43% percent of the total injuries are female. The percentage of female injuries is higher than the percentage of participants, indicating that youth female soccer players are more likely to be injured than males (Figure 1 A, Figure 3 B).

To effectively prevent injuries among youth soccer players, a comprehensive approach is essential, with strength training as a key component. Building strength, particularly in high-risk areas of the lower extremities such as the knees and ankles, which have shown in the results to be the most injured area, is crucial for reducing the injury rate. This involves not only developing the muscles around these vulnerable areas but also improving overall stability and balance through targeted exercises.

Injury prevention also requires a strong awareness from both players and coaches. Players need to understand the importance of not pushing themselves beyond their limits during practice and recognizing when their bodies need rest. They should also be educated on proper techniques for self-care, even when they are off the field, such as maintaining flexibility, staying hydrated, and getting adequate sleep. Coaches play a vital role in injury prevention by ensuring that players have sufficient rest between and during games. They should incorporate recovery sessions after intense match days to help players heal and rejuvenate. Additionally, coaches must be adaptable, adjusting training plans based on weather conditions, field quality, and the physical status of their players to minimize injury risks.

Conclusion

In conclusion, the research study, conducted between February and May 2024, revealed key insights into the injury patterns of youth soccer players aged 16-21. The data showed that the knee, ankle, and thigh are the most frequently injured areas, with muscle-related injuries being the most common and severe, often requiring longer recovery times than other types of injuries. Forwards were identified as the position with the highest injury risk, while goalkeepers had the lowest. Additionally, the data suggested that female players, although fewer in number, are more prone to injuries than their male participants. The findings emphasize the need for targeted injury prevention strategies, particularly for older youth players, to reduce the risk of significant injuries that could impact their playing careers.

Works Cited

- [1] The football landscape – the vision 2020-2023. FIFA Publications. (n.d.).
<https://publications.fifa.com/en/vision-report-2021/the-football-landscape/>
- [2] Around FIFA Professional football. FIFA Publications. (n.d.-a).
<https://publications.fifa.com/en/annual-report-2021/around-fifa/professional-football-2021/>
- [3] Smith, C. (2023, September 18). How many soccer players in the world?. Soccerprime.
<https://soccerprime.com/how-many-soccer-players-in-the-world/>
- [4] Physical demands on the soccer player. Human Kinetics. (n.d.).
<https://us.humankinetics.com/blogs/excerpt/physical-demands-on-the-soccer-player>
- [5] Dragone, A. (2022, February 15). Youth soccer: All you need to know about it and the best drills to train young players. YouCoach.
<https://www.youcoach.com/article/youth-soccer#:~:text=Youth%20soccer%20refers%20to%20various,19%2Dyear%2Dold%20players.>
- [6] Esquivel, A. O., Bruder, A., Ratkowiak, K., & Lemos, S. E. (2015, July). Soccer-related injuries in children and adults aged 5 to 49 years in US emergency departments from 2000 to 2012. *Sports health*.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4481679/>
- [7] Khodaei, M., Currie, D. W., Asif, I. M., & Comstock, R. D. (2017, February 1). Nine-year study of US High School soccer injuries: Data from a national sports injury surveillance programme. *British Journal of Sports Medicine*.
<https://bjsm.bmj.com/content/51/3/185>
- [8] Research, T. (2023, March 21). Red card: Soccer injuries by the numbers.
<https://www.truvelo.com/blog/research/soccer-injuries-2023/>
- [9] Ekstrand, J., Spreco, A., Bengtsson, H., & Bahr, R. (2021, October 1). Injury rates decreased in men's professional football: An 18-year prospective cohort study of almost 12 000 injuries sustained during 1.8 million hours of play. *British Journal of Sports Medicine*.
<https://bjsm.bmj.com/content/55/19/1084>
- [10] Owuoye, O. B. A., VanderWey, M. J., & Pike, I. (2020, September 21). Reducing injuries in soccer (football): An Umbrella Review of best evidence across the Epidemiological Framework for prevention - sports medicine - open. SpringerOpen.
<https://sportsmedicine-open.springeropen.com/articles/10.1186/s40798-020-00274-7>
- [11] FIFA-Professional-Football-Report-2023. (n.d.).
<https://digitalhub.fifa.com/m/2a5dc95026d9cf8a/original/FIFA-Professional-Football-Report-2023.pdf>
- [12] Andrew Watson; Jeffrey M. Mjaanes; Cynthia R. LaBella; Greg Canty; Alex B. Diamond; William Hennrikus; Kelsey Logan; Kody Moffatt; Blaise A. Nemeth; K. Brooke Pengel; Andrew R. Peterson; Paul R. Stricker (November 01 2019). Soccer Injuries in Children and Adolescents - The American Academy of Pediatrics| Clinical Report| AAP Publications
<https://publications.aap.org/pediatrics/article/144/5/e20192759/38190/Soccer-Injuries-in-Children-and-Adolescents?autologincheck=redirected>
- [13] Giza E, Micheli LJ. *Soccer injuries*. *Med Sport Sci*. 2005;49:140-169. doi: 10.1159/000085395. PMID: 16247265.
<https://pubmed.ncbi.nlm.nih.gov/16247265/>
- [14] Diagnostic Codes - Ontario Health Insurance Plan. (n.d.).
<https://files.ontario.ca/moh-diagnostic-codes-en-2023-09-01.pdf>

The Marriage Penalty: Myra Bradwell's Fight for Married Women in the Law **By Julia Nathanson**

On January 18, 1873,^[i] lawyer and Wisconsin Senator Matthew Hale Carpenter^[ii] stood before the United States Supreme Court hoping to change the future of women's rights in America forever. At stake was the case of Illinois resident Myra Bradwell, a woman who simply wanted to practice the profession for which she had, with great determination, studied and prepared: the law. Despite her rigorous education in the field and her passing score on the bar exam,^[iii] Illinois denied her entry to the profession on account of her being not simply a woman, but a *married* woman.^[iv] Carpenter argued that all women, including married women, should have the right to be admitted to the bar. He failed.^[v]

How could the Supreme Court have denied an educated and qualified woman the right to practice law simply because she was married? Chief Justice Joseph P. Bradwell's majority opinion helps answer this question. He wrote: "The paramount destiny and mission of women is to fulfill the noble and benign offices of wife and mother. This is the law of the creator."^[vi] In Illinois, raising children as a wife and mother was deemed antithetical to the practice of law. In essence, with the *Bradwell v. Illinois* ruling, Myra Bradwell was told she would never be able to do both.

Bradwell decided to fight. She went on both to pursue her own career ambitions and to change the landscape of married women's rights forever. Understanding Bradwell's plight requires a closer look both at her own early life and experiences and at the social and cultural context in which she lived.

Myra Colby was raised in the Antebellum era, the mid-nineteenth century period of rising political tension between North and South. The daughter of Eben and Abigail Colby,^[vii] Myra grew up in a staunchly abolitionist household,^[viii] and although progressive in this manner, the Colby family firmly maintained traditional nineteenth century gender roles; her father ran the farm, and her mother ran the household.^[ix] The family even sent young Myra to a Ladies' Seminary in Elgin Illinois,^[x] instead of to university to reinforce conventional gender roles. Her family valued an education, but one that would lead her to proper motherhood, not one that would lead to acquiring knowledge.

Myra's familial structure was typical for the Antebellum period. The nineteenth century saw separate spheres for men and women: husbands should belong to the professional sphere, wives to the domestic sphere.^[xi] Men could receive a higher education, while women, like Myra, learned how to be proper (and pious) wives and mothers at Ladies' Seminaries. Having a more advanced education and being the sole income provider gave husbands enormous power over their wives, and this unequal power dynamic often forced the wife into acceptance of unearned inferiority.^[xii]

Married women were not only rendered socially inferior to their husbands, they were deemed legally inferior, as well. The U.S. legal system, based largely upon English Common Law,^[xiii] asserted that, upon marriage, a woman's identity was legally merged with her husband's.

This principle, known as “unity of person” or “coverture,”^[xiv] was rooted in the notion that women were too intellectually and emotionally fragile to make rational decisions regarding money, property, and politics. Coverture restricted the rights of married women to the point where every aspect of a woman’s legal existence was subsumed by her husband.^[xv] A married woman could not own property, could not sign a contract without her husband’s signature, and could not create a will.^[xvi] Coverture also gave a husband near absolute legal power over his wife in the case of divorce. If a couple were to separate, the husband would automatically gain full legal custody of the children.^[xvii] Furthermore, according to the doctrine of coverture, a wife was not able to sue or be sued, nor could she testify in court.^[xviii] Even though this was a direct violation of the Compulsory Process Clause of the sixth amendment of the Constitution,^[xix] created in the Bill of Rights, no one challenged coverture.

Interestingly, though, a single woman faced none of these legal restrictions. Her property, contractual, occupational and other related rights were still in place.^[xx] The day she married, these rights were stripped from her. Coverture required full commitment to her new familial obligation to bear and raise children.

Ironically, part of Myra’s escape from the domestic sphere was due to her marriage. Bradwell was only able to obtain the necessary education to practice law because of her husband. In 1852, she married James Bradwell,^[xxi] the son of immigrants who had grown up poor on the prairie.^[xxii] Because of his hardscrabble upbringing where both parents had to work to support the family, he was not attached to the “separate spheres” belief. Instead of the conventional unequal power dynamic, Myra and James formed a partnership of sorts. James was a lawyer, and he educated Myra about the law. She placed herself in an intense apprenticeship with him, dreaming that one day they could work side by side as equals. In 1861, he became the county judge for Cook County, Illinois,^[xxiii] while Myra was pregnant with their fourth child. At this point, after almost a decade both of legal education and of raising children, Myra decided to challenge the law. She now saw firsthand that she could indeed conduct business and act as a wife and mother at the same time. She knew better, so she was ready to do better.

Recognizing the cruel irony of coverture, equipped with her new legal expertise, and with the American idea of “freedom” in question with the onset of the Civil War, Myra Bradwell decided to tackle the long overdue need for freedoms for married women. In 1861, she helped write the Illinois Married Women’s Property Act,^[xxiv] passed by the Illinois State Legislature, it allowed married women the right to own and control property.^[xxv] While this act was monumental in combating one of the long-standing laws of coverture, its impact was limited. Myra Bradwell would soon find out that change occurred incrementally, not all at once. Her initial efforts were not enough to change society’s belief in the separate spheres of men and women. Legal reforms that granted women greater rights would require further legal challenges. Bradwell was ready to mount them; she only needed to be admitted to the Bar.

On August 2, 1869,^[xxvi] Myra Bradwell became the first woman in Illinois to pass the state Bar Exam. Although this proved her to be a qualified lawyer, her longtime dream ended when the Illinois Supreme Court promptly denied her application to practice law in Illinois. The

rejection came not because she was a woman, but on the grounds that she was a *married* woman. She was told she could not practice law “by reason of the disability imposed by your marriage.”^[xxvii] According to the doctrine of coverture, entering into marriage meant assuming new roles within the domestic sphere which precluded her from handling the practice of law. The Illinois Supreme Court decreed that Myra’s ability to run the household would be hindered and disabled if she simultaneously occupied a profession. Myra recognized the logical flaws of this argument and decided to push all the way to the Supreme Court.

Bradwell believed this decision disregarded her fourteenth amendment rights and thus she appealed all the way to the United States Supreme Court. She hired Wisconsin Senator Matthew Hale Carpenter to represent her^[xxviii] in front of the nine male justices. Carpenter argued that the Illinois Supreme Court decision to prohibit Bradwell from practicing the law was a violation of her Fourteenth Amendment rights. He cited the newly ratified 1868 amendment’s “Privileges and Immunities” clause, which reads, “No state shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States...”^[xxix] The “citizens” to which the fourteenth amendment refers are former slaves. Carpenter cleverly reasoned that since Bradwell was a citizen, regardless that she was a married woman, she also had the right to practice her profession of choice. Not only did Carpenter equate married women with slaves in their lack of freedoms, but he used language supporting freed slaves’ rights to argue that Bradwell should have the freedom to practice law. He posited that as a citizen of the United States, Bradwell had the constitutional right to practice any profession, and this immunity could not be diminished by the state, regardless of her marital status. The Supreme Court did not see it that way, and thus ruled against Myra eight to one.^[xxx]

Justice Joseph P. Bradley wrote the majority opinion on *Bradwell v. The State of Illinois* (1873).^[xxxi] In his opinion, he did not justify the decision based on the legal status of married women under coverture, but instead used the subjective belief in the “separate spheres” of husband and wife to legitimize the conclusion of the court. He wrote, “The harmony, not to say identity of interests and views which belong, or should belong, to the family institution is repugnant to the idea of a woman adopting a distinct and independent career from that of her husband.”^[xxxii] Bradley insisted that a married woman’s primary role was to maintain the peace of the household and care for her husband, insinuating that she was too delicate to handle the occupations of civil life at the same time. Any disturbance to this well-established order, according to Bradley, would damage the fabric of society. Not only was it then disruptive for Myra to pursue a career, it was *selfish*. Bradley concluded by articulating, “The paramount destiny and mission of women are to fulfill the noble and benign offices of wife and mother. This is the law of the creator.”^[xxxiii] By connecting a woman’s role in society to religion, Bradley created a practically indisputable argument. Bradley elevated the cultural phenomenon of separate spheres into a religious legal standard, thereby forcing all American married women to obey the guidelines put forth by the creator. Even in the United States, a country founded upon a belief in the separation of church and state and with non-religious government institutions, the law of God still triumphed in the court.

Although defeated, Myra Bradwell did not give up. If she could not practice the law, she would change the law. Her first step in this new mission was to persuade lawmakers and civilians to change their thinking regarding gender roles.

In 1868, a year before the Illinois Supreme Court denied Myra the ability to practice law, she set out on the ambitious endeavor of printing the *Chicago Legal News*.^[xxxiv] In a world where women were deemed incapable of working outside of the home, it would seem inconceivable that she would be able to start her own business. In order to do so, Myra deftly used her privilege—the backing of her well-known husband and her wealthy background—to gain a special charter from the Illinois State Legislature.^[xxxv] This unique charter contravened many of the legal implications of coverture. Bradwell could run her business as a married woman, sign contracts, and keep her own earnings. Knowing that coverture *could* be overturned, at least at the state level, empowered Myra to use the platform of the *Chicago Legal News* to criticize the United States Supreme Court, a federal body, for its antiquated views on gender roles and the conventions of coverture.

Using the power of her newly established newspaper, Bradwell now had a platform for advocacy, progress and reform. She cultivated relationships with Illinois lawmakers and earned the exclusive privilege of being the first paper to report on and publish legislative decisions as they occurred. This established the *Chicago Legal News* as a reliable, valid and timely source of information; its popularity and readership rose quickly as a result. It soon became evident that no Illinois lawyer could effectively practice law in Illinois without reading the *Chicago Legal News*.^[xxxvi] Once establishing herself in Illinois, Bradwell made similar arrangements with the United States Supreme Court. All national judicial decisions from all court levels were immediately reported to her. Thus, the *News* earned the distinction that it printed more federal court decisions faster than any other legal newspaper in the United States.

With a large following of devoted readers in hand, Myra set out to change their views on the role of married women in American society. She repeatedly used the *Chicago Legal News* to advocate for the right of married women to keep their professional earnings. In the eighth issue, for example, Bradwell presented an impassioned argument in favor of married women's rights based on the case of *Schwartz v. Sanders*. Mrs. Schwartz operated her own business and, as Bradwell wrote, “worked like a slave.”^[xxxvii] Bradwell's use of this simile was very purposeful. The article was written less than a decade after the abolition of slavery, while the evils of the practice and its societal reverberations were still prevalent. To compare Mrs. Schwartz to a slave was to say that she worked brutally hard for someone else's benefit: her husband's. The use of this rhetorical strategy highlighted the notion that if a woman could work—and in this case, even own her own business—and still be unable to control her own money, then she was no freer than a slave. These words held great power at this time in history, in the wake of the abolitionist movement. Bradwell's comparison between married women and slaves forced people to face the social reforms that still needed to be implemented, both for the benefit of Black Americans and for married women. Sadly, in the legal case of *Schwartz v. Sanders*, the Supreme Court of Illinois upheld that women could not keep their earnings.^[xxxviii]

Undefeated by the decision in *Schwartz v. Sanders*, Myra drafted the Married Women's Earnings Act and, using the *Chicago Legal News*, advocated for its passing in Illinois Congress. Her combined effort of persuasive rhetoric and legal lobbying led to the passage of the Married Women's Earnings Act of 1869^[xxxix] which allowed married women to keep their own earnings.

Coverture was another societal convention that restricted married women's freedoms that Bradwell sought to dismantle through the *Chicago Legal News*. As early as the third edition, Myra used her editorials to criticize this doctrine. She knew she needed to convince lawmakers, above all else, that coverture was not, in fact, "the law of the creator." She wrote, "But one thing we do claim--that woman has a right to think and act as an individual--believing if the great Father had intended it to be otherwise--he would have placed Eve in a cage and given Adam the key."^[xli] Bradwell's reasoning illustrated that God put Adam and Even on earth as equals. She reminded readers that they all came from the same God, who sees all as equal. By treating women as inferior, society was violating God's wishes. Ultimately, Bradwell argued that if God saw women as equal to men, then America needed to view women as equal, too. Bradwell knew that her Christian readership would respond favorably to her reasoning, even if the law ostensibly upheld a separation of church and state. Thus, with this argument, Myra successfully broke down the notion that it was "the law of the creator" for married women to remain at home if that rendered them inferior. In the eyes of God, all men and women, even married women, were equal.

Bradwell repeatedly used her editorials in the *News* to dismantle the doctrine of coverture. A further example occurred in her detailed critique of the 1861 Married Women's Property Act. Although she was the key proponent of its enactment, and commended its accomplishments--namely, that a married woman in Illinois could now hold property in her own name--this did not stop her from understanding its shortcomings and advocating for further progress for married women. The Act, Bradwell argued, stopped short of ensuring a wife's true economic autonomy, as a husband was still entitled to all of his wife's earnings. She wrote, "If a married woman is allowed to retain her real and personal property the same as a single woman, should not the act of 1861 be so amended as to embrace her earnings for labor performed outside her own household?"^[xlii] Her question underscored the hypocrisies of coverture. She recognized that there always seemed to be loopholes to avoid granting married women the rights that single women enjoyed, whether by passing bills claiming major advances but in reality only achieving minor feats, or by justifying inequality using religious jargon, as was done in her Supreme Court case. She wanted her readers to notice this, too. This argument was additionally successful in aiding the passage of the Married Women's Earnings Act in 1869.^[xliii] Using the paper to point out the absurdity of coverture led to legislative change in favor of married women.

Bradwell promoted other legislation in the *News*, as well. In 1873, she drafted a successful bill which permitted women to be elected for school office.^[xliv] In 1875, she created and lobbied for a bill which made all married women eligible to become a notary public.^[xlv] Both bills gave unprecedented freedoms to married women and would not have been possible if not for their promotion in the *News*. Myra's husband, James, further used his position as a member

of the Illinois House of Representatives to advocate for married women. Without the help of her husband, not even the most powerful woman in the legal world, someone whom all lawyers looked to for information, would have been able to push her political agenda.

With a record of early success behind her, Myra finally confronted the issue of married women and the practice of law in the nineteenth edition of the second volume of the *Chicago Legal News*. She began her argument by defining the word “attorney,” which, according to Miriam Webster, was “one who takes the turn or place of another.”^[xlv] In doing this, Bradwell set out to illustrate that the word itself implied no restrictions on who could become an attorney. She then argued that everyone, no matter their role in society, at some point is an agent to or takes the place of another; therefore, everyone is technically an attorney. This marked her first rebuttal to the Supreme Court; the word “attorney” did not contain any gender restrictions or limitations regarding who could fulfill that role. Therefore, according to this definition, even a mother could act as an attorney. By taking the place of another, including a mother advocating for her children, all people--wives and mothers included--were, in a sense, attorneys.

Bradwell strengthened this argument by citing numerous Illinois Supreme Court cases regarding the rights of married women. She asserted that since under the Married Women’s Property Act of 1861 a married woman could hire an agent to employ her business, and under the Married Women’s Earnings Act of 1869 a wife could keep her own earnings and sue in her own name,^[xlvi] she technically had the right to be an agent, or an attorney, to someone else. Plus, women were increasingly gaining access to education at an equal level as men. Therefore, as they searched for jobs, they were no less qualified than men in the same field.

Bradwell additionally utilized a logical approach to persuade lawmakers, quoting the sixth article of the United States constitution: “The citizens of each State shall be entitled to all privileges and immunities of citizens in the several States.”^[xlvii] According to this, married women were guaranteed the same privileges and immunities as any citizen in Illinois. Bradwell asserted, “Under this section of the Constitution she has a right to receive a license to practice law upon the same terms and conditions as the most favored citizen of the State of Illinois.”^[xlviii] In the *Chicago Legal News*, Bradwell put forth the idea that legal precedent and even the U.S. Constitution had already established that married women, including Bradwell, should be treated equally and granted the right to practice law.

This argument energized 18-year-old Alta May Hulett to partner with Bradwell on this issue. In 1870^[xlix], Hulett had passed the Illinois bar exam only to face a similar rejection as Myra from the Illinois Supreme Court. Inspired by the *Chicago Legal News*, Hulett worked with Myra to draft a new and progressive bill. They first argued that women, as human beings, had the right to be attorneys. More controversially, the pair argued in the bill that women had the same ability and intellectual capacity as men and therefore could practice law on an equal level. The bill passed in the Illinois state legislature on March 22, 1872,^[l] and became the very first^[li] piece of legislation in the nation to require equal hiring practices between men and women. In 1873, Alta Hulett became the first female lawyer in Illinois.^[lii]

Bradwell's words had had a profound impact on state legislatures across Illinois. They became more sympathetic to married women's rights, which allowed the passage of Alta Hulett's bill in 1872.^[liii] Although the 1872 law did not specify rights for married women, it did, along with the Married Women's Property and Earnings Acts, grant new freedoms to married women never before imagined. Bradwell had made incredible strides in breaking down the walls of coverture that restricted married women's rights.

Bradwell also succeeded in changing the minds of the same men who prohibited her from the practice of law in the first place. Although it took time, Bradwell's words and persuasive arguments ultimately influenced lawyers and judges all across the country. In 1879 married women were finally granted the ability to practice law in any federal court in the United States.^[liv] In 1890,^[lv] Myra's application to the Illinois bar was reexamined, and she officially became a lawyer in the state of Illinois. In 1892, she was admitted to practice before the U.S. Supreme Court.^[lvi]

At the helm of the *Chicago Legal News*, Bradwell acted as a pioneer for all married women looking to practice law in the United States. Although the absolute equality of married women and the dissolution of the domestic sphere were not achieved during her lifetime (she died in 1894 from cancer^[lviii]), her accomplishments allowed for later women's rights activists to make further progress. She saw how paternalistic and out-of-date the decision of the Supreme Court justices was, and before she died, she ensured that all of Illinois saw it, too.

Works Cited

- Adams, Anne Brown. "A Spotlight on a Primary Source: The Struggle for Married Women's Rights, Circa 1880s." *The Struggle for Married Women's Rights, circa 1880s* | Gilder Lehrman Institute of American History. www.gilderlehrman.org/history-resources/spotlight-primary-source/struggle-married-womens-rights-circa-1880s.
- "Amdt 6.5.4 Right to Compulsory Process - Constitution Annotated." *Constitution Annotated*. Accessed April 26, 2024. constitution.congress.gov/browse/essay/amdt6-5-4/ALDE_00013333.
- "Coverture." *Wikipedia*. Wikimedia Foundation, February 19, 2023. <https://en.wikipedia.org/wiki/Coverture>.
- "Development of Rights of Women." *Foundations of Law*. <https://lawshelf.com/coursewarecontentview/development-of-rights-of-women>.
- Friedman, Jane M. *America's First Woman Lawyer: The Biography of Myra Bradwell*. Prometheus Books, 2019.
- Harvard Mirador Viewer. [iif.lib.harvard.edu/manifests/view/drs:457607836\\$281i](http://iif.lib.harvard.edu/manifests/view/drs:457607836$281i).
- HeinOnline. <https://home.heinonline.org/>.
- "Human and Civil Rights." *Illinois State Bar Association*. www.isba.org/sections/humanrights/newsletter/2013/01/celebrating140yearsoffemaleattorney#:~:text=Happily%2C%20while%20the%20U.S.%20Supreme,the%20right%20to%20practice%20law
- "In Re Lady Lawyers: The Rise of Women Attorneys and the Supreme Court." *Supreme Court of the United States*. www.supremecourt.gov/visiting/exhibitions/LadyLawyers/section1.aspx#LadyLawyers1.
- "Married Women's Property Acts in the United States." *Wikipedia*. Wikimedia Foundation, https://en.wikipedia.org/wiki/Married_Women%27s_Property_Acts_in_the_United_States.
- "Myra Bradwell." *The Glinda Factor*. <https://theglindafactor.com/myra-bradwell/>.
- "Research Guides: American Women: Resources from the Law Library: State Law Resources." *State Law Resources - American Women: Resources from the Law Library - Research Guides at Library of Congress*. <https://guides.loc.gov/american-women-law/state-laws>.
- State of Illinois Office of the Illinois Courts. "Details: State of Illinois Office of the Illinois Courts." *Administrative Office of the Illinois Courts*, February 27, 2024. www.illinoiscourts.gov/News/1331/Illinois-Supreme-Court-history-Myra-Bradwell/news-detail/#:~:text=In%201861%2C%20the%20Illinois%20legislature,to%20earn%20money%20for%20themselves.
- STPL Reference Causeway. "Women in History- Myra Bradwell." *St Tammany Parish Library*, October 28, 2019. www.sttammanylibrary.org/blogs/post/women-in-history-myra-bradwell/#:~:text=To%20support%20women's%20suffrage%20and,control%20their%20earnings%20and%20property.
- "The Catalyst." *Illinois State Bar Association*. www.isba.org/committees/women/newsletter/2004/02/myracolbybradwellillinoisfirstwoman.

"True Womanhood."

https://web-clear.unt.edu/course_projects/HIST2610/content/03_Unit_Three/11_lesson_e leven/06_true_womanhood.htm.

"Women's History in America." WIC. www.wic.org/misc/history.htm.

"Women's Rights - Domestic Relations in the Nineteenth Century." Domestic Relations In The Nineteenth Century - Husband, Property, Children, and Custody - JRank Articles. <https://law.jrank.org/pages/11370/Women-s-Rights-Domestic-Relations-in-Nineteenth-Century.html>.

"Women's Rights and Their Money: A Timeline from Cleopatra to Lilly Ledbetter." The Guardian. Guardian News and Media, August 11, 2014. <https://www.theguardian.com/money/us-money-blog/2014/aug/11/women-rights-money-timeline-history>.

"Women in World History: A Biographical Encyclopedia." Encyclopedia.com. www.encyclopedia.com/women/encyclopedias-almanacs-transcripts-and-maps/bradwell-myra-1831-1894.

ISBA. "The Remarkable Life and Times of Alta May Hulett." ISBA. <https://www.isba.org/committees/women/newsletter/2004/02/theremarkablelifeandtimesofaltamayh>.

The Law Dictionary. "Who Was Myra Bradwell, America's First Woman Lawyer?" The Law Dictionary, <https://thelawdictionary.org/article/who-was-myra-bradwell-americas-first-woman-lawyer/>.

Oyez. "Slaughterhouse Cases." Oyez. <https://www.oyez.org/cases/1850-1900/83us130>.

Constitution Center. "Bradwell v. Illinois (1873)." Constitution Center. <https://constitutioncenter.org/the-constitution/supreme-court-case-library/bradwell-v-illinois-1873>.

"On This Day, Women First Allowed to Argue Supreme Court Cases." National Constitution Center – constitutioncenter.org. <https://constitutioncenter.org/blog/on-this-day-women-first-allowed-to-argue-supreme-court-cases>.

[i] Oyez. "Slaughterhouse Cases." Oyez. <https://www.oyez.org/cases/1850-1900/83us130>.

[ii] Friedman, Jane M. *America's First Woman Lawyer: The Biography of Myra Bradwell*. Prometheus Books, 2019.

[iii] "The Catalyst." Illinois State Bar Association. www.isba.org/committees/women/newsletter/2004/02/myracolbybradwellillinoisfirstwoman.

[iv] Ibid.

[v] Constitution Center. "Bradwell v. Illinois (1873)." Constitution Center. <https://constitutioncenter.org/the-constitution/supreme-court-case-library/bradwell-v-illinois-1873>

[vi] Ibid.

[vii] Harvard Mirador Viewer. [iiif.lib.harvard.edu/manifests/view/drs:457607836\\$281i](http://iiif.lib.harvard.edu/manifests/view/drs:457607836$281i), 223.

[viii] Ibid., 223.

[ix] Ibid., 223.

[x] Ibid., 223.

[xi] "True Womanhood."

https://web-clear.unt.edu/course_projects/HIST2610/content/03_Unit_Three/11_lesson_eleven/06_true_womanhood.htm.

[xii] Ibid.

[xiii] "Coverture." Wikipedia. Wikimedia Foundation, February 19, 2023.

<https://en.wikipedia.org/wiki/Coverture>.

[xiv] "Development of Rights of Women." Foundations of Law.

<https://lawshelf.com/coursewarecontentview/development-of-rights-of-women>.

[xv] "Coverture." Wikipedia. Wikimedia Foundation, February 19, 2023.

<https://en.wikipedia.org/wiki/Coverture>.

[xvi] "Development of Rights of Women." Foundations of Law.

<https://lawshelf.com/coursewarecontentview/development-of-rights-of-women>.

[xvii] Ibid.

[xviii] Ibid.

[xix] "Amdt 6.5.4 Right to Compulsory Process - Constitution Annotated." Constitution Annotated. Accessed April 26, 2024.

constitution.congress.gov/browse/essay/amdt6-5-4/ALDE_00013333.

[xx] "Coverture." Wikipedia. Wikimedia Foundation, February 19, 2023.

<https://en.wikipedia.org/wiki/Coverture>.

[xxi] Harvard Mirador Viewer. [iif.lib.harvard.edu/manifests/view/drs:457607836\\$281i](https://iif.lib.harvard.edu/manifests/view/drs:457607836$281i), 223.

[xxii] Ibid., 223.

[xxiii] "Myra Bradwell." The Glinda Factor. <https://theglindafactor.com/myra-bradwell/>.

[xxiv] State of Illinois Office of the Illinois Courts. "Details: State of Illinois Office of the Illinois Courts." Administrative Office of the Illinois Courts, February 27, 2024.

www.illinoiscourts.gov/News/1331/Illinois-Supreme-Court-history-Myra-Bradwell/news-detail/#:~:text=In%201861%2C%20the%20Illinois%20legislature,to%20earn%20money%20for%20themselves.

[xxv] "Married Women's Property Acts in the United States." Wikipedia. Wikimedia Foundation,

https://en.wikipedia.org/wiki/Married_Women%27s_Property_Acts_in_the_United_States.

[xxvi] The Law Dictionary. "Who Was Myra Bradwell, America's First Woman Lawyer?"

The Law Dictionary,

<https://thelawdictionary.org/article/who-was-myra-bradwell-americas-first-woman-lawyer/>.

[xxvii] Friedman, Jane M. *America's First Woman Lawyer: The Biography of Myra Bradwell*. Prometheus Books, 2019.

[xxviii] Ibid.

[xxix] "Amdt 6.5.4 Right to Compulsory Process - Constitution Annotated." Constitution Annotated. Accessed April 26, 2024.

constitution.congress.gov/browse/essay/amdt6-5-4/ALDE_00013333.

[xxx] Oyez. "Slaughterhouse Cases." Oyez. <https://www.oyez.org/cases/1850-1900/83us130>.

- [xxxii] Constitution Center. "Bradwell v. Illinois (1873)." Constitution Center.
<https://constitutioncenter.org/the-constitution/supreme-court-case-library/bradwell-v-illinois-1873>
- [xxxiii] Ibid.
- [xxxiii] Ibid.
- [xxxiv] "In Re Lady Lawyers: The Rise of Women Attorneys and the Supreme Court."
Supreme Court of the United States.
www.supremecourt.gov/visiting/exhibitions/LadyLawyers/section1.aspx#LadyLawyers1.
- [xxxv] Ibid.
- [xxxvi] "Myra Bradwell." The Glinda Factor. <https://theglindafactor.com/myra-bradwell/>.
- [xxxvii] Bradwell, Myra. "Law Relating to Women." Chicago Legal News, November 21, 1868, Vol. 1 edition, sec. No. 8. 57-64.
- [xxxviii] Ibid., 57-64
- [xxxix] State of Illinois Office of the Illinois Courts. "Details: State of Illinois Office of the Illinois Courts." Administrative Office of the Illinois Courts, February 27, 2024.
www.illinoiscourts.gov/News/1331/Illinois-Supreme-Court-history-Myra-Bradwell/news-detail/#:~:text=In%201861%2C%20the%20Illinois%20legislature,to%20earn%20money%20for%20themselves.
- [xl] Bradwell, Myra. "Law Relating to Women." Chicago Legal News, November 7, 1868, Vol. 1 edition, sec. No. 6. 41-48.
- [xli] Bradwell, Myra. "Law Relating to Women." Chicago Legal News , October 17, 1868, Vol. 1 edition, sec. No. 3. 17-24
- [xlii] State of Illinois Office of the Illinois Courts. "Details: State of Illinois Office of the Illinois Courts." Administrative Office of the Illinois Courts, February 27, 2024.
www.illinoiscourts.gov/News/1331/Illinois-Supreme-Court-history-Myra-Bradwell/news-detail/#:~:text=In%201861%2C%20the%20Illinois%20legislature,to%20earn%20money%20for%20themselves.
- [xliii] Harvard Mirador Viewer. [iif.lib.harvard.edu/manifests/view/drs:457607836\\$281i](http://iif.lib.harvard.edu/manifests/view/drs:457607836$281i), 225.
- [xliv] Ibid., 225.
- [xlv] Bradwell, Myra. "Additional Brief" Chicago Legal News, February 5, 1870, Vol. 2 edition, sec. No. 19. 145-152.
- [xlvi] State of Illinois Office of the Illinois Courts. "Details: State of Illinois Office of the Illinois Courts." Administrative Office of the Illinois Courts, February 27, 2024.
www.illinoiscourts.gov/News/1331/Illinois-Supreme-Court-history-Myra-Bradwell/news-detail/#:~:text=In%201861%2C%20the%20Illinois%20legislature,to%20earn%20money%20for%20themselves.
- [xlvii] Bradwell, Myra. "Additional Brief" Chicago Legal News, February 5, 1870, Vol. 2 edition, sec. No. 19. 145-152.
- [xlviii] Ibid., 145-152.
- [xlix] ISBA. "The Remarkable Life and Times of Alta May Hulett." ISBA.
<https://www.isba.org/committees/women/newsletter/2004/02/theremarkablelifeandtimesofaltamayh>.
- [l] "Human and Civil Rights." Illinois State Bar Association.
www.isba.org/sections/humanrights/newsletter/2013/01/celebrating140yearsoffemaleattorney#:~:text=Happily%2C%20while%20the%20U.S.%20Supreme,the%20right%20to%20practice%20law

[li] "Women's Rights and Their Money: A Timeline from Cleopatra to Lilly Ledbetter."

The Guardian. Guardian News and Media, August 11, 2014.

<https://www.theguardian.com/money/us-money-blog/2014/aug/11/women-rights-money-timeline-history>.

[lii] "Human and Civil Rights." Illinois State Bar Association.

www.isba.org/sections/humanrights/newsletter/2013/01/celebrating140yearsoffemaleattorney#:~:text=Happily%2C%20while%20the%20U.S.%20Supreme,the%20right%20to%20practice%20law

[liii] Ibid.

[liiv] "On This Day, Women First Allowed to Argue Supreme Court Cases." National Constitution Center – constitutioncenter.org.

<https://constitutioncenter.org/blog/on-this-day-women-first-allowed-to-argue-supreme-court-cases>.

[liv] "Myra Bradwell." The Glinda Factor. <https://theglindafactor.com/myra-bradwell/>.

[lv] Ibid.

[lvii] Ibid.

Reproduction and Fertility: From Basic Science to Clinical Treatments

By Dhaya Sarathy

Abstract

Issues relating to infertility and reproduction are extremely widespread, and as medical technologies advance, many methods have been introduced proposing to mitigate these issues and help people to reproduce. The purpose of this article is to examine the foremost medical issues relating to infertility and reproductive health and to give an overview of both the solutions in use and the solutions being tested and proposed to combat reproductive struggles.

Keywords stem cells, fertility, infertility, mitochondrial replacement therapy, assisted reproductive technologies, mitochondrial disease, CRISPR-Cas9, polycystic ovarian syndrome, sperm disorders, ovarian disorders, intrauterine insemination, in-vitro fertilization, intracytoplasmic sperm injection, preimplantation genetic diagnosis, embryonic stem cells, induced pluripotent stem cells, somatic cell nuclear transfer, mutagenesis

Background

Infertility is defined as a lack of impregnation after 12 months or more of unprotected intercourse between a couple. [1] Some of the solutions relating to infertility and reproduction have been in use for several decades or throughout the 20th century, while others are more recent developments over the past 2-3 decades or less. In 2007, the European Society of Human Reproduction and Embryology (ESHRE) reviewed a number of studies over the past century on infertility, reaching the conclusion that fertility's global frequency was 9% (ranging from 3.5%–16.7% in developed nations and from 6.9%–9.3% in developing nations). [2]

In the past, infertility was largely attributed to sexually transmitted infections and other diseases like gonorrhea. However, more commonly diagnosed causes nowadays are stress and male factor infertility, and the increasing commonness of other diseases and health complications like diabetes, obesity, hypertension/hyperthyroidism, and addictions exacerbate the overall increase in infertility. Infertility remains difficult to diagnose in a large portion of cases. [2]

Peak fertility for women occurs in the mid-20s age range, and the fertility rate necessary for population replacement in developed countries is generally 2.1 births per woman. However, in more and more large cities and developed areas, the average age of a first birth can exceed 30 years, leaving fertility rates far below levels necessary for population replacement. This has led to the increasing popularity of IVF and egg cryopreservation, in addition to other methods of reproduction. [3] Issues relating to mitochondria often play a major role in female infertility. [4]

In one study conducted from 2005 to 2013, 20.4% of male partners struggling to reproduce were diagnosed with severe male factor infertility. The primary cause was defined for around 40% of those diagnosed, meaning that the reason behind the majority of cases of male infertility is still relatively vague or difficult to pinpoint. [5]

Some current and past technologies in use to combat infertility problems are intrauterine insemination, in vitro fertilization, intracytoplasmic sperm injection, and mitochondrial replacement therapy; these fall into the category of assisted reproductive technologies, and these methods can be used hand-in-hand with preimplantation genetic diagnosis. [6]

Research is still being conducted on many newer forms of technology, specifically types of stem cell research to aid in reproduction. These include embryonic stem cells, induced pluripotent stem cells, and somatic cell nuclear transfer; another important treatment worth mentioning is the use of CRISPR-Cas9 gene editing. [6]

Each of these methods targets different causes behind reproductive struggles, and results are still somewhat unreliable for the majority of these solutions. They each have their own specific problems which still need to be targeted through further research.

Issues of Infertility

Common Infertility Disorders in Women

As determined by the 2007 ESHRE study, female age was the greatest deciding factor in fecundity, but the age of the marriage also provided significant data. For couples married for under 5 years, polycystic ovarian syndrome (PCOS) and tubal causes related to STIs appeared more often; on the other hand, couples married for over 5 years were more likely to either have unattributable infertility issues or to have male factor infertility.

Female infertility is often caused by ovulatory disorders such as oligo-ovulation (irregular ovulation) or anovulation (lack of ovulation), wherein an oocyte is not released monthly from the uterus. [7]

Most prevalent among these ovulatory disorders is polycystic ovary syndrome (PCOS), which is defined as a hormonal imbalance in women leading to several health complications. PCOS is the most common endocrine pathology in females of reproductive age worldwide. The prevalence of PCOS typically ranges from 5-15%, affecting an estimated 5 million women in the US. PCOS is not only extremely common but also underdiagnosed; diagnosis usually takes place over several visits across a year or more. This delay can aggravate effects of the disease and make lifestyle intervention harder and harder to put in place. [7]

In addition to infertility, PCOS causes a host of other health complications, including metabolic syndrome, obesity, impaired glucose tolerance, type 2 diabetes, cardiovascular risk, depression, endometrial cancer, and nonalcoholic fatty liver disease. PCOS is more common when an individual has first-degree relatives suffering from PCOS, prepubertal obesity, congenital virilizing disorders, above-average or low birth weight for gestational age, premature adrenarche, or use of valproic acid as an antiepileptic drug. [8] Many of these ovulatory disorders do not have any known cure as of now.

Additionally, with advanced age, both the quality and quantity of the oocytes decline. This decline also becomes more and more evident over time, occurring at an especially increased rate around the mid-30s age range. Aged oocytes can lead to cytoplasmic deficiency, which in

turn causes chromosomal abnormalities within the oocytes, making fetal development not viable. Infertility rates in the 15-34 age group were 7.3-9.1%; rates in the 35-39 age group were 25%, and rates in 40-44 age group were 30%. Thus, there is a clear link between increased age and infertility in women. [7]

Endometriosis is a chronic disease caused by an excess of estrogen production. It is characterized by histological elements normally constituting the inner uterine lining instead being present on the outside of the uterus. These endometrial tissues, glands, and stroma typically develop in the ovaries, but they can also be present in other areas. [9]

Endometriosis is one of the most common benign gynecological diseases; it is found in approximately 6–10% of reproductive-aged women. The condition is mainly identified by severe and chronic pain including pelvic pain (experienced by 75% of patients), dysmenorrhea, deep dyspareunia, and infertility. [10] However, some studies indicate that endometriosis can be associated with a higher risk of subsequent infertility only among women over the age of 35. [11]

Endometriosis is present in 25-40% of infertile patients; adjusting for age, women with endometriosis had a 2-fold increased risk of incident infertility as compared to women with no history of endometriosis. [11] This infertility is typically a result of multiple factors such as implantation disorders, pelvic adhesions, and ovarian insufficiency. [10]

Due to fibroids (uterine growth) or other conditions known as congenital uterine abnormalities (CUAs), at times, even after getting pregnant successfully, miscarriages may occur in the second or third trimesters of pregnancy. Ovarian quality is also affected by external factors like consumption of alcohol, tobacco, and drugs. [7]

Mitochondrial Health

Mitochondria carry out important functions relating to human fertility and infertility in sperm, oocytes, and oocytes' surrounding cells. Even though sperm mitochondria are not transmitted to the future embryo, they're still important for producing the energy needed for sperm movement and overall function, as well as for sperm–oocyte fusion. In contrast, oocyte mitochondria produce energy required for oocyte meiotic division; thus, problems with oocyte mitochondria can lead to oocyte and embryo aneuploidy (a chromosome count differing from the exact multiple of the haploid karyotype). [13]

Mitochondria are known to be central players in energy production in somatic cells, providing the energy required for all cellular functions; however, the breadth of their purpose in cell metabolism of several specific cell types has long been misjudged. Besides, the metabolic uses of mitochondria surpass bioenergetics to include the processing of nutrients, generation of biosynthetic precursors for macromolecules, maintenance of redox homeostasis, and waste management. The role of mitochondria in sex cells and the early stages of embryos is more obfuscated still. [12]

A vulnerability of mitochondria in reproduction is that mtDNA is more exposed to cumulated oxidative damage than nuclear DNA. Many of the problems caused by sperm mitochondrial abnormalities disrupt the typical process of fertilization, but they can be resolved

with the use of assisted reproduction technologies (ARTs) and intracytoplasmic sperm injections (ICSI); this can be particularly attributed to the fact that mitochondria from sperm cells are not transmitted to the embryos. Oocyte mitochondria, however, are transmitted to early embryos. The consequence of this is that any mitochondrial abnormalities can create long-term issues for the embryo's subsequent development. [12] Among human metabolic disorders, primary mitochondrial diseases are the most common, being present in 1 out of 4300 cases. [14]

MDS, or mtDNA depletion syndromes, are a subset of disorders characterized by an extreme decrease in mtDNA levels; as a result, affected tissues and organs may not be able to produce sufficient energy. MDS are a result of mutations in nuclear genes functioning in mitochondrial nucleotide synthesis or mtDNA replication. To produce subunits of mitochondrial respiratory chain complexes necessary for energy, a certain amount of mtDNA is necessary. If this benchmark is not met, it can lead to organ dysfunction. [12]

Myoclonic epilepsy with ragged red fibers, or MERRF, is similarly a result of genetic mutations in mtDNA. [15] Myoclonus is defined as involuntary muscle contraction, causing sudden jerking movements of a muscle, set of muscles, or the body; some examples are hiccups or sleep starts. [16] MERRF causes progressive myoclonus and seizures, in addition to a number of other symptoms such as myopathy, cardiac arrhythmia, hearing loss, dementia, etc. [15]

Peripheral neuropathies are a subset of disorders affecting the abilities of peripheral nerve cells and fibers, such as cranial nerves, spinal nerve roots, nerve trunks and division, autonomic nervous system nerves, etc. Peripheral neuropathies most often cause types of paresthesia like tingling, prickling, chilling burning, or numbness, in addition to pain, weakness, and loss of deep tendon reflexes, all with a broad range of severities. [17] Peripheral neuropathies are often caused by mtDNA or nuclear DNA mutations, disrupting nervous systems' normal distribution and transport. [18]

Like the aforementioned mitochondrial diseases, mutations in mtDNA can cause myopathy, a muscle weakness disorder, which ranges in severity and can even lead to death. [19]

Primary mitochondrial disease can be difficult to diagnose despite its wide prevalence because it often presents a variety of phenotypes in patients. For this reason, current treatment methods typically involve supportive and preventative measures as opposed to widespread disease-specific therapies. [20]

Male infertility

Although many cases of inability to reproduce can be attributed to female infertility (this, at times, being a conclusion drawn without substantial evidence, a sizable portion of these cases are caused by male infertility or both male and female infertility. A primary concern is oligozoospermia and azoospermia—the inability to produce a sufficient amount of viable sperm and the inability to produce any sperm, respectively. [21]

In one study conducted from 2005 to 2013, 20.4% of male partners struggling to reproduce were diagnosed with severe male factor infertility. The primary cause was defined for

around 40% of those diagnosed, meaning that the reason behind the majority of cases of male infertility is still relatively vague or difficult to pinpoint. [22]

Azoospermia is identified in about 1% of men and about 10-15% of infertile men. Male factor infertility may result from an underlying medical condition that is often treatable but could possibly be life-threatening. Historically, when a diagnosis of azoospermia was reached, a sperm donor was seen as the best option to conceive. In more recent times, however, medical practitioners can take advantage of the general knowledge that many root causes of azoospermia are reversible. [21]

Intracytoplasmic sperm injection (ICSI), which will be covered further in this article, has come to be utilized as a suitable treatment for most male factor infertility cases. Even men with potentially treatable causes of infertility are most commonly treated instead with assisted reproductive techniques (ARTs). However, in the case of an azoospermia diagnosis, no sperm will be produced through ejaculation, so ARTs are inapplicable. [21]

A study conducted in Ukraine found that 17% of patients with sperm disorders had chromosomal abnormalities; this figure breaks down to 35% of patients with azoospermia and 12.7% of patients with oligozoospermia in whom the researchers detected chromosomal abnormalities. The detection of these abnormalities not only provides a more complete diagnosis but also significantly affects the method with which infertility is treated. [23]

The male reproductive system relies on the arrangement of several hormones. For example, the testicular testosterone level must be much greater than that in the serum to create suitable conditions for normal sperm production. This testicular testosterone indirectly increases germ cell maturation. Even though endocrinopathies are associated with male infertility only in around 1-2% of cases, treatment of these hormone-related disorders provides patients with a more case-specific approach to combating infertility—the causes of which can be classified into either the lack of or excess of certain hormones. [24]

Male factor infertility is typically most responsive not to primary measures to combat infertility but instead to secondary measures; these include artificial/intrauterine insemination, in-vitro fertilization (IVF) and embryo transfer, and intra-cytoplasmic sperm transfer (ICSI), in addition to simply choosing adoption or fostering of children. Some risk factors that have been speculated to worsen struggles with infertility include smoking, alcohol intake, drugs, obesity, testicular infections, environmental toxins, excessive heat exposure in testicles, hormonal disorders, ejaculatory/erectile disorders, etc. However, there is a dearth of studies comparing and evaluating these potential causes. [25]

Idiopathic infertility, or infertility wherein the cause is not known, is the diagnosis for 44% of male infertility cases, making it the most common diagnosis. Studies spanning the last 50 years have demonstrated a steady rise in male infertility rates which can't be completely accounted for with obesity, genetics, and lifestyle (including diet) changes. This fertility decline occurred contemporaneously with a dramatic change in human exposures. Of the many synthetic chemicals that are now all around humans, some are harmless, and others can disrupt the hypothalamic-pituitary-gonadal axis and impair sperm production abilities. [26]

Past and Current Mainstream Solutions

Assisted Reproductive Technologies

Assisted reproductive technologies (ART) is an umbrella term including several different solutions for infertility- and reproductive health-related issues. By definition, ARTs are techniques that involve manipulation of eggs outside the body. [27] The most famous of these solutions is in-vitro fertilization, or IVF. [28]

ARTs are defined as treatments that manipulate eggs or embryos for fertility purposes, but they do not include manipulation of sperm or stimulation of ovaries. ARTs may involve the uterus, the fallopian tubes, and the ovaries, and they can also be used in non-fertility related treatments such as preimplantation genetic testing, which will be covered in greater detail later in this article. [28]

Intrauterine Insemination

Intrauterine insemination (IUI) is an ART involving the insertion of a processed semen sample into the upper uterus, circumventing the female reproductive tract's barriers to sperm-oocyte proximity. Compared to other techniques, IUI is a cost-effective, noninvasive treatment for select eligible patients. It, however, has limited use in patients with endometriosis, severe male factor infertility, tubal factor infertility, advanced maternal age ≥ 35 years, and many reproductive system-concerned abnormalities. [29]

Even though newer procedures such as IVF and ICSI have been introduced since IUI came into use, IUI remains an inexpensive, noninvasive, and effective first-line therapy for selected couples with typically male factor but also more recently female factor infertility issues of various origins. Even though its basis is the same, IUI has advanced in terms of stimulation protocols, gonadotropins, sperm preparation techniques, and ultrasound monitoring. Strict patient selection criteria and individualized stimulation protocols tailored according to the age and etiology of the patient help minimize risk of complications. [29]

The rationale behind IUI is increasing the gamete density to improve basic chances of fertilization. Historically, artificial insemination was used in farm animals to increase the rate of genetic improvement by accelerating the reproduction of certain species. For humans, however, IUI was originally developed to induce pregnancy in couples struggling with physical or psychological male factor subfertility. Nowadays, IUI with homologous semen (semen coming from the male partner, as opposed to a sperm donor) is most common for idiopathic and mild male factor subfertility. [30]

Semen preparation techniques— notably the gradient, swim-up, and wash techniques— are a key step in IUI to separate viable and not viable sperm to inject a greater concentration of viable sperm and increase likelihood of conception. [31]

In Vitro Fertilization

The term “in-vitro fertilization” can be broken down into two parts. “In vitro” means outside the organism: through in-vitro fertilization, while oocytes mature in the ovary and embryos develop inside the uterus, the fertilization of the oocytes occurs in a petri dish. IVF now accounts for 1.6% of U.S. live births and 4.5% of European live births. Similar to IUI, the procedure is now widely applied for the treatment of infertility due to a variety of causes, including endometriosis, male factor, and unexplained infertility. IVF is also a compatible procedure with donor oocytes, making it an option for women with primary ovarian insufficiency or age-related decline in egg count. Potential IVF patients are evaluated in several rounds prior to treatment; if the ovarian reserve is insufficient in any value, IVF is still an option, but it may need to be pursued with donor oocytes. Similarly, male semen is analyzed for necessity of ICSI. [\[27\]](#)

The IVF cycle begins with ovarian stimulation. Ovarian stimulation is performed in most IVF cycles, allowing for about 10-20 oocytes to be retrieved. This retrieval process uses ultrasound-guided transvaginal aspiration and intravenous sedation. The ovaries are visualized using an ultrasound probe, and an attached needle guide is directed into each follicle. [\[27\]](#)

Decreased sperm count or abnormal shape or motility is medically and/or surgically treatable in approximately 50% of men, especially with IUI. However, if treatments fail, IVF with or without intracytoplasmic sperm injection (ICSI) can be utilized. Sperm extracted in cases of obstructive azoospermia or testicular hypofunction can only be used in an IVF cycle with ICSI because the sperm have not undergone the final in-vivo maturation process which would allow sperm to fertilize an oocyte. It is well established that women’s fertility dramatically decreases in their 40s due to a decline in oocyte quality and quantity; oocyte cryopreservation is an option for women who have illnesses such as cancer that may affect oocyte viability or who simply wish to delay childbearing. [\[27\]](#)

Intracytoplasmic Sperm Injection

ICSI is typically used along with IVF, assisting the latter procedure in embryo formation and development. ICSI is most helpful for couples where the male partner struggles with azoospermia or oligozoospermia. It’s also used in gametes that inherently struggle with fertilization, such as thawed oocytes that had previously undergone cryopreservation. ICSI forces the fertilization process to take place by injecting a single sperm into the egg in vitro, bypassing clinical and biological obstacles to embryo formation. [\[32\]](#) However (though the full picture involving the risks of both ICSI and IVF is still not known), ICSI may have consequences including ovarian hyperstimulation and congenital abnormalities. [\[33\]](#)

Preimplantation Genetic Diagnosis

The predecessors of preimplantation genetic testing (PGT) were born in 1890 when Walter Heape successfully transferred embryos into Belgian Hare doe rabbits. Building on this experiment in 1935, Gregory Pincus cultured rabbit oocytes to the metaphase stage of meiosis. Then, in the 1960s, Robert Edwards (who also paved the way for IVF treatment) came up with

the idea of PGT. These developments in PGT occurred in parallel to the development of IVF. [\[34\]](#)

The biopsy techniques found for preimplantation genetic diagnosis (PGD) are blastomere biopsy, trophoctoderm biopsy, and polar body biopsy. Wilton and Trounson discovered blastomere biopsy from the cleavage stage of mouse embryos. In 1988, Marilyn Monk and Audrey Muggleton-Harris formed the trophoctoderm biopsy technique and PGD. Yury Velinsky and his team used this biopsy technique to find maternal unaffected genes. [\[34\]](#)

PGD had its first successful trial in 1990 using mice. Handyside and team were able to impregnate mice using PGD for sex-related disease. They used the polymerase chain reaction technique to find male embryos that weren't affected by disease passed down from x-linked recessive inheritance. They went on to introduce cases of PGD identification for male, female, and Turner Syndrome embryos. [\[34\]](#) To do so, they used a method known as fluorescence in situ hybridization, wherein a fluorescent reporter molecule containing a target sequence of a DNA sample is connected to DNA or RNA probes which are to be recombined. This method was used to apply PGD techniques to humans starting in 2001, and it is an ongoing process. [\[35\]](#)

Mitochondrial Replacement Therapy

Mitochondrial replacement therapy (MRT) involves the introduction of healthy donor mitochondria into patients' mutated cells. MRT is typically used for mitochondrial disease or issues caused by female age-related infertility. MRT is the only solution that actually proposes to eliminate risks of passing down mitochondrial disease to future generations. [\[36\]](#)

There are several types of MRT: pronuclear transfer (PNT), spindle transfer (ST), and polar body transfer (PBT). Pronuclear transfer involves the implantation of both pronuclei (one from sperm and one from the oocyte) from one zygote to another. PNT was first successfully accomplished with mice in the 1980s. However, high mitochondrial DNA (mtDNA) carryover levels have affected the success rate. [\[36\]](#)

Spindle transfer was first introduced with the rhesus monkey while the oocyte was still unfertilized. Since mitochondrial distribution in oocyte is far more even than mitochondrial distribution in zygotes, the mtDNA carryover level problem was largely addressed. In ST, karyoplasts were fused to donor cytoplasts, and ICSI and embryo culture were then used to form the blastocyst. [\[36\]](#)

Polar body transfer involves two polar bodies formed in the process of meiosis. The first has a diploid chromosome set and the second has a haploid chromosome set. PBT proposes that one of these polar bodies be implanted into an oocyte or the cytoplasm of a zygote. The advantage of this technique would be the efficient handling of a fewer number of mitochondria which would improve the effectiveness of PNT and ST used in combination with PBT. Research still needs to be done on this treatment in humans; one limitation is that the polar bodies tend to have a relatively short life because of DNA fragmentation and degradation. [\[36\]](#)

Stem Cell Technologies

Embryonic Stem Cells

There are 2 types of pluripotent stem cells (PSCs): embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs). Pluripotent stem cells can differentiate into any bodily tissue cells— everything except the placenta. [37] PSCs can be differentiated into oocytes and sperm, making them useful for fertility and reproduction. PSCs proliferate and regenerate very quickly, so they're suitable for drug screening as well as modeling and treating diseases. [6]

ESCs were first isolated from inner cell walls of blastocysts of mice in 1981, and human ESCs were established in 1998. As discovered with mice in 1981, ESCs could be maintained indefinitely in culture because they would renew by themselves through cell division and continue to be able to differentiate into all somatic cell lineages. [37]

Induced Pluripotent Stem Cells

iPSCs are pluripotent stem cells derived from somatic cells. They were introduced by Japanese scientist Shinya Yamanaka in 2006, who reprogrammed murine skin cells (specifically fibroblasts) in adult mice into iPSCs. The iPSCs were very similar in morphology and function to ESCs. [6]

iPSCs provide a solution to some of the ethical issues presented by human ESC solutions, but they are limited in that they have some genetic instability and cannot fully erase certain genetic markings. [6]

Somatic Cell Nuclear Transfer

SCNTs involve enucleation or removal of the nucleus from the oocyte, followed by donor cell injection and fusion with the enucleated oocyte, then artificial activation of the reconstructed oocyte, and finally formation of a blastocyst. [6]

Somatic cell nuclear transplants, or SCNTs, were first performed in 1962. After stimulating the oocyte, it was fertilized and could develop into an organism. The 2006 study by Yamanaka and colleague Takahashi used this knowledge in reprogramming murine fibroblasts into iPSCs. [6]

CRISPR/Cas9

CRISPR (clustered regularly interspaced short palindromic repeats) is a tool which is very useful for gene editing because it allows for knocking genes in and out in addition to making insertions and deletions for cell lines. It relies on a small guide RNA which leads the Cas9 protein to the area of modification for cleavage. [38]

The CRISPR/Cas9 technique has been used to create new mutations in DNA that would produce homozygous traits in organisms (this has been tested at length in mice). [38]

These techniques can be applied to mutations concerned with infertility and reproductive disorders and diseases in combination with other techniques such as, notably, somatic cell nuclear transfer. [38]

CRISPR/Cas9 is also very relevant to human cell health. Several studies have been done on CRISPR/Cas9 in human stem cells which pave the way for further research on oocytes and embryos. Other trials have determined that certain methods of Cas9 editing such as dCas9-VP64 targeted by small guide RNA can increase certain gene expressions for humans. [38]

CRISPR/Cas9 additionally demonstrates promise in eliminating many of the ethical concerns that arise from methods such as mitochondrial replacement therapy because unlike MRT, CRISPR/Cas9 does not require a donor. [39]

Problems with Stem Cell Technologies

Production of Gametes from Opposite Sex Primordial Cells

Not enough research has been done to support the suggestion that oocytes can be produced from male cells with XY chromosomes, and less so that sperm cells can be produced from female cells with XX chromosomes. Additionally, there is no empirical indication that either sperm or oocytes can be produced without the presence of a Y-chromosome. There have been some studies that suggest the possibility of XY primordial germ cells producing XX oocytes from male multipotent cells, but significantly more research is required on this topic. [6]

Haploid iPSCs

Haploidization is defined as the splitting of a diploid cell into 2 haploid cells through meiosis (a process typically specific to germ cells). The potential for somatic cell haploidization is currently very unclear. Some studies have suggested the possibility using insertion of diploid cell nuclei into enucleated oocytes to induce haploidization, but any promising results of these studies were outliers in the data. There's also the problem of epigenetic memory, as mentioned previously: iPSCs cannot erase certain genetic markings, leaving potential bias in the ultimate demonstration of the gene. [6]

Mutagenesis

Mutagenesis is a phenomenon that can occur due to DNA replication errors, damage, and lab techniques. Mutagenesis describes the unintended effects on completely separate genes caused by the editing of a singular gene; this issue becomes particularly relevant in context of the usage of CRISPR/Cas9 and other gene editing tools. [40]

Specifically when it comes to stem cells and gene editing, modification of genes can have effects beyond what is predicted in a lab setting. This can have severe health-related and ethical implications. [37]

Conclusion

There are many solutions, both established and developing, proposing to identify and/or mitigate different aspects of fertility and reproductive health issues. Different techniques produce different results on a case-by-case basis. With many stem-cell based and biotechnology-involved

solutions, there is still a lot of research yet to be done before it can be widely implemented, because each technique has both its advantages and its drawbacks. It's important that research continues to develop on this issue due to its gravity; up-and-coming research in clinical approaches to fertility holds the potential to improve genetic lines not only in isolated cases but also over many generations.

Works Cited

- [1] Vander Borgh, Mélodie, and Christine Wyns. "Fertility and infertility: Definition and epidemiology." *Clinical biochemistry* vol. 62 (2018): 2-10.
doi:10.1016/j.clinbiochem.2018.03.012, Published 12 March 2018,
pubmed.ncbi.nlm.nih.gov/29555319/.
- [2] Deshpande, Priyanka Sanjay, and Alka ShantiPrakash Gupta. "Causes and Prevalence of Factors Causing Infertility in a Public Health Facility." *Journal of human reproductive sciences* vol. 12,4 (2019): 287-293. doi:10.4103/jhrs.JHRS_140_18, Published 17 Dec 2019, www.ncbi.nlm.nih.gov/pmc/articles/PMC6937760/.
- [3] Kushnir, Vitaly A et al. "The Future of IVF: The New Normal in Human Reproduction." *Reproductive sciences (Thousand Oaks, Calif.)* vol. 29,3 (2022): 849-856.
doi:10.1007/s43032-021-00829-3, Published online 3 Jan 2022,
www.ncbi.nlm.nih.gov/pmc/articles/PMC8722744/.
- [4] Rodríguez-Varela, Cristina, and Elena Labarta. "Role of Mitochondria Transfer in Infertility: A Commentary." *Cells* vol. 11,12 1867. 8 Jun. 2022, doi:10.3390/cells11121867,
Published 8 June 2022, www.ncbi.nlm.nih.gov/pmc/articles/PMC9221194/.
- [5] Punab, M et al. "Causes of male infertility: a 9-year prospective monocentre study on 1737 patients with reduced total sperm counts." *Human reproduction (Oxford, England)* vol. 32,1 (2017): 18-31. doi:10.1093/humrep/dew284, 17 November 2016,
pubmed.ncbi.nlm.nih.gov/27864361/.
- [6] Lee, Yeonmi and Kang, Eunju. "Stem Cells and Reproduction". PubMed Central, National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, 31 July 2019,
www.ncbi.nlm.nih.gov/pmc/articles/PMC6726211/#b12-bmb-52-482.
- [7] Walker, Matthew H. and Tobler, Kyle J. "Female Infertility". StatPearls, Trends Mol Med, National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, 19 December 2022, www.ncbi.nlm.nih.gov/books/NBK556033/.
- [8] Rasquin, Lorena I. et al. "Polycystic Ovarian Disease". StatPearls, National Library of Medicine, National Center for Biotechnology Information, National Institute of Health, Updated 15 November 2022, www.ncbi.nlm.nih.gov/books/NBK459251/.
- [9] Tsamantioti, Eleni S. and Mahdy, Heba. "Endometriosis". StatPearls, National Library of Medicine, National Center for Biotechnology Information, National Institute of Health, Updated 23 Jan 2023, www.ncbi.nlm.nih.gov/books/NBK567777/.
- [10] Chantalat, Elodie et al. "Estrogen Receptors and Endometriosis." *International journal of molecular sciences* vol. 21,8 2815, doi:10.3390/ijms21082815, 17 Apr 2020,
www.ncbi.nlm.nih.gov/pmc/articles/PMC7215544/.
- [11] Prescott, J et al. "A prospective cohort study of endometriosis and subsequent risk of infertility." *Human reproduction (Oxford, England)* vol. 31,7 (2016): 1475-82.
doi:10.1093/humrep/dew085, Published 1 May 2016,
www.ncbi.nlm.nih.gov/pmc/articles/PMC4901880/.

- [12] Tesarik, Jan, and Raquel Mendoza-Tesarik. "Mitochondria in Human Fertility and Infertility." *International journal of molecular sciences* vol. 24,10 8950, doi:10.3390/ijms24108950, 18 May 2023, www.ncbi.nlm.nih.gov/pmc/articles/PMC10218931/.
- [13] Torres, Eduardo M et al. "Aneuploidy: cells losing their balance." *Genetics* vol. 179,2 (2008): 737-46. doi:10.1534/genetics.108.090878, Jun 2008, www.ncbi.nlm.nih.gov/pmc/articles/PMC2429870/.
- [14] Lopriore, Piervito et al. "Mitochondrial Epilepsy, a Challenge for Neurologists." *International journal of molecular sciences* vol. 23,21 13216, doi:10.3390/ijms232113216, 30 Oct. 2022, www.ncbi.nlm.nih.gov/pmc/articles/PMC9656379/.
- [15] Hameed, Sajid and Tadi, Prasanna. "Myoclonic Epilepsy and Ragged Red Fibers". StatPearls, National Library of Medicine, National Center for Biotechnology Information, National Institute of Health, Updated 16 July 2023, www.ncbi.nlm.nih.gov/books/NBK555923/.
- [16] Sanders, Amy E. et al. "Myoclonus". StatPearls, National Library of Medicine, National Center for Biotechnology Information, National Institute of Health, Updated 26 February 2024, www.ncbi.nlm.nih.gov/books/NBK537015/.
- [17] Hammi, Claudia and Yeung, Brent. "Neuropathy". StatPearls, National Library of Medicine, National Center for Biotechnology Information, National Institute of Health, Updated 15 October 2022, www.ncbi.nlm.nih.gov/books/NBK542220/.
- [18] Horvath, Rita et al. "Peripheral neuropathy in mitochondrial disease." *PubMed, Handbook of clinical neurology* vol. 194 (2023): 99-116. doi:10.1016/B978-0-12-821751-1.00014-2, pubmed.ncbi.nlm.nih.gov/36813324/.
- [19] Ahuja, Abhimanyu S. "Understanding mitochondrial myopathies: a review." *PubMed, PeerJ* vol. 6 e4790. 21 May. 2018, doi:10.7717/peerj.4790, www.ncbi.nlm.nih.gov/pmc/articles/PMC5967365/.
- [20] Tinker, Rory J et al. "Current and Emerging Clinical Treatment in Mitochondrial Disease." *PubMed, Molecular diagnosis & therapy* vol. 25,2 (2021): 181-206. doi:10.1007/s40291-020-00510-6, Published 1 Mar 2021, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7919238/>.
- [21] Cocuzza, Marcello et al. "The epidemiology and etiology of azoospermia." *PubMed Central, Clinics (Sao Paulo, Brazil)* vol. 68 Suppl 1,Suppl 1 (2013): 15-26. doi:10.6061/clinics/2013(sup01)03, Feb 2013, www.ncbi.nlm.nih.gov/pmc/articles/PMC3583160/.
- [22] Punab, M et al. "Causes of male infertility: a 9-year prospective monocentre study on 1737 patients with reduced total sperm counts." *PubMed, Human reproduction (Oxford, England)* vol. 32,1 (2017): 18-31. doi:10.1093/humrep/dew284, Published 17 Nov 2016, pubmed.ncbi.nlm.nih.gov/27864361/.

- [23] Pylyp, Larysa Y et al. "Chromosomal abnormalities in patients with oligozoospermia and non-obstructive azoospermia." PubMed Central, *Journal of assisted reproduction and genetics* vol. 30,5 (2013): 729-32. doi:10.1007/s10815-013-9990-4, Published online 11 Apr 2013, www.ncbi.nlm.nih.gov/pmc/articles/PMC3663966/.
- [24] Sengupta, Pallav et al. "Endocrinopathies and Male Infertility." PubMed Central, *Life* (Basel, Switzerland) vol. 12,1 10. 22 Dec. 2021, doi:10.3390/life12010010, Published online 22 Dec 2021, www.ncbi.nlm.nih.gov/pmc/articles/PMC8779600/.
- [25] Okonofua, Friday Ebhodaghe et al. "Causes and Risk Factors for Male Infertility: A Scoping Review of Published Studies." PubMed Central, *International journal of general medicine* vol. 15 5985-5997. 4 Jul. 2022, doi:10.2147/IJGM.S363959, www.ncbi.nlm.nih.gov/pmc/articles/PMC9268217/.
- [26] Krzastek, Sarah C et al. "Impact of environmental toxin exposure on male fertility potential." PubMed Central, *Translational andrology and urology* vol. 9,6 (2020): 2797-2813. doi:10.21037/tau-20-685, Dec 2020, www.ncbi.nlm.nih.gov/pmc/articles/PMC7807371/.
- [27] Choe, Jennifer and Shanks, Anthony L. "In Vitro Fertilization". StatPearls, National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, Updated 4 Sep 2023, www.ncbi.nlm.nih.gov/books/NBK562266/.
- [28] Jain, Meaghan and Singh, Manvinder. "Assisted Reproductive Technology (ART) Techniques". StatPearls, National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, Updated 7 June 2023, www.ncbi.nlm.nih.gov/books/NBK576409/.
- [29] Allahbadia, Gautam N. "Intrauterine Insemination: Fundamentals Revisited." PubMed Central, *Journal of obstetrics and gynaecology of India* vol. 67,6 (2017): 385-392. doi:10.1007/s13224-017-1060-x, Published online 25 Oct 2017, www.ncbi.nlm.nih.gov/pmc/articles/PMC5676579/.
- [30] Ombelet, W, and J Van Robays. "Artificial insemination history: hurdles and milestones." PubMed Central, *Facts, views & vision in ObGyn* vol. 7,2 (2015): 137-43, www.ncbi.nlm.nih.gov/pmc/articles/PMC4498171/.
- [31] Boomsma, Carolien M et al. "Semen preparation techniques for intrauterine insemination." PubMed Central, *The Cochrane database of systematic reviews* vol. 10,10 CD004507. 15 Oct. 2019, doi:10.1002/14651858.CD004507.pub4, www.ncbi.nlm.nih.gov/pmc/articles/PMC6792139/.
- [32] Chamayou, Sandrine et al. "What Does Intracytoplasmic Sperm Injection Change in Embryonic Development? The Spermatozoon Contribution." PubMed Central, *Journal of clinical medicine* vol. 12,2 671. 14 Jan. 2023, doi:10.3390/jcm12020671, www.ncbi.nlm.nih.gov/pmc/articles/PMC9867417/.
- [33] Alukal, Joseph P, and Dolores J Lamb. "Intracytoplasmic sperm injection (ICSI)--what are the risks?." PubMed Central, *The Urologic clinics of North America* vol. 35,2 (2008):

- 277-88, ix-x. doi:10.1016/j.ucl.2008.01.004, May 2008,
www.ncbi.nlm.nih.gov/pmc/articles/PMC2424218/.
- [34] Parikh, Firuza Rajesh et al. "Preimplantation Genetic Testing: Its Evolution, Where Are We Today?" PubMed Central, National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, Oct-Dec 2018,
www.ncbi.nlm.nih.gov/pmc/articles/PMC6333033/.
- [35] Shakoori, Abdul Rauf. "Fluorescence In Situ Hybridization (FISH) and Its Applications". PubMed Central, National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, Edited by Tariq Ahmad Bhat and Aijaz Ahmad Wani, 10 February 2017, www.ncbi.nlm.nih.gov/pmc/articles/PMC7122835/.
- [36] Wolf, Don P. et al. "Mitochondrial Replacement Therapy in Reproductive Medicine". PubMed Central, Trends Mol Med, National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, February 2015,
www.ncbi.nlm.nih.gov/pmc/articles/PMC4377089/.
- [37] Romito, Antonio and Cobellis, Gilda. "Pluripotent Stem Cells: Current Understanding and Future Directions". PubMed Central, National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, 20 Dec. 2015,
www.ncbi.nlm.nih.gov/pmc/articles/PMC4699068/.
- [38] Zhang, Yi-ran et al. "CRISPR/Cas9 Technology: Applications In Oocytes And Early Embryos". Journal of Translational Medicine, PubMed, National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, 24 October 2023, pubmed.ncbi.nlm.nih.gov/37875936/.
- [39] Fogleman, Sarah et al. "CRISPR/Cas9 and mitochondrial gene replacement therapy: promising techniques and ethical considerations". Am J Stem Cells, PubMed Central, National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, 20 Aug 2016, www.ncbi.nlm.nih.gov/pmc/articles/PMC5043096/.
- [40] Durland, Justin and Ahmadian-Moghadam, Hamid. "Genetics, Mutagenesis". StatPearls, National Center for Biotechnology Information, National Library of Medicine, National Institute of Health, Updated 19 September 2022,
www.ncbi.nlm.nih.gov/books/NBK560519/.

Designing an Affordable Prosthetic Hand with Bipolex Functionality By Taneesh Matharasi

Summary

According to the World Health Organization, close to 40 million people worldwide are in need of a prosthetic, while over 75% of developing countries lack access to them (Burt). This lack of opportunity demonstrates the need for an affordable, accessible option to be developed. Many prosthetic hand designs are available today, but while some are inexpensive, they lack specific control over individual joints or are unable to match the range of motion of a real human hand. To meet this need, a prosthetic was designed to offer heightened functionality to users while remaining affordable and accessible. To compensate for lack of functionality in a prosthetic, this design involves two thumbs, allowing a far wider range of actions for the user to perform with just one hand. In order to maintain low costs, the materials used are mostly 3D printed parts, aside from the motors. Each finger will contain two hinge joints, while the thumbs will only contain one hinge joint and remain angled. The hinge motion of each digit is controlled by a motor, and for the fingers with two hinge joints these will be connected by a linkage. The thumbs and wrist will be placed on rolling joints, to allow for a wider range of motion. The rolling of each thumb will be controlled by a motor while the rolling of the wrist will be passive but limited mechanically. After designing the hand in CAD, various simulations were performed to determine how much functionality was achieved. A stress analysis was conducted to evaluate strength and it was determined that the hand would be able to withstand over 500 lbs of force, allowing the user to lean their body weight on the prosthetic without having to be concerned about damage. The prosthetic was also able to achieve a larger range of motion than a real human hand in its hinge joints.

Introduction

Context:

In the world, upper extremity amputations make up 30% of all amputees, affecting 3 million people (LeBlanc). While prosthetic solutions exist, they are extremely costly and often unrealistic for amputees to acquire. Patients without insurance can expect to pay up to \$5,000 for a simple cosmetic limb, and an additional \$5,000 to make it functional (Vandersea). These numbers are often cost prohibitive to patients in need, and they are thus unable to carry out basic human functions required of everyday life. In order to solve this problem, a functional but affordable prosthetic hand model was designed.

To approach this problem, we must first observe the functionality of a real human hand. A human hand contains 4 fingers and 1 thumb (for the purposes of this project the thumb is distinguished from the fingers due to difference in function). These 5 digits all connect to the palm and are able to move in a jointed motion towards the palm and extend back out. The thumb is unique, in that it has a wider range of horizontal motion, allowing for a variety of grips including the key grip, the chuck grip, and the hook grip, as shown below.



Fig. 3 Key pinch.



Fig. 4 Chuck grip/directional grip.



Fig. 5 Hook grip.

Depiction of three mentioned grips that demonstrate the versatility of the thumb's motion (Duncan).

The human hand controls each digit using muscles that pull tendons, creating tension. Joints between different sections exist between bones. While the human hand is very functional, it is unable to do many things on its own, without the use of another hand. One example is opening a bottle in the air, since one hand is required to grip the bottle while the other twists it open.

Most prosthetic models on the market aim to return the user to original human functionality by offering a realistic design of the human hand. An example of the average prosthetic hand on the market is the one designed by WPI. This design uses DC motors, worm gears, and linkages to drive the joints on each finger as well as to control the rolling motion of the thumb. The prosthetic uses toe switches within each shoe to allow the user to control the prosthetic by selecting preset grips. The price of production for this prosthetic comes out to be around \$3,000 which is actually on the cheaper end of the market (Ventimiglia).

Technical Specs/Design Constraints:

The hand designed consists of 4 fingers and two thumbs, each with the ability to hinge on two separate joints, except for the thumbs. Because of the structure of the thumb required in everyday tasks such as gripping, the thumb will also be able to roll on a joint at its base, although it will only be able to hinge at the base due to its short length. The wrist is also a necessary component that needs to be able to rotate in many actions. This joint will not be powered, but will be able to passively roll, however with a restricted range of motion to avoid the hand falling limp. For all the powered motion, DC motors will be used since they are cheap and effective. Each finger with two joints will use one motor, by connecting the two joints with a linkage. The housing for the hand will be made from plastic and encased in rubber, which are both affordable materials. The plastic is rigid and allows all internal components to maintain stability, while the rubber on the outside simulates the grip and texture of human skin, which improves the functionality. The arm is controlled by a remote that can either individually set each joint or select from an option of preset grips. The remote can be operated by the other hand of the user.

The major constraint of this design is the price. In order to be widely affordable, the price should be below \$500. Despite this, the design should still be functional. This can be measured by performing calculations based on the stress capacity of the material and the strength of the motors, as well as measuring range of motion in the number of degrees of motion. Additionally,

in order to be widely accessible, the design should be able to fit all the components while mimicking the size and shape of a human hand.

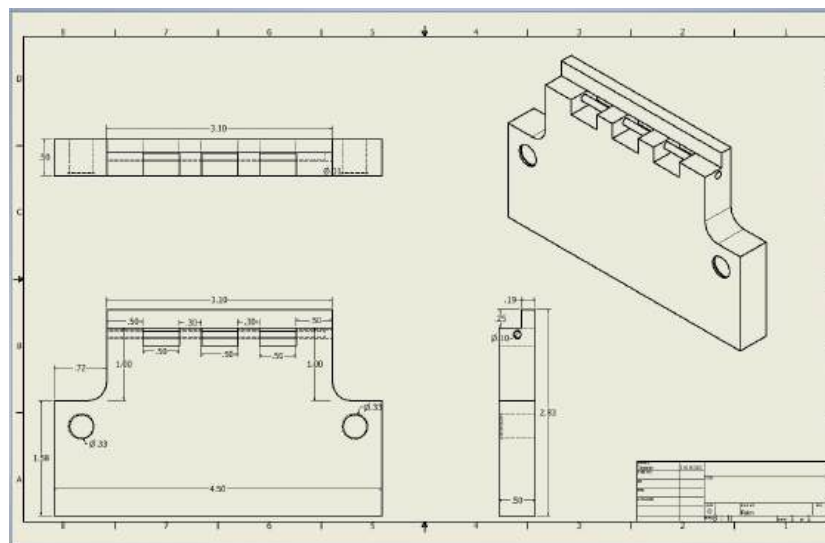
Design Objectives:

One of the main objectives of this design include achieving 60-70% of full human hand functionality. The arm is meant to be cost effective, and therefore achieving full capability is not feasible, but focusing on the important functionality seems to be the most reasonable approach in designing this prosthetic. Within this definition of the functionality, the arm should achieve the designated threshold for range of motion in degrees of freedom, strength based on pinch test, size, and weight as compared to a human hand. Because the prosthetic is designed to be cost-efficient and therefore requires a simple remote control, the implementation of two thumbs in the design allows the user to use the prosthetic for many tasks that would usually require both hands, such as gripping something with one hand while twisting or tearing it open with the other.

Analysis

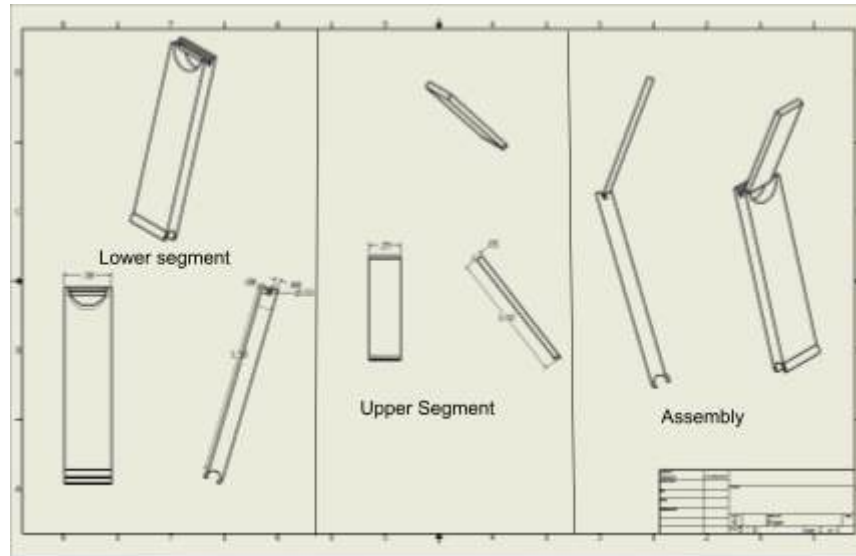
Design Overview:

For any hand design, the base of all functions is in the palm. Each finger is jointed to the palm, and it provides support for the entire limb. For this design, the palm tries to replicate the shape of a human hand while remaining symmetric due to the two thumbs of the design. For the hinge joints of each finger, the pivots for each joint were designed first, accounting for the width of each finger along with extra space on the sides to prevent friction. The three functional fingers each had a pivot on the palm, however the palm only contained the rolling pivot for the thumbs, as the joint would need to be on the platform that rolls with the thumb. The palm, along with most of the parts for this design, are meant to be 3D printable, allowing the design to remain affordable and widely accessible.



Design sketch of palm component including front, top, side, and isometric view

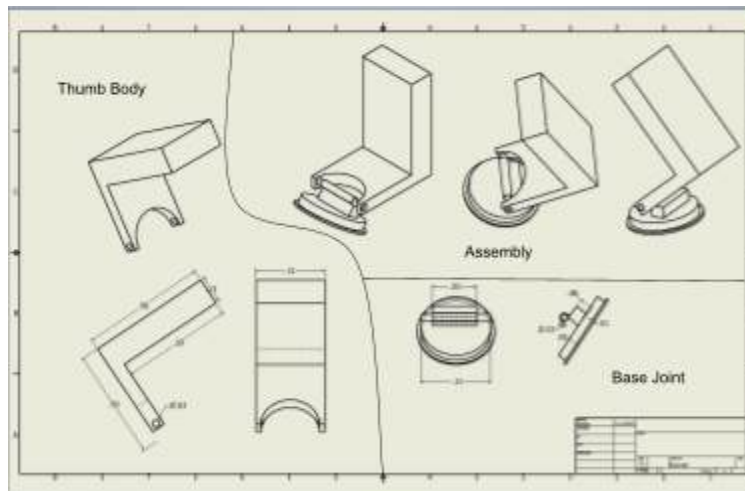
After the palm, the fingers were designed, and in an effort to keep the design simple the three fingers were all identical. The first part of the finger designed was the hook for the connection to the hinge pivot already on the palm. After this, the first segment of the finger was attached to this joint. Each of the two segments of the finger were roughly the same size, each being around half of the full length of a human finger, so that when put together they would accurately represent a real finger. After the first segment was created, a section was cut out and a pivot was placed for the second segment to be connected to, and a similar process to creating the first segment was followed for the second.



Design sketch of finger including front, side and isometric view of lower and upper segments, as well as side and isometric view for the assembly.

Since the thumbs in this design are different from the rest of the fingers because they are only jointed once at the base, the body of the thumb was designed first. Due to the lack of a second joint, two simple segments to the thumb were created with a rigid bend, allowing for wider application of use for the thumb. Due to time and resource constraints, the thumb was approximated as a 90 degree bend. After creating this body, I recognized that the thumb not only had to be on a hinge, but had to be rolling as well. A circular body was created to fit the socket outlined in the palm design, and the hook was mounted to that. Unlike the finger, the pivot itself

was placed on the thumb, because of the way it had to work out with the rolling joint.

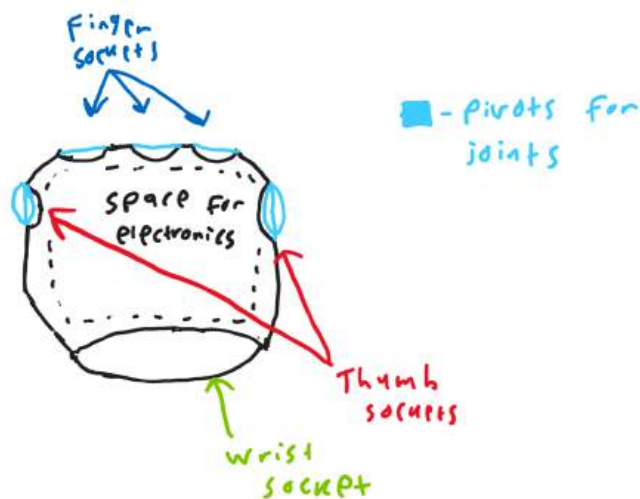


Design sketch of thumb including front, side, and isometric view of the body and assembly, as well as top and side view of the base joint.

Prototypes:

Before the design was ideated in CAD, a few simple design sketches were used to help gauge various details of the design. By sketching each part individually, many specifics became solidified that made the CAD process go much smoother, such as how to implement the various types of joints and the general shape of the design.

Palm



Prototype sketch of Palm component, including designated space for electronics and labels for joint sockets.

Finger

side view



front view



Prototype sketch of finger component, including labels and ideation of joints

Thumb

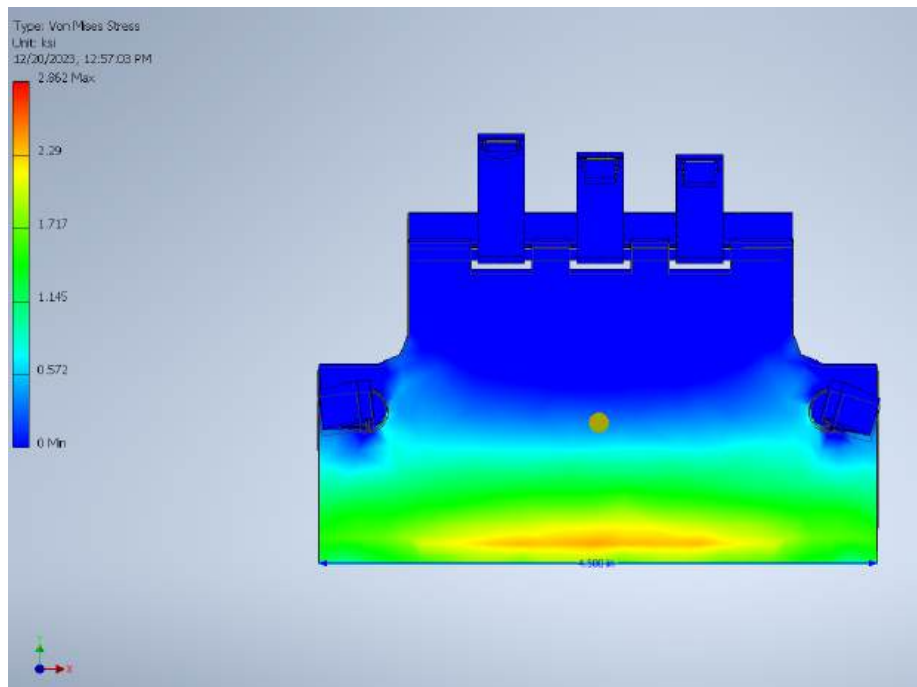


Prototype sketch of thumb, including labeled hinge and rolling joint.

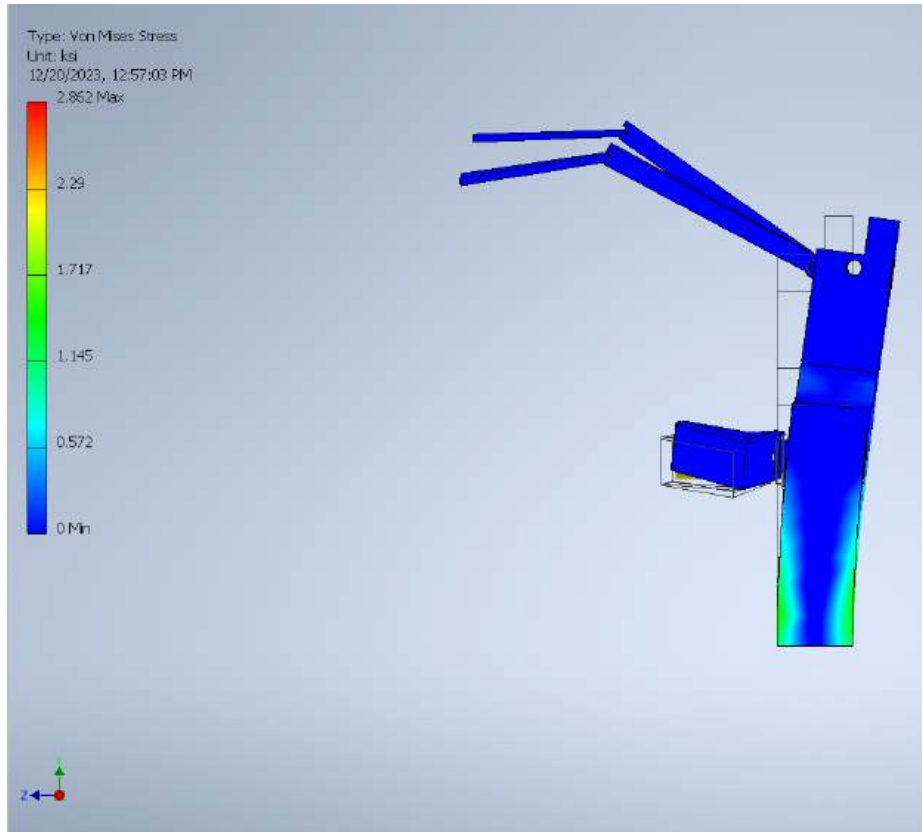
Design Tests:

The range of motion of each component was measured using Autodesk Inventor's Joint function, and mechanical stops were added where needed. The drive joint function in Inventor has a feature that allows detection of collision, allowing calculation of the exact number of degrees that each joint is able to move, from collision on one side to the other. For the joint connecting each finger to the palm, the range of motion was 109.03 degrees. This can be compared to the range of motion for the MCP joint of the human hand, which is around 90 degrees ("Wrist/Hand Active Range of Motion (AROM): Basic Assessment"). The joint between the first and second stages of each finger had a range of motion of 116 degrees, which can be compared to the PIP joint of the human hand, which reaches a maximum range of motion of 115 degrees ("Wrist/Hand Active Range of Motion (AROM): Basic Assessment"). Finally, the hinging of the thumb joint achieved 180 degrees of motion, which can be compared to the native value for the CMC joint of the thumb, which is only 50 degrees ("Wrist/Hand Active Range of Motion (AROM): Basic Assessment"). This additional flexibility in the prosthetic is able to account for the fact that there are fewer joints within each finger.

The prosthetic also needs to remain durable, which was tested by running a stress analysis using the tool incorporated within Autodesk Inventor. The analysis reports values in pressure in terms of ksi(kilopound force per square inch). PLA, the material which would be used to 3D print the design, is able to withstand up to 7.25 ksi (Central Scanning). According to this simulation, the hand would be able to withstand over 500 lbs of force distributed across the palm, although the structure would bend as shown in the images below. The most stress would be applied on the wrist.



Front view of stress analysis of force on palm, with scale on left (ksi)



Side view of stress analysis of force on palm, with scale on left (ksi)

Cost Analysis:

As the design currently stands, the frame of the prosthetic (not including electronics) would be entirely 3D printed. Autodesk inventor calculates the volume to be 5.3 cubic inches, and 3D printer filament is estimated to cost around \$0.75, on average (“Material Cost for Printing”). This results in a cost of around \$4, which is extremely cheap, keeping in mind the lack of electronics. Assuming the cost of electronics will not exceed \$200, the design is far more affordable than many currently available market options.

Conclusions

Design Summary:

The prosthetic hand design is a simple structure involving two thumbs and three fingers, for a total of 5 fingers. Each of the thumbs are identical, and are able to roll on the palm as well as hinge against it. The thumb is modeled by a rigid 90 degree bend, while the fingers each have two joints.

The range of motions for all of the joints measurable in CAD exceeded the comparable native values of a human hand. These results are shown in the table below.

Joint	Native Range of Motion	Prosthetic Range of Motion
-------	------------------------	----------------------------

Thumb (CMC)	40-50°	181.1°
Finger (MCP)	85-90°	109.03°
Finger (PIP)	100-115°	116°

Table comparing native range of motion values to corresponding measured range of motion values for the prosthetic design (“Wrist/Hand Active Range of Motion (AROM): Basic Assessment”).

The stress analysis conducted using Inventor showed that the prosthetic is able to handle over 500 lbs of pressure, although it resulted in a significant amount of bending in the structure of the hand. A simple cost analysis showed that the plastic frame of the design would cost less than \$4 to produce, assuming a 3D printer is available.

Design Shortcomings

While the design was successful, there were several areas in which the design could be improved on to be much more effective. First, simply using 3D printed parts would be ineffective as a hand as they wouldn’t have grip. To address this, rubber caps could be added to the tips of each finger and the palms. The placement of the thumb joint is also suboptimal, as well as the design, since the native hand involves a ball and socket joint rather than a rolling one, although this would be difficult to control using a motor. The stress analysis showed that although the design was able to withstand 500 lbs, there was a significant amount of bending. This could be fixed by reinforcing the structure or making the palm thicker. The stress analysis conducted could also have increased accuracy, as the tool provided in Autodesk Inventor analyzes a force applied to a face, rather than a specific point. This allows the weight to be distributed across the whole palm, which is not realistic in most scenarios. A better way to conduct testing, with available resources, is to prototype the 3D printed frame and conduct experiments with varying amount of weight.

Future Directions

This design is far from being ready for large scale production, and many improvements and additions must be made for the design to be functional. Firstly, the electronics have not yet been accounted for in the design, so creating space for those would be the first priority. This could be done by hollowing out a section inside the palm, potentially creating space for the motors, battery, etc. Following the completion of the design including all electronics and the control system, another cost analysis would need to be conducted to ensure that the design is maintaining the goal of remaining affordable, and any changes to increase affordability should be made. After the design is ideated to a functional state, it would need to go through rigorous testing to ensure durability and quality, which could be done by rolling out a beta version to a few consumers and receiving feedback to make any necessary improvements.

Works Cited

- Burt, Sheila. "Facts about Limb Loss." Shirley Ryan AbilityLab, 4 Apr. 2018, www.sralab.org/research/labs/bionic-medicine/news/facts-about-limb-loss.
- Central Scanning. "Are 3D Printed Parts Strong?" Central Scanning, 20 Oct. 2023, www.central-scanning.co.uk/are-3d-printed-parts-strong/.
- Duncan, Scott F.M., et al. "Biomechanics of the Hand - Hand Clinics." Hand Clinics, 17 Oct. 2013, [www.hand.theclinics.com/article/S0749-0712\(13\)00058-9/fulltext](http://www.hand.theclinics.com/article/S0749-0712(13)00058-9/fulltext).
- LeBlanc, Maurice. "Give Hope - Give a Hand." Stanford, 9 Nov. 2008, web.stanford.edu/class/engr110/2011/LeBlanc-03a.pdf.
- "Material Cost for Printing." ToyBuilder Labs, 14 Dec. 2013, www.toybuilderlabs.com/blogs/news/13055597-material-cost-for-printing.
- Vandersea, Jamie. "The Complete Guide to Arm & Hand Amputations and Prosthetics: MCOP." MCOP Prosthetics, 1 Apr. 2020, mcopro.com/blog/resources/arm-hand-prosthetics/.
- Ventimiglia, Paul. Worcester Polytechnic Institute, pp. 1–74, Design of a Human Hand Prosthesis, <https://digital.wpi.edu/downloads/pk02cc64v?locale=en>.
- "Wrist/Hand Active Range of Motion (AROM): Basic Assessment." Physiotutors, 8 Nov. 2022, www.physiotutors.com/wiki/wrist-hand-active-range-of-motion/.

Schrödinger's Equation: Variations, Potentials, and Modern Applications

Neev Jain
Supervisor: Mo Vali

2480 words

Abstract

Schrödinger's wave equation stands as a monumental foundation in quantum mechanics, presenting a framework that describes the evolution of quantum states across time. This research delves into the dynamic nature of Schrödinger's equation, emphasizing its adaptability when applied to various potential structures. The equation's utility is explored across different contexts, from infinite and finite potential wells to potential barriers and three-dimensional problems, each illustrating unique quantum phenomena such as discrete energy levels, quantum tunneling, and the probabilistic distribution of particles. Furthermore, this paper highlights the equation's application in avant-garde research fields such as quantum computing, where it provides essential insights into qubit behavior, and in material science, guiding the development of novel materials with bespoke quantum properties.

Contents

1 Introduction	1	4.2 Material Science and Nanotechnology	7
1.1 The Quantum Revolution	1	4.3 Drug Discovery and Molecular Biology	7
1.2 Introduction to Schrödinger's Equation	2	4.4 Chemical Reaction Dynamics	7
1.3 Electrons as Both Waves and Particles: Introducing the Wavefunction	2	4.5 Astrophysics and Cosmology	8
1.4 Born Rule, Probabilities, and Normalization	3		
2 Solutions to Schrödinger's Equation in Different Potentials	3		
2.1 Motivation	3		
2.2 Infinite Potential Well	3		
2.3 Finite Potential Well	4		
2.4 Potential Step with ($E > V$) and ($E < V$)	4		
2.4.1 Scenario ($E > V$)	4		
2.4.2 Scenario ($E < V$)	5		
2.5 Finite Barrier	5		
2.6 3D Schrödinger Equation	5		
3 Density Functional Theory (DFT)	6		
3.1 Introduction to Density Functional Theory	6		
3.2 Theoretical Foundations	6		
3.3 Many-Body Problem and Independent Electron Approximation	6		
3.4 Inclusion of Electron-Electron Interactions	6		
3.5 Kohn-Sham DFT and the Exchange-Correlation Potential	6		
3.6 Exchange-Correlation Functional Approximations	7		
3.7 Self-Consistent Solution of Kohn-Sham Equations	7		
4 Modern Applications in Contemporary Research	7		
4.1 Quantum Computing and Information Theory	7		

1 Introduction

1.1 The Quantum Revolution

Quantum mechanics, a realm where the very fabric of nature reveals its probabilistic and wave-like character, underwent a revolutionary transformation with the advent of Schrödinger's wave equation in the early 20th century. This transformation was not just a leap in understanding the microscopic world but also a foundational shift in how physical reality is conceptualized. Schrödinger's equation, embodying the wave-particle duality, serves as a linchpin in the vast domain of quantum mechanics, offering a mathematical framework to predict the behavior of particles at the quantum level.

The narrative of quantum mechanics is punctuated by profound discoveries that challenged the classical conceptions of physics [1]. Before Schrödinger introduced his wave theory, our understanding of the tiny quantum world was shaped by Planck's ideas about energy coming in packets and by what we learned from the photoelectric effect, which Einstein explained (see Figure 1). The photoelectric effect is an experiment where light hits a surface and knocks out electrons, but only if the light is strong enough. This showed us that light acts like a bunch of tiny particles, called photons, and helped set the stage for the quantum theories that came later [1-3]. However, it was the formulation of the wave equation by Erwin Schrödinger in 1925 that provided a comprehensive mathematical structure for quantum mechanics. Schrödinger's equation emerged not only as a

theoretical construct but as a practical tool to unravel the mysteries of atomic and subatomic particles.

In this tumultuous era, the concept of wave-particle duality became a cornerstone of quantum theory, thanks to the pivotal experiments and theories proposed by physicists like Heisenberg, Bohr, and de Broglie [4,5]. These developments led to a new understanding of matter and energy, fundamentally altering our perception of the universe at its most elemental level.

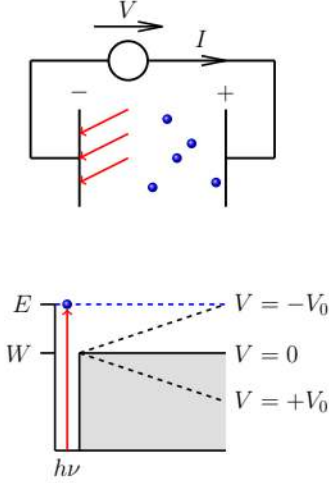


Figure 1: A diagram illustrating the photoelectric effect experiment. This experiment validated Einstein’s quantization hypothesis in which energy is discretized in packets, more commonly known as photons, with each photon carrying an energy of $h\nu$.

1.2 Introduction to Schrödinger’s Equation

In the midst of the quantum revolution, amidst the flourishing of ideas that would drastically alter our comprehension of the microscopic world, Erwin Schrödinger in 1925 proposed an equation that stands as one of the most significant contributions to the field of quantum mechanics. This section delves into the historical backdrop leading to the inception of Schrödinger’s wave equation and provides a concise derivation, illustrating its foundational role in quantum theory.

The journey to Schrödinger’s equation began with the failure of classical physics to explain phenomena at the atomic and subatomic levels [1]. Experiments such as the photoelectric effect, blackbody radiation, and the discrete lines of atomic spectra could not be accounted for within the classical framework. This led to the birth of quantum mechanics, a theory grounded in the quantization of physical properties. Max Planck’s hypothesis of quantized energy levels and Albert Einstein’s explanation of the photoelectric effect laid the groundwork [2,3]. However, it was Schrödinger’s wave equation that provided a comprehensive mathematical framework for quantum mechanics.

Schrödinger sought to describe particles not as discrete entities but as wave-like structures spread out in space. His inspiration came from the classical equation of a vibrating string, which is a partial differential equation describing the displacement of points on a string as a function of time. Schrödinger postulated that the behavior of quantum particles could be described by a similar wave equation. The

derivation of Schrödinger’s equation starts from the classical energy conservation equation, with a quantum perspective:

$$E = \frac{p^2}{2m} + V$$

where E is the total energy, p is the momentum, m is the mass of the particle, and V represents the potential energy. In quantum mechanics, energy and momentum are represented as operators. The momentum operator, in one dimension, is given by $\hat{p} = -i\hbar\frac{\partial}{\partial x}$, where \hbar is the reduced Planck’s constant and i is the square root of -1 . Substituting the quantum mechanical operators into the classical energy equation yields:

$$\hat{E}\Psi = \left[\frac{-\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V \right] \Psi$$

Considering that the energy operator corresponds to the time derivative, $\hat{E} = i\hbar\frac{\partial}{\partial t}$, we obtain the time-dependent Schrödinger wave equation:

$$i\hbar\frac{\partial\Psi}{\partial t} = \left[-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(x,t) \right] \Psi$$

This equation describes the evolution of the wave function Ψ , which encodes the probabilistic nature of the quantum state of a particle. For time-independent potentials, separating the variables allows for the derivation of the time-independent Schrödinger equation, essential for analyzing the stationary states of a system:

$$-\frac{\hbar^2}{2m} \frac{\partial^2\Psi}{\partial x^2} + V\Psi = E\Psi$$

Schrödinger’s wave equation revolutionized physics by providing a comprehensive mathematical framework for quantum mechanics, enabling the prediction of quantum phenomena with unprecedented accuracy. Through its solutions, the wave equation elucidates the probabilistic nature of particles’ behavior at the quantum level, offering insights into the fundamental principles governing the microscopic world.

1.3 Electrons as Both Waves and Particles: Introducing the Wavefunction

The concept that electrons exhibit both wave and particle characteristics is a fundamental principle in quantum mechanics, known as wave-particle duality [4]. This principle transcended classical physics’s distinct categories of ‘wave’ and ‘particle’ and fostered a new understanding of quantum entities. In 1924, Louis de Broglie proposed that particles could exhibit wave-like properties, suggesting that matter could behave similarly to light, which was already known to display both wave and particle traits.

Experimental evidence for the electron’s dual nature was famously provided by the double-slit experiment (see Figure 2). When electrons pass through two closely spaced slits and are detected on a screen, they form an interference pattern indicative of wave-like behavior. Yet, when measured individually, electrons reveal their particle-like aspects, imparting discrete impacts on the detecting screen.

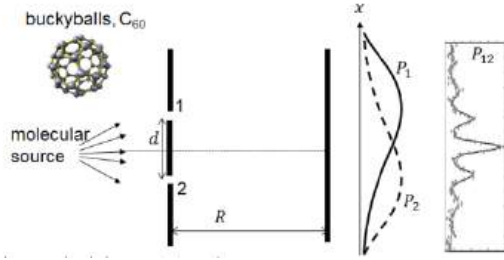


Figure 2: A diagram of the double-slit experiment utilizing buckyballs (C_{60} molecules) to demonstrate wave-particle duality. The diagram depicts the setup where a beam of buckyballs is directed towards a barrier with two slits, resulting in an interference pattern on the screen behind the barrier. This pattern of alternating light and dark fringes provides evidence for the wave-like behavior of large molecules, akin to the interference observed with light and electrons.

The mathematical representation of an electron's wave-like characteristics is encapsulated in the wavefunction Ψ , a complex function that contains all the information about the quantum state of an electron. The wavefunction is a core concept in quantum mechanics, explicitly formulated by the Schrödinger wave equation:

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, t) = \left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right] \Psi(\mathbf{r}, t)$$

Here, \hbar signifies the reduced Planck's constant, m represents the mass of the electron, ∇^2 is the Laplacian operator affecting spatial variables and describing how the space around the electron influences its behavior, and $V(\mathbf{r})$ is the potential energy as a function of position.

1.4 Born Rule, Probabilities, and Normalization

The Born Rule is the postulate that connects these abstract wave functions to concrete probabilities [6]. It provides the probability density, $P(\mathbf{r}, t)$, that a particle is observed at the position \mathbf{r} and time t , which is given by the squared magnitude of the wave function:

$$P(\mathbf{r}, t) = |\Psi(\mathbf{r}, t)|^2$$

This squared modulus translates complex probability amplitudes into real, positive values that correspond to probabilities.

Normalization imposes the condition that the total probability of finding a particle somewhere in space is one:

$$\int |\Psi(\mathbf{r}, t)|^2 dV = 1$$

To maintain this constraint, wave functions must be properly normalized when they are used to predict physical phenomena. This is especially important when the wave function changes due to different potential energy functions $V(\mathbf{r})$ in Schrödinger's wave equation:

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, t) = \left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right] \Psi(\mathbf{r}, t)$$

Different potential energy shapes lead to different wave functions, and thus, to different probability distributions.

2 Solutions to Schrödinger's Equation in Different Potentials

2.1 Motivation

Understanding the solutions of the Schrödinger equation under different potentials is crucial for several reasons [7]:

Understanding Electrons in Materials

At the heart of condensed matter physics is the study of electron behavior in various materials. The Schrödinger's equation, by predicting how electrons behave under different potential energies, enables us to understand fundamental properties of materials such as conductivity, magnetism, and superconductivity. This understanding is crucial for explaining phenomena like the quantum Hall effect and topological insulators.

Designing New Materials

The ability to predict the behavior of electrons under different potentials allows scientists and engineers to design new materials with desired properties. This is especially important in the development of semiconductors, photovoltaic cells, and batteries. By manipulating the potential landscapes within materials, we can tailor their electrical, optical, and magnetic properties, leading to innovations in technology and industry.

Serving as a Base to Solve More Complicated Scenarios

The Schrödinger's equation, in its simplicity, forms the groundwork for tackling more complex quantum mechanical problems [8]. Solutions in well-understood potential scenarios serve as a stepping stone for approximating solutions in more complicated, real-world situations. Techniques such as perturbation theory and the variational principle rely on these simpler solutions to approximate behaviors in more complex systems, including interactions between multiple particles or in the presence of external fields.

2.2 Infinite Potential Well

An infinite potential well, also known as a particle in a box (see Figure 3), is a fundamental problem in quantum mechanics that illustrates the concept of quantized energy levels.

Schrödinger's Equation and Boundary Conditions: The time-independent Schrödinger equation in one dimension is [8]:

$$-\frac{\hbar^2}{2m} \frac{d^2 \Psi}{dx^2} + V(x) \Psi = E \Psi$$

For $0 < x < a$, where $V(x) = 0$, the equation simplifies to:

$$\frac{d^2 \Psi}{dx^2} = -\frac{2mE}{\hbar^2} \Psi = -k^2 \Psi$$

where $k = \sqrt{\frac{2mE}{\hbar^2}}$. The general solution is:

$$\Psi(x) = A \sin(kx) + B \cos(kx)$$

Applying the boundary conditions $\Psi(0) = \Psi(a) = 0$ leads to the quantization of energy levels:

$$k = \frac{n\pi}{a}$$

$$E_n = \frac{n^2 \pi^2 \hbar^2}{2ma^2}$$

where $n = 1, 2, 3, \dots$ denotes the quantum number.

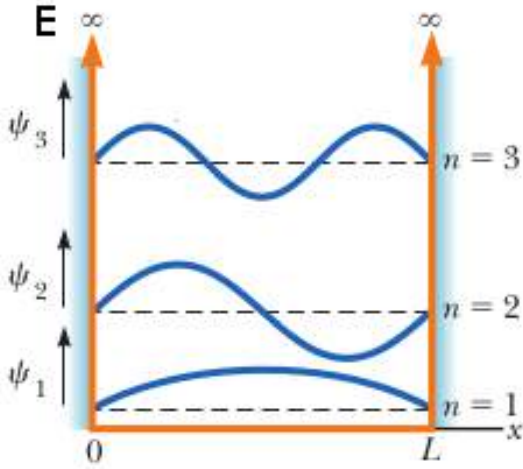


Figure 3: An infinite potential well, a conceptual model used in quantum mechanics to describe a particle that is confined to a region of space with impenetrable boundaries. The well is depicted with vertical lines representing the boundaries, which the particle cannot pass through, and a flat bottom indicating a region with zero potential energy. Within the well, the particle can only occupy certain energy levels, shown as horizontal lines, which correspond to the allowed quantum states of the particle.

2.3 Finite Potential Well

The finite potential well represents a scenario where a particle is trapped in a well with finite walls (see Figure 4). Unlike the infinite potential well, the particle has a non-zero probability of existing outside the well due to quantum tunneling.

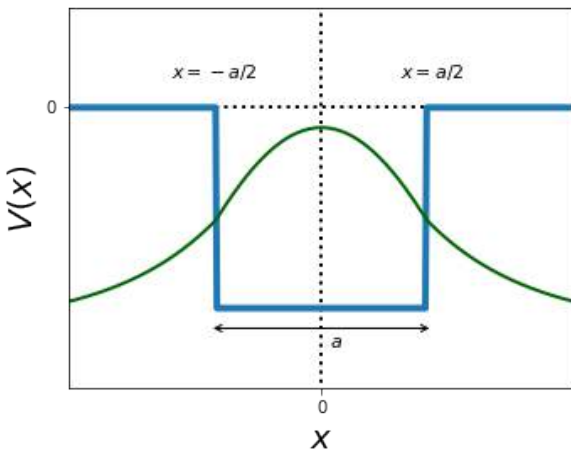


Figure 4: A graphical representation of a finite potential well, a quantum mechanical model used to describe a particle confined to a potential trough where it is energetically favorable to reside. Unlike the infinite potential well, the boundaries here are represented by finite potential barriers, suggesting that the particle has a non-zero probability of tunneling through these barriers.

The time-independent Schrödinger equation remains:

$$-\frac{\hbar^2}{2m} \frac{d^2\Psi}{dx^2} + V(x)\Psi = E\Psi$$

For $E < V_0$, within the well $0 < x < a$, the wavefunction solutions are given by a combination of sin and cos functions as in the infinite potential case. However, in the regions where the potential is non-zero, the solution takes the form:

$$\Psi(x) = Ce^{-\alpha x} + De^{\alpha x}$$

where $\alpha = \sqrt{\frac{2m(V_0-E)}{\hbar^2}}$.

Boundary conditions at $x = 0$ and $x = a$ and the continuity of Ψ and $\frac{d\Psi}{dx}$ at the walls of the well give us the allowed energy levels, which remain discrete but less straightforward than the infinite well case.

2.4 Potential Step with ($E > V$) and ($E < V$)

For a particle encountering a potential step, there are two scenarios depending on whether the energy of the particle (E) is greater than or less than the step potential (V).

2.4.1 Scenario ($E > V$)

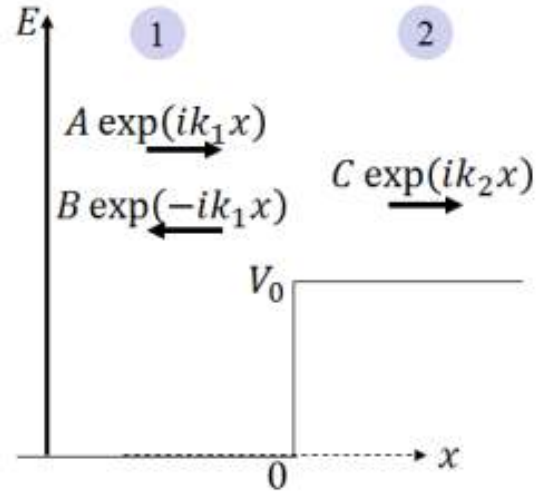


Figure 5: This figure illustrates a quantum mechanical scenario where the energy E of a particle is greater than the potential energy V of the barrier it encounters.

The wavefunction solutions on either side of the potential step are:

$$\Psi_1(x) = Ae^{ik_1 x} + Be^{-ik_1 x} \quad \text{for } x < 0$$

$$\Psi_2(x) = Ce^{ik_2 x} \quad \text{for } x > 0$$

where k_1 and k_2 relate to the energy and potential as $k_1 = \sqrt{\frac{2mE}{\hbar^2}}$ and $k_2 = \sqrt{\frac{2m(E-V)}{\hbar^2}}$. The coefficients A , B , and C are found using boundary conditions.

2.4.2 Scenario ($E < V$)

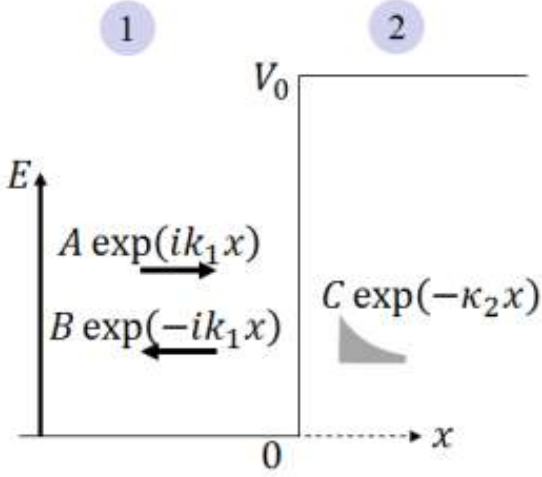


Figure 6: This figure depicts a quantum mechanical scenario where the energy of a particle, denoted by E , is less than the potential energy V of the barrier. In classical physics, a particle with insufficient energy would be unable to surmount the barrier. However, quantum mechanics allows for the phenomenon of quantum tunneling, where there is a probability that the particle can penetrate and cross the barrier despite having lower energy.

Here, the wavefunction within the region of the step ($x > 0$) decays exponentially as the particle exhibits tunneling:

$$\Psi_2(x) = D e^{-\kappa x}$$

where $\kappa = \sqrt{\frac{2m(V-E)}{\hbar^2}}$.

Boundary conditions are applied at the interface ($x = 0$) to match the wavefunctions and their derivatives.

2.5 Finite Barrier

This scenario describes a finite height and width barrier (see Figure 7) which the particle may tunnel through, reflect off, or transmit depending on its energy relative to the barrier height.

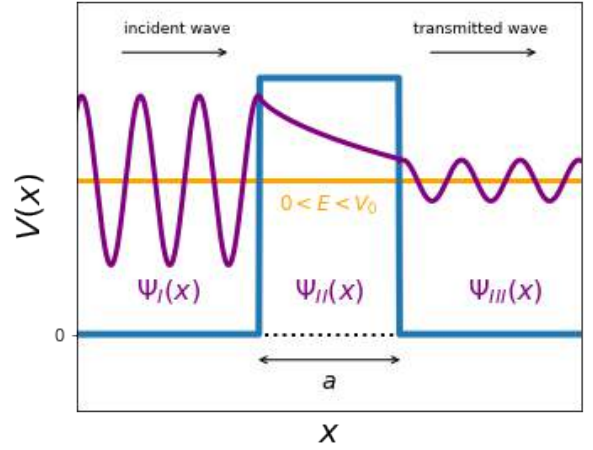


Figure 7: This figure illustrates the concept of a finite barrier potential in quantum mechanics. It shows a barrier with a specific height, representing the potential energy that a quantum particle must overcome. If a particle's energy is lower than the barrier height, it may still penetrate the barrier through quantum tunneling.

The time-independent Schrödinger equation is applied to both the free regions ($x < b$) and ($x > c$) and the barrier region ($b < x < c$). The general form of the wavefunction in each region is matched at the boundaries of the barrier, illustrating tunneling and reflection phenomena. Energy levels are not quantized in this scenario.

2.6 3D Schrödinger Equation

In three dimensions, Schrödinger's equation needs to accommodate the additional degrees of freedom and is given by [8]:

$$-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \Psi(\mathbf{r}, t) + V(\mathbf{r})\Psi(\mathbf{r}, t) = E\Psi(\mathbf{r}, t)$$

The solutions to the 3D Schrödinger equation are more complex because they depend on the specific form of the potential $V(\mathbf{r})$ [8]. For simple geometries like a cubic box, the solutions are separable in Cartesian coordinates (see Figure 8). For spherical symmetries, as in atoms, solutions are expressed in terms of spherical harmonics and radial functions.

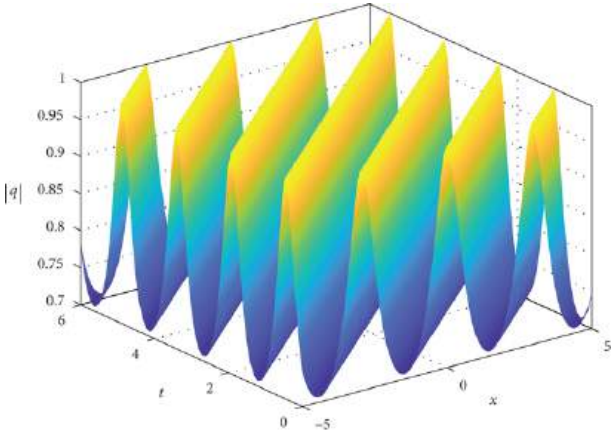


Figure 8: This figure provides an example of a three-dimensional (3D) wave function solution derived from Schrödinger's equation, which is central to quantum mechanics. The visualization captures the complex spatial form of a quantum wave function, with peaks and troughs representing the areas where the particle described by the wave function is more or less likely to be found.

3 Density Functional Theory (DFT)

3.1 Introduction to Density Functional Theory

Density Functional Theory (DFT) is a computational quantum mechanical theory used to investigate the electronic structure of many-body systems. It simplifies the complex problem of interacting electrons by focusing on the electron density, a function of three spatial coordinates, making it an essential tool for condensed matter physics, material science, and chemistry [9].

3.2 Theoretical Foundations

The journey from the many-body Schrödinger equation to DFT begins by considering that materials consist of electrons and nuclei. The Schrödinger equation for a system of interacting particles includes kinetic energy terms for both electrons and nuclei, as well as potential energy terms representing electron-electron, nucleus-nucleus, and electron-nucleus interactions. The many-body Hamiltonian is expressed as:

$$\hat{H} = -\frac{\hbar^2}{2m_e} \sum_i \nabla_i^2 - \frac{\hbar^2}{2M} \sum_I \nabla_I^2 + \sum_{i,I} \frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|}$$

The Born-Oppenheimer approximation allows us to focus on the electronic problem by assuming that the nuclear motion can be separated due to the large mass difference between electrons and nuclei. This leads to an electronic Hamiltonian:

$$\hat{H}_{el} = -\frac{\hbar^2}{2m_e} \sum_i \nabla_i^2 + \sum_{i,I} \frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}$$

3.3 Many-Body Problem and Independent Electron Approximation

Solving the electronic Hamiltonian directly is intractable for systems with more than a few electrons. The independent electron approximation simplifies the problem by assuming that electrons do not interact with each other, which eliminates the electron-electron interaction term. However, this approximation ignores electron correlation, which is significant in real materials.

3.4 Inclusion of Electron-Electron Interactions

The Hartree-Fock method reintroduces electron-electron interactions by including a Fock potential, which accounts for the average effect of electron repulsion. It is expressed as:

$$V_{\text{Fock}}(\mathbf{r}_i) = \sum_j \left(\int \frac{e^2 |\psi_j(\mathbf{r}_j)|^2}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_j - \int \frac{e^2 \psi_j^*(\mathbf{r}_j) \psi_i(\mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|} d\mathbf{r}_j \right)$$

This potential includes the Hartree term (classical electron-electron repulsion) and the exchange term (quantum mechanical exchange interaction). However, it still does not fully account for electron correlation.

3.5 Kohn-Sham DFT and the Exchange-Correlation Potential

The Kohn-Sham method of DFT introduces an effective potential $V_{\text{eff}}(\mathbf{r})$ that simplifies the many-body problem into a set of single-particle equations. This effective potential includes contributions from the external potential, the Hartree potential, and the exchange-correlation potential. The Kohn-Sham equations are:

$$\left[-\frac{\hbar^2}{2m_e} \nabla^2 + V_{\text{eff}}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r})$$

The effective potential $V_{\text{eff}}(\mathbf{r})$ is defined as the sum of the external potential $V_{\text{ext}}(\mathbf{r})$, which is typically due to the nuclei, the Hartree potential $V_{\text{H}}(\mathbf{r})$, and the exchange-correlation potential $V_{\text{xc}}(\mathbf{r})$:

$$V_{\text{eff}}(\mathbf{r}) = V_{\text{ext}}(\mathbf{r}) + V_{\text{H}}(\mathbf{r}) + V_{\text{xc}}(\mathbf{r})$$

The Hartree potential $V_{\text{H}}(\mathbf{r})$ accounts for the classical electron-electron repulsion and is given by:

$$V_{\text{H}}(\mathbf{r}) = \int \frac{e^2 \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'$$

The exchange-correlation potential $V_{\text{xc}}(\mathbf{r})$ captures the complex quantum mechanical interactions between electrons beyond the classical repulsion and is derived from the exchange-correlation energy functional $E_{\text{xc}}[\rho]$ as its functional derivative with respect to the electron density:

$$V_{\text{xc}}(\mathbf{r}) = \frac{\delta E_{\text{xc}}[\rho]}{\delta \rho(\mathbf{r})}$$

By solving the Kohn-Sham equations self-consistently, we obtain a set of orbitals that are used to construct the ground-state electron density, which in turn allows us to determine the electronic properties of the system.

3.6 Exchange-Correlation Functional Approximations

The exchange-correlation functional $E_{xc}[\rho]$ is approximated using various models. The Local Density Approximation (LDA) and the Generalized Gradient Approximation (GGA) are two widely used approximations (see Figure 9). LDA assumes that the exchange-correlation energy at each point in space is that of a uniform electron gas:

$$E_{xc}^{\text{LDA}}[\rho] = \int \rho(\mathbf{r}) \epsilon_{xc}^{\text{hom}}(\rho(\mathbf{r})) d\mathbf{r}$$

GGA refines this by including the gradient of the electron density:

$$E_{xc}^{\text{GGA}}[\rho] = \int f(\rho(\mathbf{r}), \nabla\rho(\mathbf{r})) d\mathbf{r}$$

where f is a function that accounts for the inhomogeneity of the electron density.

3.7 Self-Consistent Solution of Kohn-Sham Equations

The Kohn-Sham DFT formalism requires a self-consistent solution of the Kohn-Sham equations. Starting with an initial guess for the electron density, one solves the Kohn-Sham equations to obtain a new set of orbitals $\psi_i(\mathbf{r})$, from which a new electron density is constructed:

$$\rho(\mathbf{r}) = \sum_i |\psi_i(\mathbf{r})|^2$$

This process is iterated until the input and output densities converge to a predetermined accuracy, yielding the ground-state electron density and energy.

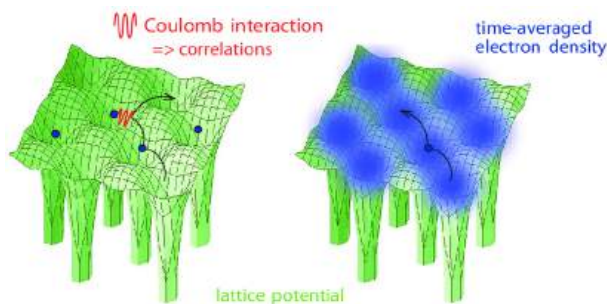


Figure 9: The many-body Hamiltonian schematically visualized. Black arrows represent the kinetic energy and the background ionic potential is shown in green. The Coulomb interaction (red) correlates the movements of the electrons. On the right, the dramatic simplification achieved by LDA is depicted. Each electron moves independently within a time-averaged local density of the other electrons.

4 Modern Applications in Contemporary Research

4.1 Quantum Computing and Information Theory

The Schrödinger wave equation and the extensions of it play a foundational role in quantum computing and information

theory. By describing the quantum states of particles, it allows for the development of quantum algorithms that leverage superposition and entanglement to solve problems intractable for classical computers. Research in this field often focuses on the development of quantum bits (qubits), quantum error correction, and quantum cryptography [10,11].

4.2 Material Science and Nanotechnology

In material science and nanotechnology, the Schrödinger equation is used to model the electronic properties of novel materials [7]. This includes the design of semiconductors, nanostructures, and photovoltaic materials (see Figure 10). The predictive power of quantum mechanical calculations based on Schrödinger's equation enables the discovery and optimization of materials with specific electronic, magnetic, and optical properties.

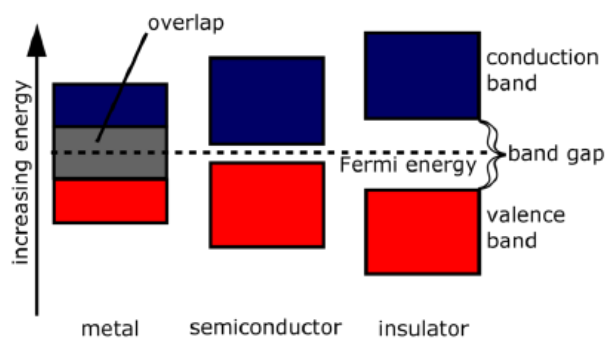


Figure 10: Schematic illustration of a semiconductor's electronic band structure. The valence band (VB) represents filled electron states, and the conduction band (CB) represents empty states. The energy difference between the VB and the CB is the bandgap (E_g), indicated by the horizontal gap. The Fermi level (E_F) is the energy at which the probability of finding an electron is 50%.

4.3 Drug Discovery and Molecular Biology

Schrödinger's wave equation is a critical tool in computational chemistry and molecular biology, particularly in the field of drug discovery [12]. Quantum mechanical simulations help in understanding the interactions between drugs and their target molecules, predicting the binding affinities, and identifying potential side effects. This approach accelerates the drug development process by reducing the reliance on empirical testing.

4.4 Chemical Reaction Dynamics

The study of chemical reaction dynamics relies heavily on the Schrödinger equation to model the behavior of electrons (see Figure 11) during chemical reactions [13,14]. This includes the exploration of reaction mechanisms, transition states, and the prediction of reaction rates. Quantum dynamics simulations provide insights into the fundamental processes that govern chemical reactivity.

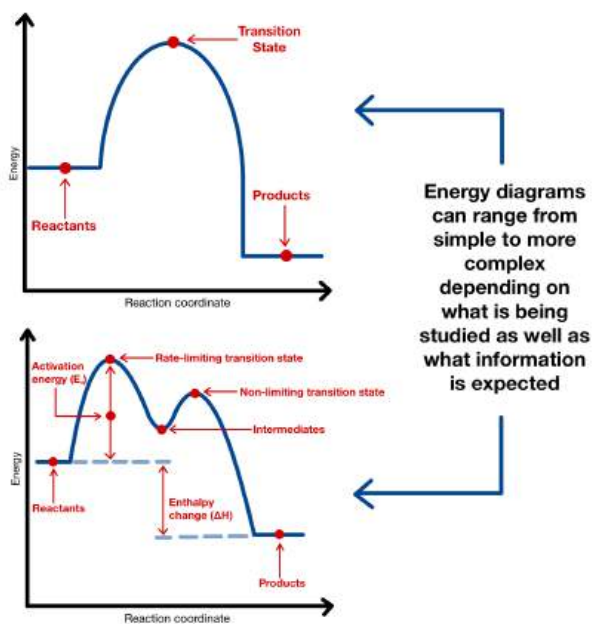


Figure 11: A graphical representation of the potential energy surface for a chemical reaction. The reactants (R) and products (P) are shown on the x-axis, with their corresponding potential energies on the y-axis. The transition state (TS) represents the highest energy point along the reaction pathway, requiring an activation energy (E_a) for the reaction to proceed. The arrows indicate the overall energy change (ΔE) during the reaction.

4.5 Astrophysics and Cosmology

Astrophysics and cosmology benefit from the Schrödinger equation in the study of quantum phenomena in extreme environments, such as white dwarfs and neutron stars [15,16]. The equation is also used in the context of dark matter research and the search for explanations of quantum gravitational effects in the early universe.

References

- [1] C. Johnson, “Mathematical physics of blackbody radiation,” *Icarus iDucation*, 2012.
- [2] R. A. Millikan, “A direct photoelectric determination of Planck’s constant,” *Physical Review*, vol. 7, pp. 355–388, 1916.
- [3] A. Einstein, “On a heuristic point of view about the creation and conversion of light,” in *The Old Quantum Theory*, pp. 91–107, Elsevier, 1967.
- [4] L. de Broglie, “Recherches sur la théorie des quanta,” *Migration - université en cours d’affectation*, 1924.
- [5] N. Bohr, “I. On the constitution of atoms and molecules,” *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 26, pp. 1–25, 1913.
- [6] M. Born, “Zur quantenmechanik der stoßvorgänge,” *Zeitschrift für Physik*, vol. 37, pp. 863–867, 1926.
- [7] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*. Holt, Rinehart and Winston, 1976.
- [8] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum Mechanics*. Wiley, 1977.
- [9] A. J. Cohen, P. Mori-Sánchez, and W. Yang, “Insights into current limitations of density functional theory,” *Science*, vol. 321, pp. 792–794, 2008.
- [10] A. W. Harrow, A. Hassidim, and S. Lloyd, “Quantum algorithm for linear systems of equations,” *Physical Review Letters*, vol. 103, p. 150502, 2009.
- [11] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, “Quantum error correction via codes over $GF(4)$,” 1997.
- [12] W. L. Jorgensen, “The many roles of computation in drug discovery,” *Science*, vol. 303, pp. 1813–1818, 2004.
- [13] G. Worth, “Non-adiabatic dynamics,” *Faraday Discussions*, 2004.
- [14] S. Gómez, E. Spinlove, and G. Worth, “Benchmarking non-adiabatic quantum dynamics using the molecular tully models,” *Physical Chemistry Chemical Physics*, vol. 26, pp. 1829–1844, 2024.
- [15] G. Fontaine, P. Brassard, and P. Bergeron, “The potential of white dwarf cosmochronology,” *Publications of the Astronomical Society of the Pacific*, vol. 113, pp. 409–435, 2001.
- [16] J. A. Faber and F. A. Rasio, “Hydrodynamics of neutron star mergers,” in *AIP Conference Proceedings*, vol. 575, pp. 130–142, 2001.

The Optimization of Neural Networks

Zubin Gupta¹, Cyrus Ayubcha²

1. Lexington High School, 51 Waltham St, Lexington, Massachusetts, 02421, United States

2. Harvard University, Massachusetts Hall, Cambridge, Massachusetts, 02138, United States

* Corresponding author email: zubin.gupta.mail@gmail.com

Abstract

A neural network is a type of Artificial Intelligence consisting of multiple layers of neurons. Neural networks can be trained to perform a specific task, modeled after the human brain. This paper will discuss the efficiency problem in algorithms containing neural networks. It will demonstrate how to optimize these neural networks by creating a neural network and measuring its accuracy and speed on specific tests. Afterward, some neural network features will be changed and measured again. Eventually, the data will be enough to create a more optimized neural network that will be more accurate and precise than the others. One of the most influential factors in the speed of a neural network is its max pooling size or the number of values pooled together into one. Another result is that most other modifications, like dropout size or number of convolution layers, do not significantly affect the accuracy and speed of the neural network.

Keywords

Robotics; Intelligent Machines; Machine Learning; Convolution Layer; Neural Network



Introduction

Through the recent surge in the development and usage of Artificial Intelligence (AI), computer scientists have been able to create machine learning algorithms for various applications. However, the upkeep of large models can be expensive, so it is necessary to ensure that the value of money is maximized for a large algorithm.^{1,2} We can do that by changing the attributes of the neural network. If it were to work, it would be an efficient solution that does not require additional money spent.² That brings us to the question: How does changing the attributes of a neural network change its efficiency?

A convolutional network is a type of AI that turns all the pixels in an image into a 3d matrix of values going from 0 to 1, puts the values through different layers, and calculates values based on that.¹ Some values include the convolution layer, which takes the dot product of each 2d value matrix and one created by the algorithm, resulting in another 3d. Afterward, there is the batch normalization layer, which essentially rescales the weights that modify the matrix of values. The max pooling layer follows after, which takes subsets of a specific area out of the value matrix, finds the maximum value for each subset, and uses those values to construct another matrix.^{1,3} This process is repeated for all matrices. The dropout layer comes next: its job is to drop out random values in the neural network.²

Those four layers can be repeated multiple times, but not infinitely, as each sequence through the four layers makes the matrix of values smaller.¹ After that, a few more layers are applied. The flattening first converts the 3d matrix into a 1d vector. The dense layer is next, and weights are applied to the vector to make the data smaller. A batch normalization layer follows, and there is one more dense layer, which gives the output.^{1,2}

While optimization studies have long attempted to develop complex methods and structures to improve training efficiency while retaining performance, this study utilizes the Keras package's most widely used convolutional neural network structures. This study will augment common hyperparameters and CNN size and assess performance in various datasets.

I can optimize a neural network without paying or waiting excessive time. Other companies can look at this research and determine what would be best for their models. Even if my example could be more satisfactory for some companies due to my algorithm's specific purpose, computer scientists can observe my research and even attempt a scaled-down version of their AI similarly, leading to sufficient results. In sum, my goal is to catalyze the growth of AI by giving an example of optimizing a neural network.

Methods

Other key terms include the epoch, which is how long the model is trained and tested. The batch size is the number of images propagated through the neural network.

The program used in this study requires a convolutional neural network to classify different images. This means a dataset and a convolutional neural network are required.^{3,4} The experiment requires a neural network with set attributes and progressive modifications over time. The resulting data will be kept in a spreadsheet for further analysis.

Python will be used as the coding language in the program, and the TensorFlow Keras library will be used for the neural network. A traffic light classification dataset will be used initially. The program's goal is to see what percentage of traffic lights it can adequately identify after given some information about the traffic lights.

Using the time taken, I will measure the approximate average time per epoch for the dataset. This is because if I run a different number of epochs on each neural network, the timing will not be unbiased. I will use the peak validation accuracy achieved at any time in the neural network as the accuracy, as there would be no reason to use a less accurate neural network if a more accurate one already exists.⁴

Results and Discussion

In this dataset, the neural network has to identify if a traffic light has a yellow light, a red light, a green light, or none. The experiment begins using three convolution layers, a batch size of 15, and a max pooling size of (2,2) (Figure 1). This will serve as the baseline for the project.

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 126, 126, 32)	896
batch_normalization (Batch Normalization)	(None, 126, 126, 32)	128
max_pooling2d (MaxPooling2D)	(None, 63, 63, 32)	0
dropout (Dropout)	(None, 63, 63, 32)	0
conv2d_1 (Conv2D)	(None, 61, 61, 64)	18496
batch_normalization_1 (Batch Normalization)	(None, 61, 61, 64)	256
max_pooling2d_1 (MaxPooling2D)	(None, 30, 30, 64)	0
dropout_1 (Dropout)	(None, 30, 30, 64)	0
conv2d_2 (Conv2D)	(None, 28, 28, 128)	73856
batch_normalization_2 (Batch Normalization)	(None, 28, 28, 128)	512
max_pooling2d_2 (MaxPooling2D)	(None, 14, 14, 128)	0
dropout_2 (Dropout)	(None, 14, 14, 128)	0
flatten (Flatten)	(None, 25088)	0
dense (Dense)	(None, 512)	12845568
batch_normalization_3 (Batch Normalization)	(None, 512)	2048
dropout_3 (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 4)	2052

=====
Total params: 12,943,812
Trainable params: 12,942,340
Non-trainable params: 1,472
=====

Figure 1 - Model 1 (optimizer = rmsprop, batch size = 15)

```
Epoch 1/50
142/142 [=====] - ETA: 0s - loss: 0.2887 - accuracy: 0.9220WARNING:tensorflow:Learning rate reduction is conditioned on |
142/142 [=====] - 1105s 8s/step - loss: 0.2887 - accuracy: 0.9220 - val_loss: 8.5466 - val_accuracy: 0.2587 - lr: 0.0010
Epoch 2/50
142/142 [=====] - ETA: 0s - loss: 0.1155 - accuracy: 0.9643WARNING:tensorflow:Learning rate reduction is conditioned on |
142/142 [=====] - 150s 1s/step - loss: 0.1155 - accuracy: 0.9643 - val_loss: 6.2853 - val_accuracy: 0.2827 - lr: 0.0010
Epoch 3/50
142/142 [=====] - ETA: 0s - loss: 0.0665 - accuracy: 0.9817WARNING:tensorflow:Learning rate reduction is conditioned on |
142/142 [=====] - 143s 1s/step - loss: 0.0665 - accuracy: 0.9817 - val_loss: 1.0319 - val_accuracy: 0.7600 - lr: 0.0010
Epoch 4/50
142/142 [=====] - ETA: 0s - loss: 0.0529 - accuracy: 0.9859WARNING:tensorflow:Learning rate reduction is conditioned on |
142/142 [=====] - 144s 1s/step - loss: 0.0529 - accuracy: 0.9859 - val_loss: 0.0588 - val_accuracy: 0.9813 - lr: 0.0010
Epoch 5/50
142/142 [=====] - ETA: 0s - loss: 0.0455 - accuracy: 0.9878WARNING:tensorflow:Learning rate reduction is conditioned on |
142/142 [=====] - 157s 1s/step - loss: 0.0455 - accuracy: 0.9878 - val_loss: 0.0114 - val_accuracy: 0.9947 - lr: 0.0010
Epoch 6/50
142/142 [=====] - ETA: 0s - loss: 0.0428 - accuracy: 0.9887WARNING:tensorflow:Learning rate reduction is conditioned on |
142/142 [=====] - 148s 1s/step - loss: 0.0428 - accuracy: 0.9887 - val_loss: 0.0088 - val_accuracy: 0.9947 - lr: 0.0010
Epoch 7/50
142/142 [=====] - ETA: 0s - loss: 0.0258 - accuracy: 0.9930WARNING:tensorflow:Learning rate reduction is conditioned on |
142/142 [=====] - 145s 1s/step - loss: 0.0258 - accuracy: 0.9930 - val_loss: 0.0100 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 8/50
142/142 [=====] - ETA: 0s - loss: 0.0547 - accuracy: 0.9859WARNING:tensorflow:Learning rate reduction is conditioned on |
142/142 [=====] - 144s 1s/step - loss: 0.0547 - accuracy: 0.9859 - val_loss: 0.0170 - val_accuracy: 0.9947 - lr: 0.0010
Epoch 9/50
142/142 [=====] - ETA: 0s - loss: 0.0321 - accuracy: 0.9911WARNING:tensorflow:Learning rate reduction is conditioned on |
142/142 [=====] - 152s 1s/step - loss: 0.0321 - accuracy: 0.9911 - val_loss: 0.2217 - val_accuracy: 0.9307 - lr: 0.0010
Epoch 10/50
142/142 [=====] - ETA: 0s - loss: 0.0133 - accuracy: 0.9953WARNING:tensorflow:Learning rate reduction is conditioned on |
142/142 [=====] - 144s 1s/step - loss: 0.0133 - accuracy: 0.9953 - val_loss: 0.0386 - val_accuracy: 0.9920 - lr: 0.0010
Epoch 11/50
142/142 [=====] - ETA: 0s - loss: 0.0408 - accuracy: 0.9897WARNING:tensorflow:Learning rate reduction is conditioned on |
142/142 [=====] - 143s 1s/step - loss: 0.0408 - accuracy: 0.9897 - val_loss: 0.0061 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 12/50
142/142 [=====] - ETA: 0s - loss: 0.0350 - accuracy: 0.9897WARNING:tensorflow:Learning rate reduction is conditioned on |
142/142 [=====] - 144s 1s/step - loss: 0.0350 - accuracy: 0.9897 - val_loss: 0.0175 - val_accuracy: 0.9920 - lr: 0.0010
Epoch 13/50
142/142 [=====] - ETA: 0s - loss: 0.0279 - accuracy: 0.9939WARNING:tensorflow:Learning rate reduction is conditioned on |
142/142 [=====] - 148s 1s/step - loss: 0.0279 - accuracy: 0.9939 - val_loss: 0.0184 - val_accuracy: 0.9947 - lr: 0.0010
Epoch 14/50
142/142 [=====] - ETA: 0s - loss: 0.0167 - accuracy: 0.9953WARNING:tensorflow:Learning rate reduction is conditioned on |
142/142 [=====] - 145s 1s/step - loss: 0.0167 - accuracy: 0.9953 - val_loss: 0.5549 - val_accuracy: 0.8800 - lr: 0.0010
```

Figure 2 - Model 1 Results (Epochs 1-14)

```
Epoch 15/50
142/142 [=====] - ETA: 0s - loss: 0.0370 - accuracy: 0.9915WARNING:tensorflow:Learning rate reduction is conditioned on me|
142/142 [=====] - 149s 1s/step - loss: 0.0370 - accuracy: 0.9915 - val_loss: 0.0322 - val_accuracy: 0.9813 - lr: 0.0010
Epoch 16/50
142/142 [=====] - ETA: 0s - loss: 0.0315 - accuracy: 0.9930WARNING:tensorflow:Learning rate reduction is conditioned on me|
142/142 [=====] - 145s 1s/step - loss: 0.0315 - accuracy: 0.9930 - val_loss: 0.0151 - val_accuracy: 0.9920 - lr: 0.0010
Epoch 17/50
142/142 [=====] - ETA: 0s - loss: 0.0167 - accuracy: 0.9953WARNING:tensorflow:Learning rate reduction is conditioned on me|
142/142 [=====] - 143s 1s/step - loss: 0.0167 - accuracy: 0.9953 - val_loss: 0.0679 - val_accuracy: 0.9893 - lr: 0.0010
Epoch 18/50
142/142 [=====] - ETA: 0s - loss: 0.0198 - accuracy: 0.9934WARNING:tensorflow:Learning rate reduction is conditioned on me|
142/142 [=====] - 143s 1s/step - loss: 0.0198 - accuracy: 0.9934 - val_loss: 0.0024 - val_accuracy: 1.0000 - lr: 0.0010
Epoch 19/50
142/142 [=====] - ETA: 0s - loss: 0.0259 - accuracy: 0.9939WARNING:tensorflow:Learning rate reduction is conditioned on me|
142/142 [=====] - 145s 1s/step - loss: 0.0259 - accuracy: 0.9939 - val_loss: 0.0118 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 20/50
142/142 [=====] - ETA: 0s - loss: 0.0254 - accuracy: 0.9948WARNING:tensorflow:Learning rate reduction is conditioned on me|
142/142 [=====] - 144s 1s/step - loss: 0.0254 - accuracy: 0.9948 - val_loss: 0.0606 - val_accuracy: 0.9893 - lr: 0.0010
Epoch 21/50
142/142 [=====] - ETA: 0s - loss: 0.0393 - accuracy: 0.9901WARNING:tensorflow:Learning rate reduction is conditioned on me|
142/142 [=====] - 147s 1s/step - loss: 0.0393 - accuracy: 0.9901 - val_loss: 0.0063 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 22/50
142/142 [=====] - ETA: 0s - loss: 0.0274 - accuracy: 0.9915WARNING:tensorflow:Learning rate reduction is conditioned on me|
142/142 [=====] - 144s 1s/step - loss: 0.0274 - accuracy: 0.9915 - val_loss: 0.0544 - val_accuracy: 0.9867 - lr: 0.0010
Epoch 23/50
142/142 [=====] - ETA: 0s - loss: 0.0089 - accuracy: 0.9967WARNING:tensorflow:Learning rate reduction is conditioned on me|
142/142 [=====] - 142s 1s/step - loss: 0.0089 - accuracy: 0.9967 - val_loss: 0.0190 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 24/50
142/142 [=====] - ETA: 0s - loss: 0.0413 - accuracy: 0.9939WARNING:tensorflow:Learning rate reduction is conditioned on me|
142/142 [=====] - 144s 1s/step - loss: 0.0413 - accuracy: 0.9939 - val_loss: 0.3698 - val_accuracy: 0.8960 - lr: 0.0010
Epoch 25/50
142/142 [=====] - ETA: 0s - loss: 0.0199 - accuracy: 0.9958WARNING:tensorflow:Learning rate reduction is conditioned on me|
142/142 [=====] - 141s 990ms/step - loss: 0.0199 - accuracy: 0.9958 - val_loss: 0.0906 - val_accuracy: 0.9653 - lr: 0.0010
Epoch 26/50
142/142 [=====] - ETA: 0s - loss: 0.0172 - accuracy: 0.9953WARNING:tensorflow:Learning rate reduction is conditioned on me|
142/142 [=====] - 144s 1s/step - loss: 0.0172 - accuracy: 0.9953 - val_loss: 0.0239 - val_accuracy: 0.9920 - lr: 0.0010
Epoch 27/50
142/142 [=====] - ETA: 0s - loss: 0.0298 - accuracy: 0.9925WARNING:tensorflow:Learning rate reduction is conditioned on me|
142/142 [=====] - 141s 993ms/step - loss: 0.0298 - accuracy: 0.9925 - val_loss: 0.0408 - val_accuracy: 0.9947 - lr: 0.0010
Epoch 28/50
142/142 [=====] - ETA: 0s - loss: 0.0180 - accuracy: 0.9958WARNING:tensorflow:Learning rate reduction is conditioned on me|
142/142 [=====] - 143s 1s/step - loss: 0.0180 - accuracy: 0.9958 - val_loss: 0.0488 - val_accuracy: 0.9920 - lr: 0.0010
```

Figure 3 - Model 1 Results (Epochs 15-28)

The experiment is continued by removing convolution layers, thinking a smaller number of layers would mean fewer parameters and a faster neural network. However, the number of parameters doubled due to the increased work of the dense layer, which had to compress more data (Figure 4). The results will be compared to the original, and if there are changes, the experiment will continue in that way.

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 126, 126, 32)	896
batch_normalization (Batch Normalization)	(None, 126, 126, 32)	128
max_pooling2d (MaxPooling2D)	(None, 63, 63, 32)	0
dropout (Dropout)	(None, 63, 63, 32)	0
conv2d_1 (Conv2D)	(None, 61, 61, 64)	18496
batch_normalization_1 (Batch Normalization)	(None, 61, 61, 64)	256
max_pooling2d_1 (MaxPooling2D)	(None, 30, 30, 64)	0
dropout_1 (Dropout)	(None, 30, 30, 64)	0
flatten (Flatten)	(None, 57600)	0
dense (Dense)	(None, 512)	29491712
batch_normalization_2 (Batch Normalization)	(None, 512)	2048
dropout_2 (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 4)	2052

=====
 Total params: 29,515,588
 Trainable params: 29,514,372
 Non-trainable params: 1,216
 =====

Figure 4 - Model 2 (optimizer = rmsprop, batch size = 15)


```

Epoch 1/50
142/142 [=====] - ETA: 0s - loss: 0.2784 - accuracy: 0.9239WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 837s 6s/step - loss: 0.2784 - accuracy: 0.9239 - val_loss: 7.5530 - val_accuracy: 0.2560 - lr: 0.0010
Epoch 2/50
142/142 [=====] - ETA: 0s - loss: 0.0984 - accuracy: 0.9695WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 151s 1s/step - loss: 0.0984 - accuracy: 0.9695 - val_loss: 6.8778 - val_accuracy: 0.2747 - lr: 0.0010
Epoch 3/50
142/142 [=====] - ETA: 0s - loss: 0.0676 - accuracy: 0.9803WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 161s 1s/step - loss: 0.0676 - accuracy: 0.9803 - val_loss: 1.2983 - val_accuracy: 0.7573 - lr: 0.0010
Epoch 4/50
142/142 [=====] - ETA: 0s - loss: 0.0859 - accuracy: 0.9807WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 154s 1s/step - loss: 0.0859 - accuracy: 0.9807 - val_loss: 0.2967 - val_accuracy: 0.9227 - lr: 0.0010
Epoch 5/50
142/142 [=====] - ETA: 0s - loss: 0.0440 - accuracy: 0.9845WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 151s 1s/step - loss: 0.0440 - accuracy: 0.9845 - val_loss: 0.0157 - val_accuracy: 0.9920 - lr: 0.0010
Epoch 6/50
142/142 [=====] - ETA: 0s - loss: 0.0412 - accuracy: 0.9887WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 152s 1s/step - loss: 0.0412 - accuracy: 0.9887 - val_loss: 0.0175 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 7/50
142/142 [=====] - ETA: 0s - loss: 0.0347 - accuracy: 0.9901WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 151s 1s/step - loss: 0.0347 - accuracy: 0.9901 - val_loss: 0.0548 - val_accuracy: 0.9893 - lr: 0.0010
Epoch 8/50
142/142 [=====] - ETA: 0s - loss: 0.0442 - accuracy: 0.9892WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 150s 1s/step - loss: 0.0442 - accuracy: 0.9892 - val_loss: 0.0072 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 9/50
142/142 [=====] - ETA: 0s - loss: 0.0402 - accuracy: 0.9911WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 151s 1s/step - loss: 0.0402 - accuracy: 0.9911 - val_loss: 0.1486 - val_accuracy: 0.9627 - lr: 0.0010

```

Figure 5 - Model 2 Results (Epochs 1-9)

```

Epoch 10/50
142/142 [=====] - ETA: 0s - loss: 0.0389 - accuracy: 0.9892WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 150s 1s/step - loss: 0.0389 - accuracy: 0.9892 - val_loss: 0.0142 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 11/50
142/142 [=====] - ETA: 0s - loss: 0.0359 - accuracy: 0.9925WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 149s 1s/step - loss: 0.0359 - accuracy: 0.9925 - val_loss: 0.0126 - val_accuracy: 0.9920 - lr: 0.0010
Epoch 12/50
142/142 [=====] - ETA: 0s - loss: 0.0344 - accuracy: 0.9930WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 150s 1s/step - loss: 0.0344 - accuracy: 0.9930 - val_loss: 0.0226 - val_accuracy: 0.9947 - lr: 0.0010
Epoch 13/50
142/142 [=====] - ETA: 0s - loss: 0.0258 - accuracy: 0.9934WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 149s 1s/step - loss: 0.0258 - accuracy: 0.9934 - val_loss: 0.0110 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 14/50
142/142 [=====] - ETA: 0s - loss: 0.0234 - accuracy: 0.9925WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 149s 1s/step - loss: 0.0234 - accuracy: 0.9925 - val_loss: 0.0276 - val_accuracy: 0.9947 - lr: 0.0010
Epoch 15/50
142/142 [=====] - ETA: 0s - loss: 0.0397 - accuracy: 0.9930WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 149s 1s/step - loss: 0.0397 - accuracy: 0.9930 - val_loss: 0.0121 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 16/50
142/142 [=====] - ETA: 0s - loss: 0.0453 - accuracy: 0.9906WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 150s 1s/step - loss: 0.0453 - accuracy: 0.9906 - val_loss: 0.1113 - val_accuracy: 0.9600 - lr: 0.0010
Epoch 17/50
142/142 [=====] - ETA: 0s - loss: 0.0357 - accuracy: 0.9915WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 150s 1s/step - loss: 0.0357 - accuracy: 0.9915 - val_loss: 0.0119 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 18/50
142/142 [=====] - ETA: 0s - loss: 0.0314 - accuracy: 0.9901WARNING:tensorflow:Learning rate reduction is conditioned on
142/142 [=====] - 149s 1s/step - loss: 0.0314 - accuracy: 0.9901 - val_loss: 0.0528 - val_accuracy: 0.9867 - lr: 0.0010

```

Figure 6 - Model 2 Results (Epochs 10-18)

As a convolution layer had already been added, the experiment continued by removing convolution layers, and the number of parameters sharply decreased (Figure 7).

Model: "sequential_2"

Layer (type)	Output Shape	Param #
conv2d_6 (Conv2D)	(None, 126, 126, 32)	896
batch_normalization_8 (Batch Normalization)	(None, 126, 126, 32)	128
max_pooling2d_6 (MaxPooling2D)	(None, 63, 63, 32)	0
dropout_8 (Dropout)	(None, 63, 63, 32)	0
conv2d_7 (Conv2D)	(None, 61, 61, 64)	18496
batch_normalization_9 (Batch Normalization)	(None, 61, 61, 64)	256
max_pooling2d_7 (MaxPooling2D)	(None, 30, 30, 64)	0
dropout_9 (Dropout)	(None, 30, 30, 64)	0
conv2d_8 (Conv2D)	(None, 28, 28, 64)	36928
batch_normalization_10 (Batch Normalization)	(None, 28, 28, 64)	256
max_pooling2d_8 (MaxPooling2D)	(None, 14, 14, 64)	0
dropout_10 (Dropout)	(None, 14, 14, 64)	0
conv2d_9 (Conv2D)	(None, 12, 12, 64)	36928
batch_normalization_11 (Batch Normalization)	(None, 12, 12, 64)	256
max_pooling2d_9 (MaxPooling2D)	(None, 6, 6, 64)	0
dropout_11 (Dropout)	(None, 6, 6, 64)	0
flatten_2 (Flatten)	(None, 2304)	0
dense_4 (Dense)	(None, 512)	1180160
batch_normalization_12 (Batch Normalization)	(None, 512)	2048
dropout_12 (Dropout)	(None, 512)	0
dense_5 (Dense)	(None, 4)	2052

=====
Total params: 1,278,404
Trainable params: 1,276,932

Figure 7 - Model 3 (optimizer = rmsprop, batch size = 15)

```

Epoch 1/50
<ipython-input-37-63cb34b6e89a>:2: UserWarning: `Model.fit_generator` is deprecated and will be removed in a future version. Please use `Model.fit`,
  history = model.fit_generator(
142/142 [=====] - ETA: 0s - loss: 0.0640 - accuracy: 0.9859WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 106s 746ms/step - loss: 0.0640 - accuracy: 0.9859 - val_loss: 0.0902 - val_accuracy: 0.9680 - lr: 0.0010
Epoch 2/50
142/142 [=====] - ETA: 0s - loss: 0.0497 - accuracy: 0.9859WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 103s 729ms/step - loss: 0.0497 - accuracy: 0.9859 - val_loss: 0.1890 - val_accuracy: 0.9280 - lr: 0.0010
Epoch 3/50
142/142 [=====] - ETA: 0s - loss: 0.0663 - accuracy: 0.9840WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 106s 743ms/step - loss: 0.0663 - accuracy: 0.9840 - val_loss: 0.0195 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 4/50
142/142 [=====] - ETA: 0s - loss: 0.0415 - accuracy: 0.9897WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 108s 759ms/step - loss: 0.0415 - accuracy: 0.9897 - val_loss: 0.0130 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 5/50
142/142 [=====] - ETA: 0s - loss: 0.0537 - accuracy: 0.9878WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 111s 778ms/step - loss: 0.0537 - accuracy: 0.9878 - val_loss: 0.1623 - val_accuracy: 0.9707 - lr: 0.0010
Epoch 6/50
142/142 [=====] - ETA: 0s - loss: 0.0505 - accuracy: 0.9883WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 105s 741ms/step - loss: 0.0505 - accuracy: 0.9883 - val_loss: 0.0075 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 7/50
142/142 [=====] - ETA: 0s - loss: 0.0499 - accuracy: 0.9887WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 103s 723ms/step - loss: 0.0499 - accuracy: 0.9887 - val_loss: 0.0238 - val_accuracy: 0.9947 - lr: 0.0010
Epoch 8/50
142/142 [=====] - ETA: 0s - loss: 0.0247 - accuracy: 0.9925WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 107s 751ms/step - loss: 0.0247 - accuracy: 0.9925 - val_loss: 0.0256 - val_accuracy: 0.9920 - lr: 0.0010
Epoch 9/50
142/142 [=====] - ETA: 0s - loss: 0.0437 - accuracy: 0.9915WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 107s 747ms/step - loss: 0.0437 - accuracy: 0.9915 - val_loss: 0.0433 - val_accuracy: 0.9813 - lr: 0.0010
Epoch 10/50
142/142 [=====] - ETA: 0s - loss: 0.0440 - accuracy: 0.9897WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 102s 714ms/step - loss: 0.0440 - accuracy: 0.9897 - val_loss: 0.0157 - val_accuracy: 0.9947 - lr: 0.0010
Epoch 11/50
142/142 [=====] - ETA: 0s - loss: 0.0203 - accuracy: 0.9962WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 106s 744ms/step - loss: 0.0203 - accuracy: 0.9962 - val_loss: 0.0030 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 12/50
142/142 [=====] - ETA: 0s - loss: 0.0200 - accuracy: 0.9939WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 106s 744ms/step - loss: 0.0200 - accuracy: 0.9939 - val_loss: 0.0168 - val_accuracy: 0.9947 - lr: 0.0010

```

Figure 8 - Model 3 Results (Epochs 1-12)

```

Epoch 13/50
142/142 [=====] - ETA: 0s - loss: 0.0431 - accuracy: 0.9920WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 103s 727ms/step - loss: 0.0431 - accuracy: 0.9920 - val_loss: 0.0224 - val_accuracy: 0.9893 - lr: 0.0010
Epoch 14/50
142/142 [=====] - ETA: 0s - loss: 0.0238 - accuracy: 0.9953WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 101s 708ms/step - loss: 0.0238 - accuracy: 0.9953 - val_loss: 0.0023 - val_accuracy: 1.0000 - lr: 0.0010
Epoch 15/50
142/142 [=====] - ETA: 0s - loss: 0.0414 - accuracy: 0.9944WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 102s 718ms/step - loss: 0.0414 - accuracy: 0.9944 - val_loss: 0.0138 - val_accuracy: 0.9947 - lr: 0.0010
Epoch 16/50
142/142 [=====] - ETA: 0s - loss: 0.0213 - accuracy: 0.9939WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 103s 726ms/step - loss: 0.0213 - accuracy: 0.9939 - val_loss: 0.0478 - val_accuracy: 0.9893 - lr: 0.0010
Epoch 17/50
142/142 [=====] - ETA: 0s - loss: 0.0072 - accuracy: 0.9977WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 107s 751ms/step - loss: 0.0072 - accuracy: 0.9977 - val_loss: 0.0138 - val_accuracy: 0.9947 - lr: 0.0010
Epoch 18/50
142/142 [=====] - ETA: 0s - loss: 0.0107 - accuracy: 0.9981WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 104s 732ms/step - loss: 0.0107 - accuracy: 0.9981 - val_loss: 0.0379 - val_accuracy: 0.9893 - lr: 0.0010
Epoch 19/50
142/142 [=====] - ETA: 0s - loss: 0.0270 - accuracy: 0.9958WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 104s 728ms/step - loss: 0.0270 - accuracy: 0.9958 - val_loss: 0.2600 - val_accuracy: 0.9760 - lr: 0.0010
Epoch 20/50
142/142 [=====] - ETA: 0s - loss: 0.0300 - accuracy: 0.9948WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 108s 759ms/step - loss: 0.0300 - accuracy: 0.9948 - val_loss: 0.2202 - val_accuracy: 0.9760 - lr: 0.0010
Epoch 21/50
142/142 [=====] - ETA: 0s - loss: 0.0387 - accuracy: 0.9925WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 104s 732ms/step - loss: 0.0387 - accuracy: 0.9925 - val_loss: 0.0120 - val_accuracy: 0.9920 - lr: 0.0010
Epoch 22/50
142/142 [=====] - ETA: 0s - loss: 0.0396 - accuracy: 0.9939WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 103s 727ms/step - loss: 0.0396 - accuracy: 0.9939 - val_loss: 0.0173 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 23/50
142/142 [=====] - ETA: 0s - loss: 0.0369 - accuracy: 0.9930WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 106s 746ms/step - loss: 0.0369 - accuracy: 0.9930 - val_loss: 0.0052 - val_accuracy: 0.9973 - lr: 0.0010
Epoch 24/50
142/142 [=====] - ETA: 0s - loss: 0.0377 - accuracy: 0.9915WARNING:tensorflow:Learning rate reduction is conditioned on met
142/142 [=====] - 102s 719ms/step - loss: 0.0377 - accuracy: 0.9915 - val_loss: 0.2210 - val_accuracy: 0.9680 - lr: 0.0010

```

Figure 9 - Model 3 Results (Epochs 13-24)

At this point, predicting the traffic light color would be a relatively simple task for a neural network. So the dataset was replaced by a lung disease classification dataset. For convolution models, distinguishing different medical conditions is more complicated than other datasets, like traffic light classification.

The max pooling size was changed to (3,3) instead of (2,2) to lower vector sizes, and the program's dropout size was changed to 0.25 for consistency (Figure 10).

Model: "sequential_2"

Layer (type)	Output Shape	Param #
conv2d_6 (Conv2D)	(None, 126, 126, 32)	896
batch_normalization_8 (Batch Normalization)	(None, 126, 126, 32)	128
max_pooling2d_6 (MaxPooling2D)	(None, 63, 63, 32)	0
dropout_8 (Dropout)	(None, 63, 63, 32)	0
conv2d_7 (Conv2D)	(None, 61, 61, 64)	18496
batch_normalization_9 (Batch Normalization)	(None, 61, 61, 64)	256
max_pooling2d_7 (MaxPooling2D)	(None, 30, 30, 64)	0
dropout_9 (Dropout)	(None, 30, 30, 64)	0
conv2d_8 (Conv2D)	(None, 28, 28, 64)	36928
batch_normalization_10 (Batch Normalization)	(None, 28, 28, 64)	256
max_pooling2d_8 (MaxPooling2D)	(None, 14, 14, 64)	0
dropout_10 (Dropout)	(None, 14, 14, 64)	0
conv2d_9 (Conv2D)	(None, 12, 12, 64)	36928
batch_normalization_11 (Batch Normalization)	(None, 12, 12, 64)	256
max_pooling2d_9 (MaxPooling2D)	(None, 6, 6, 64)	0
dropout_11 (Dropout)	(None, 6, 6, 64)	0
flatten_2 (Flatten)	(None, 2304)	0
dense_4 (Dense)	(None, 512)	1180160
batch_normalization_12 (Batch Normalization)	(None, 512)	2048
dropout_12 (Dropout)	(None, 512)	0
dense_5 (Dense)	(None, 4)	2052
Total params: 1,278,404		
Trainable params: 1,276,932		

Figure 10 - Model 4 (optimizer = rmsprop, batch size = 15)

```
Epoch 1/50
566/566 [=====] - ETA: 0s - loss: 1.2412 - accuracy: 0.4717WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 1424s 3s/step - loss: 1.2412 - accuracy: 0.4717 - val_loss: 1.0967 - val_accuracy: 0.5447 - lr: 0.0010
Epoch 2/50
566/566 [=====] - ETA: 0s - loss: 1.0409 - accuracy: 0.5117WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 523s 923ms/step - loss: 1.0409 - accuracy: 0.5117 - val_loss: 1.2842 - val_accuracy: 0.5461 - lr: 0.0010
Epoch 3/50
566/566 [=====] - ETA: 0s - loss: 1.0060 - accuracy: 0.5241WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 529s 935ms/step - loss: 1.0060 - accuracy: 0.5241 - val_loss: 1.0190 - val_accuracy: 0.5433 - lr: 0.0010
Epoch 4/50
566/566 [=====] - ETA: 0s - loss: 0.9936 - accuracy: 0.5426WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 518s 914ms/step - loss: 0.9936 - accuracy: 0.5426 - val_loss: 1.1445 - val_accuracy: 0.5385 - lr: 0.0010
Epoch 5/50
566/566 [=====] - ETA: 0s - loss: 0.9894 - accuracy: 0.5389WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 516s 911ms/step - loss: 0.9894 - accuracy: 0.5389 - val_loss: 1.0417 - val_accuracy: 0.5442 - lr: 0.0010
Epoch 6/50
566/566 [=====] - ETA: 0s - loss: 0.9790 - accuracy: 0.5393WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 508s 898ms/step - loss: 0.9790 - accuracy: 0.5393 - val_loss: 1.0325 - val_accuracy: 0.5414 - lr: 0.0010
Epoch 7/50
566/566 [=====] - ETA: 0s - loss: 0.9693 - accuracy: 0.5449WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 532s 940ms/step - loss: 0.9693 - accuracy: 0.5449 - val_loss: 1.0644 - val_accuracy: 0.5475 - lr: 0.0010
Epoch 8/50
566/566 [=====] - ETA: 0s - loss: 0.9664 - accuracy: 0.5478WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 506s 894ms/step - loss: 0.9664 - accuracy: 0.5478 - val_loss: 1.1774 - val_accuracy: 0.5343 - lr: 0.0010
Epoch 9/50
566/566 [=====] - ETA: 0s - loss: 0.9692 - accuracy: 0.5469WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 530s 937ms/step - loss: 0.9692 - accuracy: 0.5469 - val_loss: 1.1131 - val_accuracy: 0.5456 - lr: 0.0010
Epoch 10/50
566/566 [=====] - ETA: 0s - loss: 0.9628 - accuracy: 0.5476WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 518s 916ms/step - loss: 0.9628 - accuracy: 0.5476 - val_loss: 1.0729 - val_accuracy: 0.5433 - lr: 0.0010
```

Figure 11 - Model 4 Results (Epochs 1-10)

```
Epoch 11/50
566/566 [=====] - ETA: 0s - loss: 0.9603 - accuracy: 0.5462WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 508s 897ms/step - loss: 0.9603 - accuracy: 0.5462 - val_loss: 1.0524 - val_accuracy: 0.5447 - lr: 0.0010
Epoch 12/50
566/566 [=====] - ETA: 0s - loss: 0.9622 - accuracy: 0.5485WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 517s 913ms/step - loss: 0.9622 - accuracy: 0.5485 - val_loss: 1.0008 - val_accuracy: 0.5551 - lr: 0.0010
Epoch 13/50
566/566 [=====] - ETA: 0s - loss: 0.9570 - accuracy: 0.5537WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 514s 909ms/step - loss: 0.9570 - accuracy: 0.5537 - val_loss: 1.5134 - val_accuracy: 0.5584 - lr: 0.0010
Epoch 14/50
566/566 [=====] - ETA: 0s - loss: 0.9586 - accuracy: 0.5613WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 511s 902ms/step - loss: 0.9586 - accuracy: 0.5613 - val_loss: 1.0241 - val_accuracy: 0.5092 - lr: 0.0010
Epoch 15/50
566/566 [=====] - ETA: 0s - loss: 0.9518 - accuracy: 0.5593WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 518s 916ms/step - loss: 0.9518 - accuracy: 0.5593 - val_loss: 1.3043 - val_accuracy: 0.5404 - lr: 0.0010
Epoch 16/50
566/566 [=====] - ETA: 0s - loss: 0.9542 - accuracy: 0.5555WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 515s 910ms/step - loss: 0.9542 - accuracy: 0.5555 - val_loss: 0.9925 - val_accuracy: 0.5343 - lr: 0.0010
Epoch 17/50
566/566 [=====] - ETA: 0s - loss: 0.9512 - accuracy: 0.5537WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 515s 910ms/step - loss: 0.9512 - accuracy: 0.5537 - val_loss: 0.9151 - val_accuracy: 0.5773 - lr: 0.0010
Epoch 18/50
566/566 [=====] - ETA: 0s - loss: 0.9437 - accuracy: 0.5626WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 495s 875ms/step - loss: 0.9437 - accuracy: 0.5626 - val_loss: 1.0758 - val_accuracy: 0.5480 - lr: 0.0010
Epoch 19/50
566/566 [=====] - ETA: 0s - loss: 0.9454 - accuracy: 0.5586WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 492s 870ms/step - loss: 0.9454 - accuracy: 0.5586 - val_loss: 0.9763 - val_accuracy: 0.5783 - lr: 0.0010
Epoch 20/50
566/566 [=====] - ETA: 0s - loss: 0.9414 - accuracy: 0.5605WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 495s 875ms/step - loss: 0.9414 - accuracy: 0.5605 - val_loss: 0.9416 - val_accuracy: 0.5650 - lr: 0.0010
```

Figure 12 - Model 4 Results (Epochs 11-20)

The program progressed with a change of optimizer. The other models used an optimizer called RMSProp, so the program's optimizer was changed to a different one called Adam (Figure 13).

Model: "sequential_2"

Layer (type)	Output Shape	Param #
conv2d_6 (Conv2D)	(None, 126, 126, 32)	896
batch_normalization_8 (Batch Normalization)	(None, 126, 126, 32)	128
max_pooling2d_6 (MaxPooling2D)	(None, 63, 63, 32)	0
dropout_8 (Dropout)	(None, 63, 63, 32)	0
conv2d_7 (Conv2D)	(None, 61, 61, 64)	18496
batch_normalization_9 (Batch Normalization)	(None, 61, 61, 64)	256
max_pooling2d_7 (MaxPooling2D)	(None, 30, 30, 64)	0
dropout_9 (Dropout)	(None, 30, 30, 64)	0
conv2d_8 (Conv2D)	(None, 28, 28, 64)	36928
batch_normalization_10 (Batch Normalization)	(None, 28, 28, 64)	256
max_pooling2d_8 (MaxPooling2D)	(None, 14, 14, 64)	0
dropout_10 (Dropout)	(None, 14, 14, 64)	0
conv2d_9 (Conv2D)	(None, 12, 12, 64)	36928
batch_normalization_11 (Batch Normalization)	(None, 12, 12, 64)	256
max_pooling2d_9 (MaxPooling2D)	(None, 6, 6, 64)	0
dropout_11 (Dropout)	(None, 6, 6, 64)	0
flatten_2 (Flatten)	(None, 2304)	0
dense_4 (Dense)	(None, 512)	1180160
batch_normalization_12 (Batch Normalization)	(None, 512)	2048
dropout_12 (Dropout)	(None, 512)	0
dense_5 (Dense)	(None, 4)	2052

=====
Total params: 1,278,404
Trainable params: 1,276,932

Figure 13 - Model 5 (optimizer = Adam, batch size = 15)

```
Epoch 1/50
566/566 [=====] - ETA: 0s - loss: 1.2051 - accuracy: 0.4720WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 1057s 2s/step - loss: 1.2051 - accuracy: 0.4720 - val_loss: 1.0301 - val_accuracy: 0.5442 - lr: 0.0010
Epoch 2/50
566/566 [=====] - ETA: 0s - loss: 1.0384 - accuracy: 0.5141WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 493s 871ms/step - loss: 1.0384 - accuracy: 0.5141 - val_loss: 1.1178 - val_accuracy: 0.4629 - lr: 0.0010
Epoch 3/50
566/566 [=====] - ETA: 0s - loss: 1.0145 - accuracy: 0.5244WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 502s 886ms/step - loss: 1.0145 - accuracy: 0.5244 - val_loss: 0.9968 - val_accuracy: 0.5504 - lr: 0.0010
Epoch 4/50
566/566 [=====] - ETA: 0s - loss: 1.0030 - accuracy: 0.5268WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 472s 833ms/step - loss: 1.0030 - accuracy: 0.5268 - val_loss: 0.9979 - val_accuracy: 0.5173 - lr: 0.0010
Epoch 5/50
566/566 [=====] - ETA: 0s - loss: 0.9955 - accuracy: 0.5284WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 497s 878ms/step - loss: 0.9955 - accuracy: 0.5284 - val_loss: 1.0153 - val_accuracy: 0.5475 - lr: 0.0010
Epoch 6/50
566/566 [=====] - ETA: 0s - loss: 0.9825 - accuracy: 0.5427WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 488s 862ms/step - loss: 0.9825 - accuracy: 0.5427 - val_loss: 0.9426 - val_accuracy: 0.5456 - lr: 0.0010
Epoch 7/50
566/566 [=====] - ETA: 0s - loss: 0.9785 - accuracy: 0.5402WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 475s 840ms/step - loss: 0.9785 - accuracy: 0.5402 - val_loss: 0.9732 - val_accuracy: 0.5229 - lr: 0.0010
Epoch 8/50
566/566 [=====] - ETA: 0s - loss: 0.9771 - accuracy: 0.5382WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 479s 847ms/step - loss: 0.9771 - accuracy: 0.5382 - val_loss: 0.9770 - val_accuracy: 0.5418 - lr: 0.0010
Epoch 9/50
566/566 [=====] - ETA: 0s - loss: 0.9724 - accuracy: 0.5434WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 456s 806ms/step - loss: 0.9724 - accuracy: 0.5434 - val_loss: 0.9155 - val_accuracy: 0.5844 - lr: 0.0010
Epoch 10/50
566/566 [=====] - ETA: 0s - loss: 0.9676 - accuracy: 0.5500WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 480s 848ms/step - loss: 0.9676 - accuracy: 0.5500 - val_loss: 0.9653 - val_accuracy: 0.5636 - lr: 0.0010
```

Figure 14 - Model 5 Results (Epochs 1-10)

```
Epoch 11/50
566/566 [=====] - ETA: 0s - loss: 0.9700 - accuracy: 0.5481WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 474s 838ms/step - loss: 0.9700 - accuracy: 0.5481 - val_loss: 1.0471 - val_accuracy: 0.5385 - lr: 0.0010
Epoch 12/50
566/566 [=====] - ETA: 0s - loss: 0.9653 - accuracy: 0.5456WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 481s 850ms/step - loss: 0.9653 - accuracy: 0.5456 - val_loss: 0.9568 - val_accuracy: 0.5650 - lr: 0.0010
Epoch 13/50
566/566 [=====] - ETA: 0s - loss: 0.9537 - accuracy: 0.5508WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 477s 843ms/step - loss: 0.9537 - accuracy: 0.5508 - val_loss: 0.9236 - val_accuracy: 0.5759 - lr: 0.0010
Epoch 14/50
566/566 [=====] - ETA: 0s - loss: 0.9603 - accuracy: 0.5489WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 475s 840ms/step - loss: 0.9603 - accuracy: 0.5489 - val_loss: 0.9508 - val_accuracy: 0.5504 - lr: 0.0010
Epoch 15/50
566/566 [=====] - ETA: 0s - loss: 0.9532 - accuracy: 0.5530WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 475s 839ms/step - loss: 0.9532 - accuracy: 0.5530 - val_loss: 0.9372 - val_accuracy: 0.5626 - lr: 0.0010
Epoch 16/50
566/566 [=====] - ETA: 0s - loss: 0.9484 - accuracy: 0.5551WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 472s 834ms/step - loss: 0.9484 - accuracy: 0.5551 - val_loss: 1.0088 - val_accuracy: 0.5428 - lr: 0.0010
Epoch 17/50
566/566 [=====] - ETA: 0s - loss: 0.9453 - accuracy: 0.5595WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 447s 789ms/step - loss: 0.9453 - accuracy: 0.5595 - val_loss: 1.0152 - val_accuracy: 0.5645 - lr: 0.0010
Epoch 18/50
566/566 [=====] - ETA: 0s - loss: 0.9521 - accuracy: 0.5627WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 439s 774ms/step - loss: 0.9521 - accuracy: 0.5627 - val_loss: 0.9416 - val_accuracy: 0.5707 - lr: 0.0010
Epoch 19/50
566/566 [=====] - ETA: 0s - loss: 0.9436 - accuracy: 0.5591WARNING:tensorflow:Learning rate reduction is conditioned on met
566/566 [=====] - 442s 780ms/step - loss: 0.9436 - accuracy: 0.5591 - val_loss: 0.9352 - val_accuracy: 0.5626 - lr: 0.0010
```

Figure 15 - Model 5 Results (Epochs 11-19)

The experiment progressed by raising the batch size from 15 to 50 so more images could be propagated through the neural network. (Figure 16).

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 126, 126, 32)	896
batch_normalization (Batch Normalization)	(None, 126, 126, 32)	128
max_pooling2d (MaxPooling2D)	(None, 42, 42, 32)	0
dropout (Dropout)	(None, 42, 42, 32)	0
conv2d_1 (Conv2D)	(None, 40, 40, 64)	18496
batch_normalization_1 (Batch Normalization)	(None, 40, 40, 64)	256
max_pooling2d_1 (MaxPooling2D)	(None, 13, 13, 64)	0
dropout_1 (Dropout)	(None, 13, 13, 64)	0
conv2d_2 (Conv2D)	(None, 11, 11, 64)	36928
batch_normalization_2 (Batch Normalization)	(None, 11, 11, 64)	256
max_pooling2d_2 (MaxPooling2D)	(None, 3, 3, 64)	0
dropout_2 (Dropout)	(None, 3, 3, 64)	0
flatten (Flatten)	(None, 576)	0
dense (Dense)	(None, 512)	295424
batch_normalization_3 (Batch Normalization)	(None, 512)	2048
dropout_3 (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 3)	1539

=====
Total params: 355,971
Trainable params: 354,627
Non-trainable params: 1,344
=====

Figure 16 - Model 6 (Optimizer = Adam, batch size = 50)


```

Epoch 1/50
170/170 [=====] - ETA: 0s - loss: 1.2821 - accuracy: 0.4567WARNING:tensorflow:Learning rate reduction is conditioned on metrics
170/170 [=====] - 1655s 10s/step - loss: 1.2821 - accuracy: 0.4567 - val_loss: 1.0935 - val_accuracy: 0.5405 - lr: 0.0010
Epoch 2/50
170/170 [=====] - ETA: 0s - loss: 1.0831 - accuracy: 0.5072WARNING:tensorflow:Learning rate reduction is conditioned on metrics
170/170 [=====] - 500s 3s/step - loss: 1.0831 - accuracy: 0.5072 - val_loss: 1.0934 - val_accuracy: 0.5419 - lr: 0.0010
Epoch 3/50
170/170 [=====] - ETA: 0s - loss: 1.0152 - accuracy: 0.5252WARNING:tensorflow:Learning rate reduction is conditioned on metrics
170/170 [=====] - 496s 3s/step - loss: 1.0152 - accuracy: 0.5252 - val_loss: 0.9893 - val_accuracy: 0.5390 - lr: 0.0010
Epoch 4/50
170/170 [=====] - ETA: 0s - loss: 0.9919 - accuracy: 0.5308WARNING:tensorflow:Learning rate reduction is conditioned on metrics
170/170 [=====] - 489s 3s/step - loss: 0.9919 - accuracy: 0.5308 - val_loss: 0.9546 - val_accuracy: 0.5557 - lr: 0.0010
Epoch 5/50
170/170 [=====] - ETA: 0s - loss: 0.9813 - accuracy: 0.5438WARNING:tensorflow:Learning rate reduction is conditioned on metrics
170/170 [=====] - 487s 3s/step - loss: 0.9813 - accuracy: 0.5438 - val_loss: 0.9580 - val_accuracy: 0.5481 - lr: 0.0010
Epoch 6/50
170/170 [=====] - ETA: 0s - loss: 0.9740 - accuracy: 0.5423WARNING:tensorflow:Learning rate reduction is conditioned on metrics
170/170 [=====] - 489s 3s/step - loss: 0.9740 - accuracy: 0.5423 - val_loss: 0.9646 - val_accuracy: 0.5481 - lr: 0.0010
Epoch 7/50
170/170 [=====] - ETA: 0s - loss: 0.9616 - accuracy: 0.5468WARNING:tensorflow:Learning rate reduction is conditioned on metrics
170/170 [=====] - 490s 3s/step - loss: 0.9616 - accuracy: 0.5468 - val_loss: 1.1281 - val_accuracy: 0.5371 - lr: 0.0010
Epoch 8/50
170/170 [=====] - ETA: 0s - loss: 0.9700 - accuracy: 0.5471WARNING:tensorflow:Learning rate reduction is conditioned on metrics
170/170 [=====] - 459s 3s/step - loss: 0.9700 - accuracy: 0.5471 - val_loss: 0.9947 - val_accuracy: 0.5381 - lr: 0.0010
Epoch 9/50
170/170 [=====] - ETA: 0s - loss: 0.9613 - accuracy: 0.5480WARNING:tensorflow:Learning rate reduction is conditioned on metrics
170/170 [=====] - 477s 3s/step - loss: 0.9613 - accuracy: 0.5480 - val_loss: 0.9367 - val_accuracy: 0.5619 - lr: 0.0010
Epoch 10/50
170/170 [=====] - ETA: 0s - loss: 0.9565 - accuracy: 0.5492WARNING:tensorflow:Learning rate reduction is conditioned on metrics
170/170 [=====] - 450s 3s/step - loss: 0.9565 - accuracy: 0.5492 - val_loss: 0.9243 - val_accuracy: 0.5710 - lr: 0.0010
Epoch 11/50
170/170 [=====] - ETA: 0s - loss: 0.9514 - accuracy: 0.5558WARNING:tensorflow:Learning rate reduction is conditioned on metrics
170/170 [=====] - 467s 3s/step - loss: 0.9514 - accuracy: 0.5558 - val_loss: 0.9371 - val_accuracy: 0.5529 - lr: 0.0010
Epoch 12/50
170/170 [=====] - ETA: 0s - loss: 0.9532 - accuracy: 0.5499WARNING:tensorflow:Learning rate reduction is conditioned on metrics
170/170 [=====] - 469s 3s/step - loss: 0.9532 - accuracy: 0.5499 - val_loss: 0.9155 - val_accuracy: 0.5686 - lr: 0.0010
Epoch 13/50
170/170 [=====] - ETA: 0s - loss: 0.9414 - accuracy: 0.5577WARNING:tensorflow:Learning rate reduction is conditioned on metrics
170/170 [=====] - 499s 3s/step - loss: 0.9414 - accuracy: 0.5577 - val_loss: 0.9354 - val_accuracy: 0.5576 - lr: 0.0010

```

Figure 17 - Model 6 Results (Epochs 1-13)

```
Epoch 14/50
170/170 [=====] - ETA: 0s - loss: 0.9420 - accuracy: 0.5571WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 483s 3s/step - loss: 0.9420 - accuracy: 0.5571 - val_loss: 1.1779 - val_accuracy: 0.5310 - lr: 0.0010
Epoch 15/50
170/170 [=====] - ETA: 0s - loss: 0.9382 - accuracy: 0.5608WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 485s 3s/step - loss: 0.9382 - accuracy: 0.5608 - val_loss: 0.9466 - val_accuracy: 0.5662 - lr: 0.0010
Epoch 16/50
170/170 [=====] - ETA: 0s - loss: 0.9309 - accuracy: 0.5629WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 448s 3s/step - loss: 0.9309 - accuracy: 0.5629 - val_loss: 0.9608 - val_accuracy: 0.5710 - lr: 0.0010
Epoch 17/50
170/170 [=====] - ETA: 0s - loss: 0.9370 - accuracy: 0.5591WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 445s 3s/step - loss: 0.9370 - accuracy: 0.5591 - val_loss: 0.9067 - val_accuracy: 0.5829 - lr: 0.0010
Epoch 18/50
170/170 [=====] - ETA: 0s - loss: 0.9344 - accuracy: 0.5610WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 488s 3s/step - loss: 0.9344 - accuracy: 0.5610 - val_loss: 0.9935 - val_accuracy: 0.5552 - lr: 0.0010
Epoch 19/50
170/170 [=====] - ETA: 0s - loss: 0.9319 - accuracy: 0.5645WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 487s 3s/step - loss: 0.9319 - accuracy: 0.5645 - val_loss: 0.9042 - val_accuracy: 0.5814 - lr: 0.0010
Epoch 20/50
170/170 [=====] - ETA: 0s - loss: 0.9351 - accuracy: 0.5638WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 485s 3s/step - loss: 0.9351 - accuracy: 0.5638 - val_loss: 1.0578 - val_accuracy: 0.5438 - lr: 0.0010
Epoch 21/50
170/170 [=====] - ETA: 0s - loss: 0.9303 - accuracy: 0.5655WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 478s 3s/step - loss: 0.9303 - accuracy: 0.5655 - val_loss: 0.9490 - val_accuracy: 0.5748 - lr: 0.0010
Epoch 22/50
170/170 [=====] - ETA: 0s - loss: 0.9252 - accuracy: 0.5677WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 475s 3s/step - loss: 0.9252 - accuracy: 0.5677 - val_loss: 1.0777 - val_accuracy: 0.4457 - lr: 0.0010
Epoch 23/50
170/170 [=====] - ETA: 0s - loss: 0.9317 - accuracy: 0.5654WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 488s 3s/step - loss: 0.9317 - accuracy: 0.5654 - val_loss: 0.9101 - val_accuracy: 0.5929 - lr: 0.0010
Epoch 24/50
170/170 [=====] - ETA: 0s - loss: 0.9222 - accuracy: 0.5742WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 486s 3s/step - loss: 0.9222 - accuracy: 0.5742 - val_loss: 0.9546 - val_accuracy: 0.5743 - lr: 0.0010
Epoch 25/50
170/170 [=====] - ETA: 0s - loss: 0.9151 - accuracy: 0.5733WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 451s 3s/step - loss: 0.9151 - accuracy: 0.5733 - val_loss: 0.8998 - val_accuracy: 0.5786 - lr: 0.0010
Epoch 26/50
170/170 [=====] - ETA: 0s - loss: 0.9147 - accuracy: 0.5753WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 476s 3s/step - loss: 0.9147 - accuracy: 0.5753 - val_loss: 0.8966 - val_accuracy: 0.5848 - lr: 0.0010
```

Figure 18 - Model 6 Results (Epochs 14-26)

One of the main factors that affected the traffic light dataset was the max pooling size, as increasing it decreased the accuracy and the time the program took. This resulted in a lowering of the max pooling size to (2,2) (Figure 19).

```

Model: "sequential"
-----
Layer (type)                Output Shape                Param #
-----
conv2d (Conv2D)              (None, 126, 126, 32)       896
batch_normalization (BatchN (None, 126, 126, 32)       128
ormalization)
max_pooling2d (MaxPooling2D (None, 63, 63, 32)         0
)
dropout (Dropout)            (None, 63, 63, 32)         0
conv2d_1 (Conv2D)            (None, 61, 61, 64)         18496
batch_normalization_1 (Batc (None, 61, 61, 64)         256
hNormalization)
max_pooling2d_1 (MaxPooling (None, 30, 30, 64)         0
2D)
dropout_1 (Dropout)          (None, 30, 30, 64)         0
conv2d_2 (Conv2D)            (None, 28, 28, 64)         36928
batch_normalization_2 (Batc (None, 28, 28, 64)         256
hNormalization)
max_pooling2d_2 (MaxPooling (None, 14, 14, 64)         0
2D)
dropout_2 (Dropout)          (None, 14, 14, 64)         0
flatten (Flatten)            (None, 12544)               0
dense (Dense)                 (None, 512)                 6423040
batch_normalization_3 (Batc (None, 512)                 2048
hNormalization)
dropout_3 (Dropout)          (None, 512)                 0
dense_1 (Dense)               (None, 3)                   1539
-----
Total params: 6,483,587
Trainable params: 6,482,243
Non-trainable params: 1,344
-----

```

Figure 19 - Model 7 (Optimizer = Adam, batch size = 50)

```

Epoch 1/50
170/170 [=====] - ETA: 0s - loss: 1.2300 - accuracy: 0.4881WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 1300s 8s/step - loss: 1.2300 - accuracy: 0.4881 - val_loss: 1.5148 - val_accuracy: 0.5395 - lr: 0.0010
Epoch 2/50
170/170 [=====] - ETA: 0s - loss: 1.0494 - accuracy: 0.5198WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 673s 4s/step - loss: 1.0494 - accuracy: 0.5198 - val_loss: 1.0106 - val_accuracy: 0.5267 - lr: 0.0010
Epoch 3/50
170/170 [=====] - ETA: 0s - loss: 1.0185 - accuracy: 0.5241WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 675s 4s/step - loss: 1.0185 - accuracy: 0.5241 - val_loss: 1.2810 - val_accuracy: 0.4624 - lr: 0.0010
Epoch 4/50
170/170 [=====] - ETA: 0s - loss: 0.9987 - accuracy: 0.5360WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 677s 4s/step - loss: 0.9987 - accuracy: 0.5360 - val_loss: 1.0259 - val_accuracy: 0.4895 - lr: 0.0010
Epoch 5/50
170/170 [=====] - ETA: 0s - loss: 0.9806 - accuracy: 0.5397WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 677s 4s/step - loss: 0.9806 - accuracy: 0.5397 - val_loss: 2.6474 - val_accuracy: 0.4986 - lr: 0.0010
Epoch 6/50
170/170 [=====] - ETA: 0s - loss: 0.9711 - accuracy: 0.5448WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 680s 4s/step - loss: 0.9711 - accuracy: 0.5448 - val_loss: 0.9478 - val_accuracy: 0.5610 - lr: 0.0010
Epoch 7/50
170/170 [=====] - ETA: 0s - loss: 0.9610 - accuracy: 0.5435WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 678s 4s/step - loss: 0.9610 - accuracy: 0.5435 - val_loss: 0.9867 - val_accuracy: 0.5210 - lr: 0.0010
Epoch 8/50
170/170 [=====] - ETA: 0s - loss: 0.9618 - accuracy: 0.5466WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 671s 4s/step - loss: 0.9618 - accuracy: 0.5466 - val_loss: 1.2212 - val_accuracy: 0.5367 - lr: 0.0010
Epoch 9/50
170/170 [=====] - ETA: 0s - loss: 0.9551 - accuracy: 0.5548WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 657s 4s/step - loss: 0.9551 - accuracy: 0.5548 - val_loss: 0.9541 - val_accuracy: 0.5614 - lr: 0.0010
Epoch 10/50
170/170 [=====] - ETA: 0s - loss: 0.9548 - accuracy: 0.5497WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 674s 4s/step - loss: 0.9548 - accuracy: 0.5497 - val_loss: 0.9350 - val_accuracy: 0.5724 - lr: 0.0010

```

Figure 20 - Model 7 Results (Epochs 1-10)

```

Epoch 11/50
170/170 [=====] - ETA: 0s - loss: 0.9535 - accuracy: 0.5558WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 652s 4s/step - loss: 0.9535 - accuracy: 0.5558 - val_loss: 1.5631 - val_accuracy: 0.5400 - lr: 0.0010
Epoch 12/50
170/170 [=====] - ETA: 0s - loss: 0.9550 - accuracy: 0.5531WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 675s 4s/step - loss: 0.9550 - accuracy: 0.5531 - val_loss: 1.0211 - val_accuracy: 0.5233 - lr: 0.0010
Epoch 13/50
170/170 [=====] - ETA: 0s - loss: 0.9462 - accuracy: 0.5610WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 659s 4s/step - loss: 0.9462 - accuracy: 0.5610 - val_loss: 2.1107 - val_accuracy: 0.5357 - lr: 0.0010
Epoch 14/50
170/170 [=====] - ETA: 0s - loss: 0.9472 - accuracy: 0.5594WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 663s 4s/step - loss: 0.9472 - accuracy: 0.5594 - val_loss: 0.9513 - val_accuracy: 0.5695 - lr: 0.0010
Epoch 15/50
170/170 [=====] - ETA: 0s - loss: 0.9498 - accuracy: 0.5554WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 671s 4s/step - loss: 0.9498 - accuracy: 0.5554 - val_loss: 0.9176 - val_accuracy: 0.5819 - lr: 0.0010
Epoch 16/50
170/170 [=====] - ETA: 0s - loss: 0.9442 - accuracy: 0.5581WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 663s 4s/step - loss: 0.9442 - accuracy: 0.5581 - val_loss: 5.3257 - val_accuracy: 0.4567 - lr: 0.0010
Epoch 17/50
170/170 [=====] - ETA: 0s - loss: 0.9448 - accuracy: 0.5607WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 641s 4s/step - loss: 0.9448 - accuracy: 0.5607 - val_loss: 0.9287 - val_accuracy: 0.5743 - lr: 0.0010
Epoch 18/50
170/170 [=====] - ETA: 0s - loss: 0.9385 - accuracy: 0.5648WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 663s 4s/step - loss: 0.9385 - accuracy: 0.5648 - val_loss: 0.9294 - val_accuracy: 0.5833 - lr: 0.0010
Epoch 19/50
170/170 [=====] - ETA: 0s - loss: 0.9316 - accuracy: 0.5684WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 664s 4s/step - loss: 0.9316 - accuracy: 0.5684 - val_loss: 1.6263 - val_accuracy: 0.5419 - lr: 0.0010
Epoch 20/50
170/170 [=====] - ETA: 0s - loss: 0.9282 - accuracy: 0.5720WARNING:tensorflow:Learning rate reduction is conditioned on
170/170 [=====] - 668s 4s/step - loss: 0.9282 - accuracy: 0.5720 - val_loss: 1.0865 - val_accuracy: 0.5643 - lr: 0.0010

```

Figure 21 - Model 7 Results (Epochs 11-20)

At this point, the program called for more dramatic changes, as the previous versions all had similar accuracy. It started from 3 convolution layers to 5 (Figures 22 and 23).

Layer (type)	Output Shape	Param #
conv2d_6 (Conv2D)	(None, 126, 126, 32)	896
batch_normalization_5 (Batch Normalization)	(None, 126, 126, 32)	128
max_pooling2d_5 (MaxPooling2D)	(None, 63, 63, 32)	0
dropout_5 (Dropout)	(None, 63, 63, 32)	0
conv2d_7 (Conv2D)	(None, 61, 61, 64)	18496
batch_normalization_6 (Batch Normalization)	(None, 61, 61, 64)	256
max_pooling2d_6 (MaxPooling2D)	(None, 30, 30, 64)	0
dropout_6 (Dropout)	(None, 30, 30, 64)	0
conv2d_8 (Conv2D)	(None, 28, 28, 64)	36928
batch_normalization_7 (Batch Normalization)	(None, 28, 28, 64)	256
max_pooling2d_7 (MaxPooling2D)	(None, 14, 14, 64)	0
dropout_7 (Dropout)	(None, 14, 14, 64)	0
conv2d_9 (Conv2D)	(None, 12, 12, 64)	36928
batch_normalization_8 (Batch Normalization)	(None, 12, 12, 64)	256
max_pooling2d_8 (MaxPooling2D)	(None, 6, 6, 64)	0
dropout_8 (Dropout)	(None, 6, 6, 64)	0
conv2d_10 (Conv2D)	(None, 4, 4, 64)	36928
batch_normalization_9 (Batch Normalization)	(None, 4, 4, 64)	256

Figure 22 - Model 8 part 1 (Optimizer = Adam, batch size = 50)

max_pooling2d_9 (MaxPooling 2D)	(None, 2, 2, 64)	0
dropout_9 (Dropout)	(None, 2, 2, 64)	0
flatten (Flatten)	(None, 256)	0
dense (Dense)	(None, 512)	131584
batch_normalization_10 (Batch Normalization)	(None, 512)	2048
dropout_10 (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 3)	1539

```

=====
Total params: 266,499
Trainable params: 264,899
Non-trainable params: 1,600

```

Figure 23 - Model 8 part 2

```

Epoch 1/50
566/566 [=====] - ETA: 0s - loss: 1.2429 - accuracy: 0.4671WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 5295s 9s/step - loss: 1.2429 - accuracy: 0.4671 - val_loss: 1.0558 - val_accuracy: 0.5381 - lr: 0.0010
Epoch 2/50
566/566 [=====] - ETA: 0s - loss: 1.0491 - accuracy: 0.5014WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 637s 1s/step - loss: 1.0491 - accuracy: 0.5014 - val_loss: 1.0558 - val_accuracy: 0.5447 - lr: 0.0010
Epoch 3/50
566/566 [=====] - ETA: 0s - loss: 1.0161 - accuracy: 0.5217WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 607s 1s/step - loss: 1.0161 - accuracy: 0.5217 - val_loss: 0.9687 - val_accuracy: 0.5437 - lr: 0.0010
Epoch 4/50
566/566 [=====] - ETA: 0s - loss: 1.0104 - accuracy: 0.5217WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 615s 1s/step - loss: 1.0104 - accuracy: 0.5217 - val_loss: 0.9695 - val_accuracy: 0.5461 - lr: 0.0010
Epoch 5/50
566/566 [=====] - ETA: 0s - loss: 0.9954 - accuracy: 0.5329WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 636s 1s/step - loss: 0.9954 - accuracy: 0.5329 - val_loss: 1.0168 - val_accuracy: 0.5277 - lr: 0.0010
Epoch 6/50
566/566 [=====] - ETA: 0s - loss: 0.9889 - accuracy: 0.5297WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 636s 1s/step - loss: 0.9889 - accuracy: 0.5297 - val_loss: 0.9317 - val_accuracy: 0.5730 - lr: 0.0010
Epoch 7/50
566/566 [=====] - ETA: 0s - loss: 0.9839 - accuracy: 0.5333WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 628s 1s/step - loss: 0.9839 - accuracy: 0.5333 - val_loss: 0.9324 - val_accuracy: 0.5693 - lr: 0.0010
Epoch 8/50
566/566 [=====] - ETA: 0s - loss: 0.9752 - accuracy: 0.5389WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 630s 1s/step - loss: 0.9752 - accuracy: 0.5389 - val_loss: 0.9151 - val_accuracy: 0.5726 - lr: 0.0010
Epoch 9/50
566/566 [=====] - ETA: 0s - loss: 0.9692 - accuracy: 0.5447WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 602s 1s/step - loss: 0.9692 - accuracy: 0.5447 - val_loss: 0.9513 - val_accuracy: 0.5603 - lr: 0.0010
Epoch 10/50
566/566 [=====] - ETA: 0s - loss: 0.9674 - accuracy: 0.5498WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 629s 1s/step - loss: 0.9674 - accuracy: 0.5498 - val_loss: 0.9516 - val_accuracy: 0.5641 - lr: 0.0010
Epoch 11/50
566/566 [=====] - ETA: 0s - loss: 0.9536 - accuracy: 0.5519WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 606s 1s/step - loss: 0.9536 - accuracy: 0.5519 - val_loss: 0.9208 - val_accuracy: 0.5702 - lr: 0.0010
Epoch 12/50
566/566 [=====] - ETA: 0s - loss: 0.9611 - accuracy: 0.5497WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 627s 1s/step - loss: 0.9611 - accuracy: 0.5497 - val_loss: 0.9168 - val_accuracy: 0.5787 - lr: 0.0010
Epoch 13/50
566/566 [=====] - ETA: 0s - loss: 0.9493 - accuracy: 0.5650WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 620s 1s/step - loss: 0.9493 - accuracy: 0.5650 - val_loss: 0.9108 - val_accuracy: 0.5730 - lr: 0.0010
Epoch 14/50
566/566 [=====] - ETA: 0s - loss: 0.9403 - accuracy: 0.5578WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 621s 1s/step - loss: 0.9403 - accuracy: 0.5578 - val_loss: 0.9514 - val_accuracy: 0.5678 - lr: 0.0010

```

Figure 24 - Model 8 Results

The number of convolution layers decreased in the program to 3, and a new factor, the dropout size, was lowered to 0.1. (Figure 25).

Layer (type)	Output Shape	Param #
conv2d_6 (Conv2D)	(None, 126, 126, 32)	896
batch_normalization_8 (Batch Normalization)	(None, 126, 126, 32)	128
max_pooling2d_6 (MaxPooling2D)	(None, 63, 63, 32)	0
dropout_8 (Dropout)	(None, 63, 63, 32)	0
conv2d_7 (Conv2D)	(None, 61, 61, 64)	18496
batch_normalization_9 (Batch Normalization)	(None, 61, 61, 64)	256
max_pooling2d_7 (MaxPooling2D)	(None, 30, 30, 64)	0
dropout_9 (Dropout)	(None, 30, 30, 64)	0
conv2d_8 (Conv2D)	(None, 28, 28, 64)	36928
batch_normalization_10 (Batch Normalization)	(None, 28, 28, 64)	256
max_pooling2d_8 (MaxPooling2D)	(None, 14, 14, 64)	0
dropout_10 (Dropout)	(None, 14, 14, 64)	0
flatten_2 (Flatten)	(None, 12544)	0
dense_4 (Dense)	(None, 512)	6423040
batch_normalization_11 (Batch Normalization)	(None, 512)	2048
dropout_11 (Dropout)	(None, 512)	0
dense_5 (Dense)	(None, 3)	1539
=====		
Total params: 6,483,587		
Trainable params: 6,482,243		
Non-trainable params: 1,344		

Figure 25 - Model 9 (Optimizer = Adam, batch size = 50, dropout = 0.1)

```
Epoch 1/50
566/566 [=====] - ETA: 0s - loss: 1.3469 - accuracy: 0.4679WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 677s 1s/step - loss: 1.3469 - accuracy: 0.4679 - val_loss: 1.3588 - val_accuracy: 0.5168 - lr: 0.0010
Epoch 2/50
566/566 [=====] - ETA: 0s - loss: 1.0522 - accuracy: 0.5037WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 667s 1s/step - loss: 1.0522 - accuracy: 0.5037 - val_loss: 1.2833 - val_accuracy: 0.4861 - lr: 0.0010
Epoch 3/50
566/566 [=====] - ETA: 0s - loss: 1.0174 - accuracy: 0.5169WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 672s 1s/step - loss: 1.0174 - accuracy: 0.5169 - val_loss: 1.2788 - val_accuracy: 0.5357 - lr: 0.0010
Epoch 4/50
566/566 [=====] - ETA: 0s - loss: 0.9969 - accuracy: 0.5340WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 669s 1s/step - loss: 0.9969 - accuracy: 0.5340 - val_loss: 1.0949 - val_accuracy: 0.5376 - lr: 0.0010
Epoch 5/50
566/566 [=====] - ETA: 0s - loss: 0.9952 - accuracy: 0.5342WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 657s 1s/step - loss: 0.9952 - accuracy: 0.5342 - val_loss: 2.1783 - val_accuracy: 0.5050 - lr: 0.0010
Epoch 6/50
566/566 [=====] - ETA: 0s - loss: 0.9903 - accuracy: 0.5339WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 636s 1s/step - loss: 0.9903 - accuracy: 0.5339 - val_loss: 1.0028 - val_accuracy: 0.5461 - lr: 0.0010
Epoch 7/50
566/566 [=====] - ETA: 0s - loss: 0.9920 - accuracy: 0.5352WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 657s 1s/step - loss: 0.9920 - accuracy: 0.5352 - val_loss: 1.1033 - val_accuracy: 0.5452 - lr: 0.0010
Epoch 8/50
566/566 [=====] - ETA: 0s - loss: 0.9917 - accuracy: 0.5378WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 659s 1s/step - loss: 0.9917 - accuracy: 0.5378 - val_loss: 1.5366 - val_accuracy: 0.5173 - lr: 0.0010
Epoch 9/50
566/566 [=====] - ETA: 0s - loss: 0.9883 - accuracy: 0.5330WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 642s 1s/step - loss: 0.9883 - accuracy: 0.5330 - val_loss: 1.1406 - val_accuracy: 0.5423 - lr: 0.0010
Epoch 10/50
566/566 [=====] - ETA: 0s - loss: 0.9840 - accuracy: 0.5421WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 665s 1s/step - loss: 0.9840 - accuracy: 0.5421 - val_loss: 2.8895 - val_accuracy: 0.5026 - lr: 0.0010
Epoch 11/50
566/566 [=====] - ETA: 0s - loss: 0.9788 - accuracy: 0.5305WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 666s 1s/step - loss: 0.9788 - accuracy: 0.5305 - val_loss: 0.9455 - val_accuracy: 0.5622 - lr: 0.0010
Epoch 12/50
566/566 [=====] - ETA: 0s - loss: 0.9800 - accuracy: 0.5383WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 663s 1s/step - loss: 0.9800 - accuracy: 0.5383 - val_loss: 1.9043 - val_accuracy: 0.4903 - lr: 0.0010
Epoch 13/50
566/566 [=====] - ETA: 0s - loss: 0.9670 - accuracy: 0.5505WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 665s 1s/step - loss: 0.9670 - accuracy: 0.5505 - val_loss: 0.9926 - val_accuracy: 0.5655 - lr: 0.0010
Epoch 14/50
566/566 [=====] - ETA: 0s - loss: 0.9759 - accuracy: 0.5434WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 658s 1s/step - loss: 0.9759 - accuracy: 0.5434 - val_loss: 4.1003 - val_accuracy: 0.4879 - lr: 0.0010
Epoch 15/50
566/566 [=====] - ETA: 0s - loss: 0.9748 - accuracy: 0.5441WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 659s 1s/step - loss: 0.9748 - accuracy: 0.5441 - val_loss: 0.9709 - val_accuracy: 0.5660 - lr: 0.0010
Epoch 16/50
566/566 [=====] - ETA: 0s - loss: 0.9679 - accuracy: 0.5426WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 667s 1s/step - loss: 0.9679 - accuracy: 0.5426 - val_loss: 0.9383 - val_accuracy: 0.5764 - lr: 0.0010
```

Figure 26 - Model 9 Results (Epochs 1-16)


```

Epoch 17/50
566/566 [=====] - ETA: 0s - loss: 0.9718 - accuracy: 0.5449WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 665s 1s/step - loss: 0.9718 - accuracy: 0.5449 - val_loss: 0.9324 - val_accuracy: 0.5626 - lr: 0.0010
Epoch 18/50
566/566 [=====] - ETA: 0s - loss: 0.9655 - accuracy: 0.5435WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 645s 1s/step - loss: 0.9655 - accuracy: 0.5435 - val_loss: 4.9898 - val_accuracy: 0.4600 - lr: 0.0010
Epoch 19/50
566/566 [=====] - ETA: 0s - loss: 0.9623 - accuracy: 0.5448WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 673s 1s/step - loss: 0.9623 - accuracy: 0.5448 - val_loss: 0.9571 - val_accuracy: 0.5593 - lr: 0.0010
Epoch 20/50
566/566 [=====] - ETA: 0s - loss: 0.9648 - accuracy: 0.5487WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 645s 1s/step - loss: 0.9648 - accuracy: 0.5487 - val_loss: 0.9506 - val_accuracy: 0.5778 - lr: 0.0010
Epoch 21/50
566/566 [=====] - ETA: 0s - loss: 0.9602 - accuracy: 0.5447WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 648s 1s/step - loss: 0.9602 - accuracy: 0.5447 - val_loss: 0.9421 - val_accuracy: 0.5626 - lr: 0.0010
Epoch 22/50
566/566 [=====] - ETA: 0s - loss: 0.9549 - accuracy: 0.5480WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 672s 1s/step - loss: 0.9549 - accuracy: 0.5480 - val_loss: 1.2536 - val_accuracy: 0.5546 - lr: 0.0010
Epoch 23/50
566/566 [=====] - ETA: 0s - loss: 0.9559 - accuracy: 0.5507WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 668s 1s/step - loss: 0.9559 - accuracy: 0.5507 - val_loss: 0.9251 - val_accuracy: 0.5745 - lr: 0.0010
Epoch 24/50
566/566 [=====] - ETA: 0s - loss: 0.9460 - accuracy: 0.5570WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 638s 1s/step - loss: 0.9460 - accuracy: 0.5570 - val_loss: 0.9572 - val_accuracy: 0.5641 - lr: 0.0010
Epoch 25/50
566/566 [=====] - ETA: 0s - loss: 0.9496 - accuracy: 0.5565WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 667s 1s/step - loss: 0.9496 - accuracy: 0.5565 - val_loss: 0.9281 - val_accuracy: 0.5754 - lr: 0.0010
Epoch 26/50
566/566 [=====] - ETA: 0s - loss: 0.9467 - accuracy: 0.5527WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 665s 1s/step - loss: 0.9467 - accuracy: 0.5527 - val_loss: 0.9516 - val_accuracy: 0.5674 - lr: 0.0010
Epoch 27/50
566/566 [=====] - ETA: 0s - loss: 0.9451 - accuracy: 0.5558WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 670s 1s/step - loss: 0.9451 - accuracy: 0.5558 - val_loss: 0.9485 - val_accuracy: 0.5664 - lr: 0.0010
Epoch 28/50
566/566 [=====] - ETA: 0s - loss: 0.9409 - accuracy: 0.5604WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 668s 1s/step - loss: 0.9409 - accuracy: 0.5604 - val_loss: 0.9044 - val_accuracy: 0.5863 - lr: 0.0010
Epoch 29/50
566/566 [=====] - ETA: 0s - loss: 0.9343 - accuracy: 0.5618WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 679s 1s/step - loss: 0.9343 - accuracy: 0.5618 - val_loss: 0.9452 - val_accuracy: 0.5688 - lr: 0.0010
Epoch 30/50
566/566 [=====] - ETA: 0s - loss: 0.9427 - accuracy: 0.5588WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 669s 1s/step - loss: 0.9427 - accuracy: 0.5588 - val_loss: 0.9195 - val_accuracy: 0.5730 - lr: 0.0010
Epoch 31/50
566/566 [=====] - ETA: 0s - loss: 0.9328 - accuracy: 0.5594WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 653s 1s/step - loss: 0.9328 - accuracy: 0.5594 - val_loss: 0.9793 - val_accuracy: 0.5626 - lr: 0.0010

```

Figure 27 - Model 9 Results (Epochs 17-31)

The experiment finishes with one last adjustment, increasing the dropout size to 0.5 (Figure 28).

Layer (type)	Output Shape	Param #
conv2d_11 (Conv2D)	(None, 126, 126, 32)	896
batch_normalization_11 (Batch Normalization)	(None, 126, 126, 32)	128
max_pooling2d_10 (MaxPooling2D)	(None, 63, 63, 32)	0
dropout_11 (Dropout)	(None, 63, 63, 32)	0
conv2d_12 (Conv2D)	(None, 61, 61, 64)	18496
batch_normalization_12 (Batch Normalization)	(None, 61, 61, 64)	256
max_pooling2d_11 (MaxPooling2D)	(None, 30, 30, 64)	0
dropout_12 (Dropout)	(None, 30, 30, 64)	0
conv2d_13 (Conv2D)	(None, 28, 28, 64)	36928
batch_normalization_13 (Batch Normalization)	(None, 28, 28, 64)	256
max_pooling2d_12 (MaxPooling2D)	(None, 14, 14, 64)	0
dropout_13 (Dropout)	(None, 14, 14, 64)	0
flatten_1 (Flatten)	(None, 12544)	0
dense_2 (Dense)	(None, 512)	6423040
batch_normalization_14 (Batch Normalization)	(None, 512)	2048
dropout_14 (Dropout)	(None, 512)	0
dense_3 (Dense)	(None, 3)	1539

=====
Total params: 6,483,587
Trainable params: 6,482,243
Non-trainable params: 1,344
=====

Figure 28 - Model 10 (Optimizer = Adam, batch size = 50, dropout = 0.5)

```

Epoch 1/50
566/566 [=====] - ETA: 0s - loss: 1.3593 - accuracy: 0.4642WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 674s 1s/step - loss: 1.3593 - accuracy: 0.4642 - val_loss: 1.1099 - val_accuracy: 0.4931 - lr: 0.0010
Epoch 2/50
566/566 [=====] - ETA: 0s - loss: 1.0647 - accuracy: 0.5059WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 657s 1s/step - loss: 1.0647 - accuracy: 0.5059 - val_loss: 1.1594 - val_accuracy: 0.4572 - lr: 0.0010
Epoch 3/50
566/566 [=====] - ETA: 0s - loss: 1.0131 - accuracy: 0.5252WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 649s 1s/step - loss: 1.0131 - accuracy: 0.5252 - val_loss: 1.5669 - val_accuracy: 0.4988 - lr: 0.0010
Epoch 4/50
566/566 [=====] - ETA: 0s - loss: 1.0047 - accuracy: 0.5266WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 656s 1s/step - loss: 1.0047 - accuracy: 0.5266 - val_loss: 1.7194 - val_accuracy: 0.4700 - lr: 0.0010
Epoch 5/50
566/566 [=====] - ETA: 0s - loss: 0.9998 - accuracy: 0.5256WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 651s 1s/step - loss: 0.9998 - accuracy: 0.5256 - val_loss: 1.9027 - val_accuracy: 0.5243 - lr: 0.0010
Epoch 6/50
566/566 [=====] - ETA: 0s - loss: 0.9948 - accuracy: 0.5290WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 656s 1s/step - loss: 0.9948 - accuracy: 0.5290 - val_loss: 3.5446 - val_accuracy: 0.4950 - lr: 0.0010
Epoch 7/50
566/566 [=====] - ETA: 0s - loss: 0.9946 - accuracy: 0.5342WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 642s 1s/step - loss: 0.9946 - accuracy: 0.5342 - val_loss: 1.4639 - val_accuracy: 0.5002 - lr: 0.0010
Epoch 8/50
566/566 [=====] - ETA: 0s - loss: 0.9965 - accuracy: 0.5319WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 646s 1s/step - loss: 0.9965 - accuracy: 0.5319 - val_loss: 1.0310 - val_accuracy: 0.5324 - lr: 0.0010
Epoch 9/50
566/566 [=====] - ETA: 0s - loss: 0.9915 - accuracy: 0.5315WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 645s 1s/step - loss: 0.9915 - accuracy: 0.5315 - val_loss: 1.0786 - val_accuracy: 0.5409 - lr: 0.0010
Epoch 10/50
566/566 [=====] - ETA: 0s - loss: 0.9893 - accuracy: 0.5309WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 615s 1s/step - loss: 0.9893 - accuracy: 0.5309 - val_loss: 0.9682 - val_accuracy: 0.5537 - lr: 0.0010
Epoch 11/50
566/566 [=====] - ETA: 0s - loss: 0.9832 - accuracy: 0.5316WARNING:tensorflow:Learning rate reduction is conditioned on
566/566 [=====] - 649s 1s/step - loss: 0.9832 - accuracy: 0.5316 - val_loss: 1.0302 - val_accuracy: 0.5404 - lr: 0.0010

```

Figure 29 - Model 10 Results

The most influential result in this experiment is the dataset. When the switch between datasets was made, the accuracy went from hovering around a specific number to hovering around another number. One of the defining factors for the time per epoch was the max pooling. When the max pooling had a larger area, the time was reduced, as there were fewer parameters to deal with. However, many parameters did not change the results significantly. For example, the dropout size, the number of convolution layers, the batch size, and the optimizer. Aside from that, one of the influential factors is the number of epochs that were able to run, as generally, the higher number of epochs that ran allowed for higher data accuracy.

Model #	Data Set	Convolution Layers	Max Pooling Size	Dropout Size	Batch Size	Optimizer	Activation Function	Peak Accuracy	Number of Parameters	Approximate Average Time per Epoch (seconds)
1	Traffic Light Classification	3	2,2	0.5	15	rmsprop	relu	1	12,943,812	145
2	Traffic Light Classification	2	2,2	0.5	15	rmsprop	relu	0.9973	29,515,588	152
3	Traffic Light Classification	4	2,2	0.5	15	rmsprop	relu	1	1,278,404	108
4	Lung Disease Classification	4	3,3	0.25	15	rmsprop	relu	0.5783	1,278,404	520
5	Lung Disease Classification	4	3,3	0.25	15	Adam	relu	0.5759	1,278,404	490
6	Lung Disease Classification	3	3,3	0.25	50	Adam	relu	0.5929	355,971	480
7	Lung Disease Classification	3	2,2	0.25	50	Adam	relu	0.5833	6,483,587	665
8	Lung Disease Classification	5	2,2	0.25	50	Adam	softplus	0.5787	266,499	610

9	Lung Disease Classification	3	2,2	0.1	50	Adam	softplus	0.5863	6,483,587	660
10	Lung Disease Classification	3	2,2	0.5	50	Adam	softplus	0.5537	6,483,587	650

Table 1 - Data Summary

There are a few potential reasons why those factors did not change the program as intended. For example, the number of convolution layers did not change anything. Although the increase in neurons guarantees a better-fitting trend, the model was already optimized without the additional layers. Additionally, this generates reasoning similar to the different optimizers and batch sizes. Both optimizers and batch sizes optimize the program, and they do it to a level such that the minor differences in what they do barely influence the results. The dropout size potentially did not affect the results because of the neurons chosen to be dropped.^{1,5}

Conclusion

In this paper, I demonstrated how to optimize a particular neural network with two datasets. Although the same modifications may be less effective on other datasets, the project demonstrates an effective template for how different aspects of neural networks can improve. Optimized neural networks can finish a task in a shorter time, so less money must be spent. That money can go towards even further optimizing the neural network.

Enhancing the performance of neural networks allows for higher efficiency, causing a given neural network to be faster and more computationally efficient. We discovered that the dataset heavily influences the neural networks' accuracy and speed. Still, as the specific task assigned to a neural network does not vary, neural networks must be optimized to their highest capability. Significantly larger neural networks, as non-efficient ones would end up as a large amount of money squandered into an AI that pales in comparison to more efficient ones.^{2,4}

Acknowledgements

Thank you for the guidance of Cyrus Ayubcha, a mentor from Harvard University, in developing this research paper.

Authors

Zubin Gupta is a Lexington High School sophomore pursuing a career in Computer Science. He also has a focus on charity with his nonprofit, the Twynphony Hopes Foundation.

References

1. Broderick, T. *Impact of CNNs AlexNet*. (2010).

<https://introml.mit.edu/static/spring23/VideoSlides/cnn.pdf>

2. Overview of Artificial Intelligence Technology | FINRA.org. (2013). Finra.org.

<https://www.finra.org/rules-guidance/key-topics/fintech/report/artificial-intelligence-in-the-securities-industry/overview-of-ai-tech>

3. Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021). *Learning Transferable Visual Models From Natural Language Supervision*. ArXiv.org. <https://arxiv.org/abs/2103.00020>

4. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. <https://www.deeplearningbook.org/>
5. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need*. ArXiv.org. <https://arxiv.org/abs/1706.0376>

Academic Bias in Liberal and Conservative Areas: Analysis of Gender and Sexual Orientation Bias in Papers from Democrat and Republican-Leaning Regions By Keilah Park

Abstract

This research paper investigates the bias of research papers produced by fifty universities. Twenty five each were situated in highly Democrat-leaning and Republican-leaning regions of the United States. The study aims to discern whether there is a significant difference in the latent written bias found in publications from each of these demographics and whether measurable language bias aligns with the prevailing political and social biases of the respective local electorate. By focusing on themes related to gender and sexual orientation, the research delves into areas where political and social attitudes intersect. The methodology involves a systematic sampling approach, retrieving a representative sample of research papers, and employing advanced Natural Language Processing (NLP) techniques for thematic analysis. The study reveals the complex relationship between academic research, regional political climate, and societal bias, highlighting the nuanced factors that influence academia. Through interdisciplinary inquiry, the research contributes to a deeper understanding of bias within academia and its implications for inclusivity and intellectual rigor. The findings suggest a difference in the average bias scores between papers from universities that are leaning in the opposite political directions, with Republican-leaning papers showing higher levels of detected bias.

Key words: Bias, NLP, ML, AI, GenBiT, Gender, Sexual Orientation, Politics

Introduction

The influence of political and social ideologies on everything from Large Language Models (LLMs) to trans prejudice is being actively studied (Campbell et al., 22). This paper investigates the influence of political ideologies on academic output. Some believe journal papers are less influenced by bias due to the peer review process. Data needs to be collected to measure bias in papers. In this paper, the author attempts to quantify bias based on geographic locations where the electorate differ greatly on the political spectrum. This paper explores the possible influence of regional political ideologies in research papers produced by universities situated in highly Democrat-leaning and Republican-leaning regions. Specifically, this research aims to investigate whether there is a significant difference in the latent linguistic bias found in publications and whether these align with the prevailing political and social biases of their respective regions. The importance of this question lies in its potential to uncover the nuanced ways in which regional political landscapes may influence academic discourse. Past work has highlighted the impact of political and ideologies on research agendas and scholarly perspectives (Chapman et al., 6) (Cooper, 2) (Doughman and Khreich, 1) (Gerber et al., 385) (Nosek et al., 10593) (Okruhlik, 21) (Prpic, 27) (Ringgenberg et al., 1). However, there remains a lack of

research specifically examining how written bias may mirror prevailing ideologies of the regions in which the academic institutions are located. This study aims to address this gap by focusing on the measurable bias in university publication, specifically bias regarding gender and sexual orientation - areas in which have been identified as crucial to contemporary political discourse (Waites, 137). Gender bias is defined as bias where there is a bias towards the cis-man vs the cis-woman. While there is significant work detailing gender bias in the peer review process as well as bias in scholarly work or research itself, there is no work analyzing geographic political leanings and the biases present in published output from these universities (Eitan et al., 189) (Gelber et al., 199) (Gerber et al., 385) (Honeycutt and Jussim, 73). Through the use of quantitative analysis, this research seeks to determine whether there is a significant difference in gender-biased language in journal papers. This paper's overall question is: Is there a significant difference in gender or sexual orientation-related bias in research papers produced by universities in highly Democrat-leaning regions compared to universities located in highly Republican-leaning regions? We will focus on whether these schools' publications reflect bias regarding gender and sexual orientation, where more bias may be expected from conservative-voting regions.

To address these questions, this study employs a rigorous methodology drawing on a diverse sample of research papers from universities across various regions of the United States. Utilizing advanced Natural Language Processing (NLP) techniques, the author aims to uncover subtle linguistic cues indicative of underlying ideological biases present within the academic discourse (Rekabsaz and Schedl). Language that implies bias for gender and sexual orientation has been researched and tools have been already developed to identify and quantify the bias (Doughman and Khreich, 8) (Sengupta et al., 10) (Microsoft GenBiT). The author utilizes these computational tools for this analysis.

Literature Analysis

Gender Bias (cis-man cis-woman patriarchal bias)

The pursuit of scientific knowledge is often described as an objective, relentless pursuit of the truth. However, a growing body of research has shown a constant problem in scientific research - the pervasive issue of gender bias. Gender bias is not confined to a single domain of science but permeates across multiple disciplines. Colin D. Chapman, a professor in the department of neuroscience, with his colleagues, categorizes the effects of experimenter gender bias into three overarching areas: the mind, the body, and behavior (Chapman et al., 1). In the field of medical research, inherent bias may lead to suboptimal outcomes when it comes to developing interventions or therapeutics. This underscores the need to report the gender of the researchers and study subjects to control for possible bias. Also viewed in this perspective, Okruhlik, a professor at the University of Western Ontario, emphasized the importance of feminist analyses in biology. This analysis delves into deep-rooted biases within the field of biology, highlighting the systemic nature of these issues and their impact on research

methodology and outcomes (Okruhlik, 40). Emphasizing the connection between social influences and scientific claims can lead to better integration of gender-conscious perspectives into scientific discourse and practice. The gender bias issue is not limited to the laboratory or academic research. Professor Nosek, under the department of psychology at the University of Virginia, reveals that implicit gender science stereotypes at the national level significantly predict disparities in science and math achievement between genders. This finding highlights the interconnectedness of implicit stereotypes and gender disparities in science performance, underlining the necessity for national policies to address both factors simultaneously (Nosek et al., 10594). To bridge the gender gap in scientific pursuits, addressing implicit stereotypes and the cultural realities that limit women's pursuit of scientific careers becomes imperative.

Gender bias extends to the realm of scientific productivity as well, as demonstrated by Prpic (28). Among young researchers, feminine-presenting scientists tend to publish fewer scientific papers than their more masculine-presenting counterparts, despite similar socio-demographic and educational factors influencing both ends of the gender spectrum. This study underscores the widening gender gap in scientific productivity and the need to address these disparities (Prpic, 57). In addition to this, Barbara Norrander and Clyde Wilcox, both professors of political science, highlight the gender gap in ideological alignment, which has seen both men and women shifting towards conservatism since the 1970s with a segment of women continuing to stay liberal. They attribute this trend to a growing alignment with specific ideologies among individuals and changing demographics, highlighting the importance of issues and influences such as church attendance, beliefs about religion, abortion, and women's roles. These shape an individual's social, cultural, and political beliefs. The study suggests that educational differences have become more prominent between the genders, emphasizing the role of cultural divisions in shaping ideological identities (Norrander and Wilcox, 521) (Parker). To support this, Katharine Gelber delves into the gender bias present in student evaluation of teachers, highlighting the existence of gendered expectations placed on university teachers. Female-identified teachers are often assessed based on nurturing aspects, while male-identified teachers are evaluated primarily for their technical expertise and lecture structure (Gelber et al., 200). These gendered evaluations can have profound implications for career progression, reinforcing negative impacts and underrepresentation of women in senior positions within universities.

The pervasive influence of societal norms and expectations, both implicit and explicit, manifests itself in various scientific disciplines, impacting medical therapy, policy, productivity, and career development. Recognizing and addressing these biases is crucial for achieving gender equality in the domain of science, and for preserving the integrity and validity of scientific findings (Messner).

Sexual Orientation

There have been past studies that have delved into the impact of sexual orientation on attitudes related to sexuality, race, and gender (Denise, 127). Using a nationally representative sample, J. Michael Bailey and his colleagues conducted a study which revealed that individuals with non-heterosexual orientation tend to exhibit more liberal perspectives on issues such as sexuality, racial, and gender equality. The study highlights the role of political ideology in shaping these attitudes, emphasizing the relationship between sexual identity, socialization, and distinct worldviews. The paper also acknowledged the need for further research to fully discover the relationship between sexual orientation, social attitudes, and political leanings. Research was conducted where the intersection of ongoing global political disputes regarding the acceptance of homosexuality with scientific debates on sexual orientation was examined (Bailey et al., 46). The comprehensive review emphasizes the importance of addressing sexual orientation in an unbiased and informative manner, stressing the need for tolerance rather than suppression.

Gap in the Body of Knowledge

The exploration of existing studies on how gender and sexual orientation impact academia has revealed a gap in understanding the potential correlation between regional political affiliations and biases based on gender or sexual orientation. This gap prompts further research into whether there are discernible connections between regional political leanings and the prevalence of bias in academia related to gender identity or sexual orientation. Addressing this gap in knowledge could provide valuable insights into the complex interplay between sociopolitical factors and biases within academic settings.

Methodology

To address the research question outlined in this study, a sampling approach adhering to a pre-developed methodology was employed to select the university sampling pools based on political demographics of the local region. The selection criteria was based on the electoral history and political leaning of the regions where the universities are located. A total of ten universities (five from each political-leaning category) were included in the experimental sample. Within each university, five papers were selected for analysis, with a total of twenty-five research papers representing each political demographic. Universities are often left-leaning vs the outlying region and often show differences in voting statistics vs the county or the state. For example, college professors largely lean democratic (Langbert and Stevens). In this study, schools where the micro-political environment at the university or city matches the macro-political environment at the county and/or state level were chosen. Colleges that were perceived to be liberal or conservative were identified based on the articles “Most Liberal Students” and “Most Conservative Students” by The Princeton Review (Princeton Review). These school rankings were based on student surveys surveying over 150,000 students. Down-selecting from this list, ten universities were chosen after analysis of county political results of the past US presidential election (CNN).

By leveraging reported electoral history and political affiliation data, the study ensured the accurate categorization of universities into Democrat-leaning and Republican-leaning groups. Voting data was obtained from CNN's analysis pages (CNN). By cross-referencing CNN's political demographic data with The Princeton Review's rankings, schools located in counties with strong Democratic or Republican bias were chosen. Five universities each were selected to represent Democratic-Party-leaning and Republican-Party-leaning universities.

To represent universities and regions that leaned heavily towards voting for the Democratic Party, the following five universities were chosen (the percentage of the electorate that within the county and state that voted for the Democratic Party is noted for each school): Reed College (Portland, OR; Multnomah County 79.2%; OR 56.5%), University of California Santa Cruz (UCSC) (Santa Cruz, CA; Santa Cruz County 78.9%; CA 63.5%), Eugene Lang College of Liberal Arts of The New School (New York, NY; New York County 76%; NY 60.9%), New York University (New York, NY; New York County 76%; NY 60.9%), and Seattle University (Seattle, WA; King County 74.9%; WA 58%). It is interesting to note that all of these schools were on the west coast or the northeast coast. Also, it is important to note that, for all these chosen universities, the majority of the state electorate voted for the Democratic Party. The universities selected where the electorate comparatively voted for the Republican Party are as follows (with the percentage of republican votes denoted): Clemson University (Clemson, SC; Pickens County 74.5% and Anderson County 70.3%; SC 55.1%), Wabash College (Crawfordsville, IN; Montgomery County 73.6%; IN 57%), Hillsdale College (Hillsdale, MI; Hillsdale County 73.1%; MI 49.4%), Baylor University (Waco, TX; McLennan County 60.8%; TX 52.1%), and Auburn University (Auburn, AL; Lee County 59.1%; AL 62%). Similar to the left-leaning schools, the electorate majority voted Republican. These schools are located in the midwest and the south.

The next step involved retrieving a representative sample of research papers from the selected universities. A comprehensive search strategy was devised to select a representative selection of papers that might contain bias but not be specifically about gender or sexual orientation bias. The time frame for data collection encompassed the past five years (2019-2024) to ensure a contemporary representation of scholarly output. Five papers were selected using the following method for each university. Utilizing advanced journal literature searches with Scopus, a citation database, papers that were in the subject areas of 'Social Sciences' and 'Arts and Humanities' were found that had the word "political" as the search term. The results were sorted by citation count and five papers with the highest citation numbers that were available for the author to download were selected. Importantly, papers that mentioned analysis of gender bias and sexual orientation bias were removed from the selection pool to prevent the computational analysis from analyzing examples of language within the paper as actual biased language used by the authors of the paper.

The research papers obtained from the selected universities were subjected to statistical analysis after processing the entire text of the papers with Microsoft GenBiT, a tool that utilizes

NLP to analyze bias in written language (Microsoft GenBiT). Instead of simple statistical analysis of the text for occurrences of biased phrases, the GenBiT tool utilizes statistical methods and NLP to analyze co-occurrence of words (Sengupta et al., 7). The analysis entails assessing the prevalence and distribution of the identified themes across the research papers. Average bias, non binary bias, trans-cis bias, std bias, std non binary bias, std trans-cis bias, percentage of cis gender definition words, female gender definition words, male gender definition words, non binary gender definition words, trans gender definition words, frequency of female gender definition words, male gender definition words, and non binary gender definition words are all analyzed and quantified using the GenBiT tool (Microsoft GenBiT). Although the GenBiT tool allows the researcher to see all these different categories, the overall bias score was primarily used for this research. An overall score (GenBiT score) was assessed based on the overall prevalence of gender bias detected by the machine learning model. The authors of the GenBiT tool note that a GenBiT score higher than 0.3% is considered as having a moderate to high level of bias (Sengupta et al., 6). The higher the percentage, the higher the bias.

However, to simplify analysis of how biased a paper is, the author chose to create 8 levels of bias by defining ranges of scores in order to bin the papers into levels. Each level has a range of scores with a span of 0.15%. The lower the level/GenBiT score, the lower the bias. Since the bias being measured is gender and sexual-orientation, one might expect more bias to be measured from right-leaning/conservative/Republican schools.

Table 1: Definition of levels and ranges for GenBiT scores

Level	GenBiT Score %
1	$0 \leq \text{score} < 0.15$
2	$0.15 \leq \text{score} < 0.30$
3	$0.30 \leq \text{score} < 0.45$
4	$0.45 \leq \text{score} < 0.60$
5	$0.60 \leq \text{score} < 0.75$
6	$0.75 \leq \text{score} < 0.90$
7	$0.90 \leq \text{score} < 1.05$
8	$1.05 \leq \text{score} < 1.20$

Statistical methods were used to determine whether there is a correlation between gender-bias and sexual orientation bias vs political voting history of the schools. After collecting the outputs of all fifty papers (twenty-five from each political party), these papers were categorized under the level ranking system included above. The average GenBiT score

percentage was calculated for the liberal and conservative schools and each paper was categorized into a certain bin (levels 1-8). The higher the level the paper scores under, the higher the bias that was detected.

This research adhered to ethical standards in data collection and analysis. All research papers included in the study were anonymized, and any identifying information about authors were removed to ensure confidentiality.

Data

Raw Data

GenBiT Scores

Table 2: Raw GenBiT Scores for the 50 papers that were analyzed

Democrat-Leaning Universities; GenBiT Score (%)	Republican-Leaning Universities; GenBiT Score (%)
Reed 1; 0.020	Clemson 1; 0.230
Reed 2; 0.031	Clemson 2; 0.079
Reed 3; 0.060	Clemson 3; 0.740
Reed 4; 0.046	Clemson 4; 0.014
Reed5; 0.404	Clemson 5; 0.005
UCSC 1; 0.222	Wabash 1; 0.112
UCSC 2; 0.044	Wabash 2; 0.119
UCSC 3; 0.008	Wabash 3; 0.064
UCSC 4; 0.182	Wabash 4; 0.004
UCSC 5; 0.0823	Wabash 5; 0.031
Eugene 1; 0.198	Hillsdale 1; 0.904
Eugene 2; 0.572	Hillsdale 2; 0.512
Eugene 3; 0.004	Hillsdale 3; 0.822
Eugene 4; 0.110	Hillsdale 4; 1.119
Eugene 5; 0.004	Hillsdale 5; 1.155
NYU 1; 0.003	Baylor 1; 0.046
NYU 2; 0.393	Baylor 2; 0.028
NYU 3; 0.393	Baylor 3; 0.071
NYU 4; 0.118	Baylor 4; 0.027
NYU 5; 0.007	Baylor 5; 0.548
Seattle 1; 0.093	Auburn 1; 0.300
Seattle 2; 0.050	Auburn 2; 0.057
Seattle 3; 0.621	Auburn 3; 0.112
Seattle 4; 0.019	Auburn 4; 1.047

Seattle 5; 0.0245	Auburn 5; 0.033
-------------------	-----------------

Democratic/Liberal

(Papers are identified by university name and number “University #”)

Average GenBiT Score = Sum of all Percentage / Total Number of Papers = 3.709 / 25

= **0.148%** (rounded to the nearest thousandths place)

Table 3: Categorization of liberal papers under respective levels

Level	Paper
1	Reed 1 Reed 2 Reed 3 Reed 4 UCSC 2 UCSC 3 UCSC 5 Eugene 3 Eugene 4 Eugene 5 NYU 1 NYU 4 NYU 5 Seattle 1 Seattle 2 Seattle 4 Seattle 5 Total: 17
2	UCSC 1 UCSC 4 Eugene 1 Total: 3
3	Reed 5 NYU 2 NYU 3 Total: 3
4	Eugene 2

	<u>Total: 1</u>
5	Seattle 3 <u>Total: 1</u>
6	<u>Total: 0</u>
7	<u>Total: 0</u>
8	<u>Total: 0</u>

Republican/Conservative

(Papers are identified by university name and number “University #”)

Average GenBiT Score = Sum of all Percentage / Total Number of Papers = 8.179 / 25

= **0.327%** (rounded to the nearest thousandths place)

Table 4: Categorization of conservative papers under respective levels

Level	Paper
1	Clemson 2 Clemson 4 Clemson 5 Wabash 1 Wabash 2 Wabash 3 Wabash 4 Wabash 5 Baylor 1 Baylor 2 Baylor 3 Baylor 4 Auburn 2 Auburn 3 Auburn 5 <u>Total: 15</u>
2	Clemson 1 Auburn 1 <u>Total: 2</u>
3	<u>Total: 0</u>

4	Hillsdale 2 Baylor 5 Total: 2
5	Clemson 3 Total: 1
6	Hillsdale 3 Total: 1
7	Hillsdale 1 Total: 1
8	Hillsdale 4 Hillsdale 5 Auburn 4 Total: 3

Analysis and Findings

Figure 1: Percent of liberal papers vs. bias level

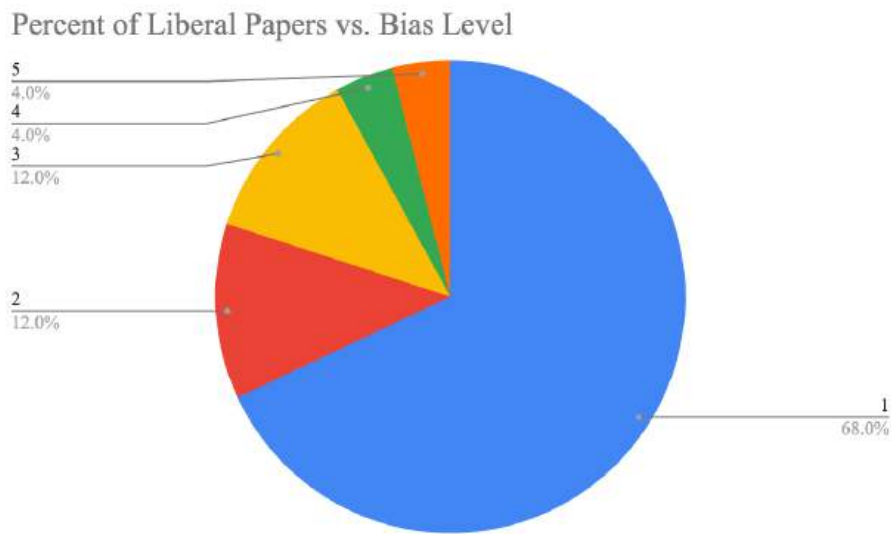


Figure 2: Percent of conservative papers vs. bias level

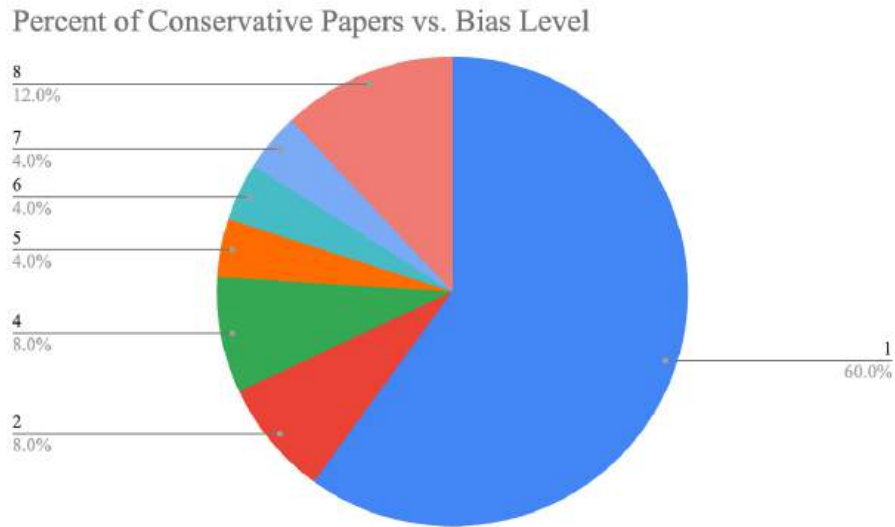
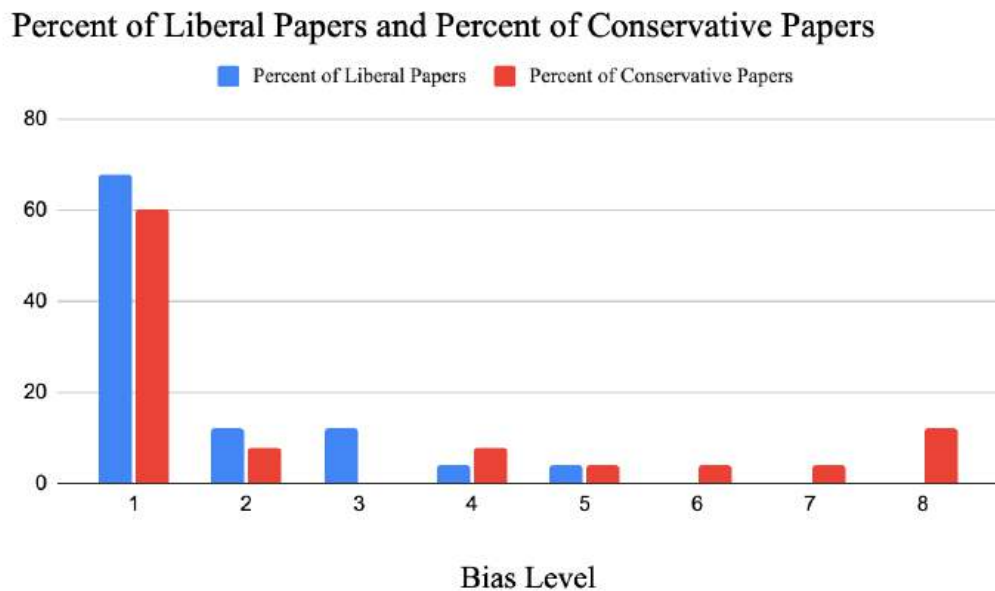


Figure 3: Percent of liberal papers and percent of conservative papers within each bias level category



The analysis of the aggregate average GenBiT scores reveals a difference, with the democratic pool of papers averaging 0.148% and the republican papers higher at 0.327%. This suggests a significant difference in average GenBiT scores as a score of 0.3 already indicates moderate to high bias. Using the definition Microsoft uses to define bias with this tool, 20% of liberal college papers showed moderate to high bias while that number rises to 32% for papers

from conservative colleges. This is a 60% higher number of papers that the GenBiT tool determines has moderate to high amounts of biased language in conservative schools vs liberal schools. The median score for either liberal or conservative colleges is still within bias level 1. This indicates that, generally, most papers do not have significantly detectable gender or sexual-orientation-related bias. This can easily be seen in Figures 1 and 2. Figure 3 clearly shows that there are more biased papers (level 3 and higher) for conservative-leaning schools. Among liberal universities, 17 papers were categorized at level 1, 3 at level 2 and 3, 1 at level 4 and 5, and none at level 6, 7, and 8. In contrast, conservative universities yield 15 papers that fall under level 1, 2 under level 2, none in level 3, 2 in level 4, 1 in level 5, 6, and 7, and 3 at level 8. Utilizing a threshold of 0.3% to signify moderate to high bias levels, levels 1 and 2 may indicate little to no bias while levels 3 to 8 indicate higher levels of gender/sexual orientation-related bias. Within liberal publications, 20 papers are classified under levels 1 and 2, while 5 fall within levels 3 to 8. Conversely, conservative papers comprise 17 at levels 1 and 2 and 8 at levels 3 to 8. From the liberal paper pool, there are no instances of the three highest bias levels (6, 7, and 8), whereas conservative university papers contain 5 total papers scoring from $0.75 \leq \text{score} < 1.20$. This discrepancy suggests a higher probability that universities within republican demographics may produce papers exhibiting greater bias. Future research would benefit from expanding the sampling pool to achieve a more accurate depiction of potential bias in scholarly works. Nevertheless, there is a clear indication of bias difference between publications from schools within regions of different political affiliations.

Limitations

The research question posed, investigating the influence of regional political ideologies on academic bias from highly Democrat-leaning versus highly Republican-leaning regions, presents several notable limitations. One of the inherent limitations of this study is that universities are not monolithic entities and may encompass diverse viewpoints even within regions traditionally associated with one political ideology. The sample size, though representative, may not capture the entire diversity within the political demographics or even each university. Thus, while there may be overarching trends aligning with regional political leanings, there likely will be significant variation and nuance within individual academic departments and the bias may also depend on the specific research agenda. Additionally, the study focuses on NLP for thematic analysis and does not delve into the specific content of individual papers which could provide further context to the findings. The methods section outlines the systematic approach employed to investigate the influence of political and social ideologies on academia. More detailed quantitative analysis using different methods is expected to yield meaningful insights into the nuanced ways in which regional political landscapes may shape academic discourse within universities. While the research question of this article aims to discern whether publications align with the political and social biases of their respective areas concerning gender and sexual orientation, it inherently assumes a direct correlation between the

political leanings of an institution's location and the content of its academic output. Such an assumption overlooks the nuanced factors that contribute to scholarly research, including individual academic interests, institutional culture, and research funding sources. Consequently, while geographic location may influence broader institutional ideologies, it may not necessarily dictate the specific thematic content of research papers produced by these universities.

Moreover, the utilization of machine learning techniques, such as GenBiT, to analyze gender bias in the datasets, introduces additional limitations. While machine learning algorithms offer valuable insights into large datasets, they are not immune to biases inherent in the data they are trained on. Biases in training data, such as gender stereotypes present in academic publications, can inadvertently influence the outcomes of machine learning analyses, potentially leading to inaccurate results. Furthermore, the effectiveness of GenBiT in detecting and mitigating gender bias may vary across different languages and cultural contexts, raising questions about the generalizability and reliability of its findings, particularly when applied to politically-charged topics.

The reliance on CNN for reported political demographics and The Princeton Review for selection of the most conservative/liberal schools introduces methodological limitations. CNN's reporting of political demographics may not capture the full spectrum of political ideologies within a given region, potentially oversimplifying the political landscape. Similarly, The Princeton Review's rankings may reflect subjective criteria and student survey bias that do not fully capture the diversity of political opinions within individual institutions. As a result, using these sources to categorize universities as highly Democrat-leaning or Republican-leaning may introduce bias and inaccuracies into the analysis, undermining the validity of the research findings. Furthermore, the focus on only the subject areas of 'Social Sciences' and 'Arts and Humanities' as well as the singular keyword of "political" may overlook other important factors that influence the thematic content of research papers, such as academic specialization, faculty expertise, and disciplinary trends. Research of all fields encompass a wide range of topics and methodologies, and while politics is undoubtedly a significant area of inquiry, it represents only a subset of the broader landscape of academic research. Failing to consider other relevant variables may limit the depth and scope of the analysis, potentially leading to incomplete or biased conclusions about the relationship between geographic location, institutional ideology, and bias. As mentioned above, the number of analyzed schools and papers is also a limitation. For example, all five papers analyzed from Hillsdale university showed significant bias. Further analysis may show that only particular universities or faculty members output significantly biased papers with very little correlation to the voter demographic.

The author does recognize that, at least in the field of gender and sexual-orientation bias, it is difficult to establish whether there is a liberal bias. Given the tools available, it was much easier to detect bias that may indicate a conservative bias. Also, the author acknowledges that due to the author's own gender and location, the author of this paper likely has a liberal bias.

Lastly, the complexity of the research question and the interdisciplinary nature of the inquiry present challenges in terms of data collection, analysis, and interpretation. All academic research is inherently multidimensional, encompassing diverse theoretical frameworks, methodological approaches, and empirical findings. As such, capturing the full extent of thematic content across research papers produced in different regions requires careful consideration of sampling strategies, data preprocessing techniques, and analytical methodologies. Failure to address these methodological challenges may compromise the validity and reliability of the study findings, limiting its contribution to the broader scholarly discourse on political ideology and academic research.

Implications

To ensure fair and rational discourse, the author believes that it is important to be aware of societal bias and believes that institutions and governments must understand the prejudices inherent in certain areas. These institutions should try to inform undergraduate students by exposing them to objective data supporting the existence of bias and to proactively address these possible complications. The academic world and the general public must analyze discourse at universities to understand whether such biases exist and, more importantly, to quantify the strength of polarization within an institution. Certain groups or populations within highly polarized universities may feel left out or otherwise ostracized. The data would also be critical to understanding the role universities play in forming or strengthening political views.

The implications of this research are multifaceted and hold significance for various stakeholders within academia and beyond. The new understanding of the gap in the body of knowledge, in which it was seen that papers originating from liberal-leaning universities displayed lesser degrees of detected bias compared to those originating from conservative institutions, helps shape future academic discourse. Firstly, by uncovering potential correlations between regional political affiliations and biases based on gender or sexual orientation, this study could inform institutional policies and practices aimed at promoting equity and inclusivity. Understanding these connections may enable universities and educational institutions to develop targeted interventions to mitigate bias and create more supportive environments for marginalized groups.

Additionally, this research could have implications for broader societal debates and policy discussions regarding diversity, equity, and inclusion. By shedding light on the intersection of political ideology and biases in academic contexts, policymakers and advocacy groups may gain insights into how sociopolitical factors influence perceptions and treatment of individuals based on gender identity or sexual orientation. This understanding could inform efforts to promote equality and combat discrimination in various sectors beyond academia.

Furthermore, the findings of this study may also have practical implications for individuals navigating academic and professional environments. Awareness of potential biases tied to regional political affiliations could empower individuals to advocate for themselves and

seek out supportive networks and resources. Ultimately, by addressing this gap in knowledge and highlighting its implications, this research contributes to ongoing efforts to create more equitable and inclusive spaces within academia and society at large.

Future Directions

Future work should try to overcome the limitations in this study by analyzing a larger amount of data. Also, work should be done to identify liberal bias (vs the mainly conservative bias measured in this paper) and other areas of bias. Analyzing non-academic bias such as articles in local newspapers may also be an interesting topic and can be compared with the analysis of academic journal publications.

Institutions and governments must try to work to address bias and prejudice in areas where measured bias is high. Quantifying the strength of bias or polarization within an institution is an important step for policy makers and leaders to develop plans and mitigation efforts. This will ensure that all individuals can feel welcome and safe. The data would also be critical to understanding the intricate relationship between academia and the general public.

Conclusion

This study has sought to investigate the influence of political ideology bias in the content of research papers produced by universities situated in highly Democrat-leaning and Republican-leaning regions measured by voter outcome in the 2020 elections. By employing advanced NLP techniques, the research has revealed notable insights into the prevalence of gender and sexual orientation bias within academic discourse. The findings indicate a significant difference in the average GenBiT scores between papers from universities aligned with different political affiliations. Papers representing liberal universities exhibited lower levels of detected bias compared to those from conservative institutions. Moreover, the distribution of papers across bias categorization levels suggests a higher likelihood of moderate to high bias in publications from universities located in Republican demographics, particularly at the higher bias levels. It is important to note that while the sample size is representative, it may not fully capture the diversity of political demographics, and rankings may oversimplify the political landscape and introduce inaccuracies. Additionally, while the study explores the link between geographic location, local politics, and academia with regards to gender and sexual orientation biases, it's crucial to acknowledge that other factors like academic specialization and disciplinary trends could also influence the paper's content. Future research could benefit from expanding the sampling pool and considering additional variables to achieve a more comprehensive understanding of the complex interplay between political ideologies and academic discourse. Moreover, efforts to raise awareness of societal biases within academic institutions and promote rational discourse are crucial in fostering inclusivity and mitigating polarization within the academic community.

Works Cited

- Bailey, J. Michael. "Sexual Orientation, Controversy, and Science." *Association for Psychological Science - APS*, 25 Apr. 2016, www.psychologicalscience.org/publications/sexual_orientation.html.
- Campbell, Marianne, et al. "A Systematic Review of the Relationship between Religion and Attitudes toward Transgender and Gender-Variant People." *The International Journal of Transgenderism*, U.S. National Library of Medicine, 19 Feb. 2019, www.ncbi.nlm.nih.gov/pmc/articles/PMC6830999/.
- Chapman, Coline D, et al. "Experimenter Gender and Replicability in Science." *Science Advances*, U.S. National Library of Medicine, Jan. 2018, pubmed.ncbi.nlm.nih.gov/29349293/.
- Cooper, Richard S. "Race in Biological and Biomedical Research." *Cold Spring Harbor Perspectives in Medicine*, U.S. National Library of Medicine, 1 Nov. 2013, www.ncbi.nlm.nih.gov/pmc/articles/PMC3808769/.
- Denise, Eric Joy. "Sexual Orientation Differences in Attitudes about Sexuality, Race, and Gender." *Social Science Research*, Academic Press, 20 May 2016, www.sciencedirect.com/science/article/abs/pii/S0049089X15300144.
- Doughman, Jad, and Wael Khreich. "Gender Bias in Text: Labeled Datasets and Lexicons." *DeepAI*, 21 Jan. 2022, deepai.org/publication/gender-bias-in-text-labeled-datasets-and-lexicons.
- Eitan, Orly, et al. "Is research in social psychology politically biased? systematic empirical tests and a forecasting survey to address the controversy." *Journal of Experimental Social Psychology*, vol. 79, Nov. 2018, pp. 188–199, <https://doi.org/10.1016/j.jesp.2018.06.004>.
- Gelber, Katharine, et al. "Gendered mundanities: Gender bias in student evaluations of teaching in political science." *Australian Journal of Political Science*, vol. 57, no. 2, 27 Feb. 2022, pp. 199–220, <https://doi.org/10.1080/10361146.2022.2043241>.
- Gerber, Alan S., et al. "Testing for publication bias in political science." *Political Analysis*, vol. 9, no. 4, 2001, pp. 385–392, <https://doi.org/10.1093/oxfordjournals.pan.a004877>.
- Honeycutt, Nathan, and Lee Jussim. "A model of political bias in Social Science Research." *Psychological Inquiry*, vol. 31, no. 1, 2 Jan. 2020, pp. 73–85, <https://doi.org/10.1080/1047840x.2020.1722600>.
- Langbert, Mitchell, and Sean Stevens. "Partisan Registration and Contributions of Faculty in Flagship Colleges by Sean Stevens." *NAS*, 17 Jan. 2020, www.nas.org/blogs/article/partisan-registration-and-contributions-of-faculty-in-flagship-colleges.
- Messner, Lindsey. "Why and How to Communicate with Bias-Free, Inclusive Language." *Harris & Associates*, 18 Aug. 2021, www.weareharris.com/resources/blog/why-and-how-to-communicate-with-bias-free-inclusive-language/.

- Microsoft, GenBit. "Responsible-Ai-Toolbox-Genbit/.Github at Main · Microsoft/Responsible-Ai-Toolbox-Genbit." *GitHub*, 2020, github.com/microsoft/responsible-ai-toolbox-genbit/tree/main/.github.
- News, CNN. "2020 Presidential Election Results." *CNN*, Cable News Network, 2020, www.cnn.com/election/2020/results/president.
- Norrander, Barbara, and Clyde Wilcox. "The Gender Gap in Ideology." *University of Arizona*, Springer New York, Dec. 2008, experts.arizona.edu/en/publications/the-gender-gap-in-ideology.
- Nosek, Brian A., et al. "National Differences in Gender–Science Stereotypes Predict National Sex Differences in Science and Math Achievement." *Psychological and Cognitive Sciences*, 30 June 2009, www.pnas.org/doi/abs/10.1073/pnas.0809921106.
- Okruhlik, Kathleen. "Gender and the Biological Sciences1: Canadian Journal of Philosophy Supplementary Volume." *Cambridge Core*, Cambridge University Press, 1 Jan. 2020, www.cambridge.org/core/journals/canadian-journal-of-philosophy-supplementary-volume/article/abs/gender-and-the-biological-sciences1/36BA9B1925E5EA76AD5917D7EEB6A490.
- Parker, Kim. "The Growing Partisan Divide in Views of Higher Education." *Pew Research Center*, Pew Research Center, 19 Aug. 2019, www.pewresearch.org/social-trends/2019/08/19/the-growing-partisan-divide-in-views-of-higher-education-2/.
- Prpic, Katarina. "Gender and Productivity Differentials in Science - Scientometrics." *SpringerLink*, Kluwer Academic Publishers, 29 July 2015, link.springer.com/article/10.1023/A:1016046819457.
- Rekabsaz, Navid, and Markus Schedl. "Do Neural Ranking Models Intensify Gender Bias?" *arXiv*, July 2020, arxiv.org/.
- Review, Princeton. "Most Conservative Students." *Most Conservative College | The Princeton Review*, 2023, www.princetonreview.com/college-rankings?rankings=most-conservative-students.
- Review, Princeton. "Most Liberal Students." *Most Liberal Colleges | The Princeton Review*, 2023, www.princetonreview.com/college-rankings?rankings=most-liberal-students.
- Ringgenberg, Matthew C, et al. "The Politics of Academic Research." *SSRN*, 19 May 2023, papers.ssrn.com/sol3/papers.cfm?abstract_id=4451697.
- Sengupta, Kinshuk, et al. "Genbit: Measure and Mitigate Gender Bias in Language Datasets." *MSJAR*, 2020, www.microsoft.com/en-us/research/uploads/prod/2021/10/MSJAR_Genbit_Final_Version-616fd3a073758.pdf.
- Waites, Matthew. "Critique of 'sexual orientation' and 'gender identity' in human rights discourse: Global queer politics beyond the yogyakarta principles." *Contemporary*