

|CURIEUX|

ACADEMIC JOURNAL

January Issue

Part 1 Issue 46

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: Are Beliefs Voluntary? By Steven Zhen-Ting Li

Page 12: How Do Traditional Dietary Practices, Cultural Perceptions, and the Role of the Familial Traditions Specific to Vietnam Collectively Influence the Nutrition Awareness and Dietary Habits of Children (Under 18) in Vietnam? By Bao Anh Pham

Page 22: Targeting Histone Modifications In Atherosclerosis and Coronary Artery Disease By Saanvi Bohrakoti

Page 33: Through the Use of Nuclear Energy How Can Humans Start an Interstellar Type 3 Civilization? By: Krishaan Somwanshi

Page 62: An Explorative Analysis into the Relationship between Age and Peak Player Performance in the NBA By Johnny Chen

Page 78: Impact of Media on Political Polarization By Tiana Settle-Bowser

Page 86: Using Machine Learning to Predict the Success of Startup Businesses By Ahmad Zayan Sadek

Page 105: Systems of Power: Analyzing Apartheid's Legacy of White Supremacy in South Africa By Saawan Duvvuri

Page 115: Programming Prejudice: Violence Against Women in Media and its Effect on Teenage Attitudes Towards Misogyny By Anita Marie Júlca

Page 150: An Analysis and Discussion of the Various Signal Acquisition Techniques Used in Speech Decoding Technology By Sai Akhil Madurakavi

Page 165: The Application of CRISPR-cas9 Gene Editing Technology for the Synthesization of Hypoallergenic Food Items By Rahul Jagadish Nadig

Page 173: Oral Immunotherapy (OIT): Evaluating An Effective But Unknown Treatment By Anya Jolly

Page 184: An Analysis of the Orthodontic Section of a Dental Clinic in Ho Chi Minh City By Khoi A. Lam, Thuy T. Nguyen, Thu N. Y. Nguyen, Trung V. Lam, Kevin Lam

Page 197: How did the COVID-19 Pandemic Impact the Mental Development of Children Across Different Socioeconomic Backgrounds? By Zhiyuan Liu

Page 204: The Impact of Fluctuating Housing Prices on Socioeconomic Groups in Singapore By Jungwoon Song

Page 210: An Analysis of Housing Affordability in Vancouver By Sophia Huang

Page 222: Energy Justice in Global Climate Change Policy By Saral Gupta

Page 237: Beliefs: a Psychological Analysis By Zhen Xu

Page 241: Revolutionizing Justice by Addressing Racial Bias in Capital Sentencing By Santo Raggiri

Page 256: Associations Between Social Media Use and Adolescent Depression By Danah AlBahar

Page 265: Food Allergies and Evolutionary Mismatch By Naiya Sapra

Page 272: A Descent into the Dark Watery Mud By Jonah Miller

Page 275: Analyzing the Impact of Climate Risk on Inflation: A Comprehensive Statistical Review By Alice Zhang

Page 285: Economic Mobility in Domestic Violence Experienced Families By Max Wu and Andrew Li

Page 298: Marketing for a Cause: An Exploration of How For-Profit and Nonprofit Brands Leverage Social Media to Engage Teenagers By Victoria Oh

Page 305: The Legacy of Chinese Theater: Beijing Opera's Trail Through Macbeth By Mei Teng Shao

Page 311: Unraveling Japanese Encephalitis: Insights into Pathogenesis, Immune Response, and Biomarker Development By Miwa Davis

Page 322: Investments and Maintenance: Their Impact on Perceived Closeness in Adolescent Friendships By Zhen Xu

Page 334: Narratives of Control: How Greek Mythology Reinforces Patriarchal Ideals through Female Archetypes By Suhana Malli

Page 348: Evolving Identities: The Impact of Climate Change on the San and Nubian Peoples By Olumisimi Owolabi

Page 359: A Comparative Analysis of the Economic Relief Efforts from the 1918 Spanish Flu and COVID-19 Pandemics By Hana Sultan, Alena Friedrich

Page 372: The Impact of Social Media on Teen Anxiety By Naiesha Shah

Page 381: Binary Black Hole Merger Models: Common Envelope Theory vs Chemically Homogenous Theory By Sophia Lu

Page 392: Investigation of the Transmembrane Regions of Neurotrophin Receptors By Angela Chou

Page 402: The Neurological Effects Of Music And How It Can Be Used To Improve Patient Well-Being By Narayan Nair

Page 413: Technologies to Combat Ocean Acidification By Diya Sammanna

Page 420: PTSD Symptoms Resulting from Incarceration in America by Siena Coburn By Siena Coburn

Page 431: Polycrystalline Thin Film Technologies for Solar Photovoltaics By Edison Y. Chen

Page 439: The Effect of Gender Stereotypes Found on Zhihu on STEM Career Perceptions for Chinese High School Girls By Zoey Zhi-ning Chen

Page 452: Imperfect Yet Impactful Reforms: How Educational Reformers Successfully Transformed the US Public Education System in the Progressive Era By Jake Kangas

Page 459: The Influence of Emotional Knowledge on Visual Perception: A Comprehensive Review of Facial Expression Interpretation By Chloe Lai

Page 470: The Inequalities of Green Policy By Armaan Sarkhel & Ishan Parikh

Page 486: Examining Japanese Perceptions of Immigrants By Kaho Yamamoto

**Page 497: Exploring the Absence and Double Peaks of Alpha Activity in Resting-State EEG
By George Rohan**

**Page 507: Trastuzumab Deruxtecan vs. Immunotherapy: A Comparative Analysis of
Survival Outcomes in HER2-Positive Breast Cancer By Dylan Gu**

**Page 517: The Effect of Non-governmental Funding on School Dropout Rates in Cambodia
By Jason Tun**

**Page 528: Demographic Influences on U.S. Election Outcomes Across Levels: A Machine
Learning Approach By Benjamin Qiao**

Page 535: The Legalization of Gambling: Does it Really Matter? By Erich Paredes

**Page 559: Livestock Ponds and Their Significance to Avian Communities By Rayyan
Ibrahim**

**Page 572: Design and Analysis of a Modified Knee Brace to Provide Stabilization and
Reduce Inflammation from Patellar Tendonitis in Volleyball Players By Boston Richardson**

**Page 580: A Thorough Analysis of Machine Learning Models For Sports Results Prediction
By Adit Singh**

Page 589: Proof of Zsigmondy's Theorem By Nayer Sharar

**Page 601: Educational Disparities in Urban and Rural India and Its Impact on Economic
Inequality By Akaanksha Chintakindi**

**Page 618: The Impact of Brand Rivalries on Consumer Behavior in the Indian Smartphone
Market By Swakshi Jain**

Are Beliefs Voluntary? By Steven Zhen-Ting Li

Introduction

Beliefs serve as the foundational building blocks upon which our understanding of the world is constructed. In essence, beliefs shape our identity as human beings, being defined as an “idea that we hold as being correct, an acceptance that something exists or is true” (Gurteen, 2009). However, this raises the question: do we truly possess autonomy over the formation of our beliefs? If so, is it possible for us to adopt a belief based solely on conscious choice, or in other words, voluntarily?

In this essay, I will contend that the beliefs human beings hold are voluntary to an extent, as they follow a nuanced trajectory throughout human lives, influenced by various power dynamics. During early stages of life or when individuals occupy subordinate positions of power, the volitional aspect of belief formation is significantly diminished, with external factors, such as socialization and authority, playing a significant role in shaping beliefs during these formative periods. However, as individuals mature and gain the capacity and resources to assert independence over their beliefs and moral systems, the potential for conscious choice emerges.

Throughout our lives, power relations can engender conditions that limit the choices of beliefs available to individuals. The social structures stemming from these skewed dynamics such as economic inequality or hierarchies of knowledge restrict access to the diverse perspectives, information, and resources necessary to critically evaluate and construct beliefs. As such, individuals may find themselves constrained within the boundaries of societal norms and what is endorsed by those in positions of power.

Although this essay will argue in favor of the progressively voluntaristic formation of beliefs with age, it must be acknowledged that complete voluntarism remains elusive. While individuals possess the ability to weigh evidence gathered and make autonomous decisions, a variety of uncontrollable hidden factors lurk beneath the surface, compromising the extent to which the adoption of beliefs is entirely voluntary. One such factor is confirmation bias, which inherently predisposes individuals to favor beliefs and courses of action that aligns with their preconceived opinion (Nikolopoulou, 2022) Thus, even when engaging in a voluntary exploration of beliefs, many people cannot fully extricate themselves from the sway of subconscious biases. However, appropriate educational interventions and provisions of resources can do a great deal to increase the extent to which individuals overcome the limitations imposed upon them by external factors or pre-existing beliefs.

Childhood

During infancy, the cognitive capabilities necessary for intricate belief formation are still nascent, limiting our comprehension of rudimentary sensory perceptions and preventing us from engaging in reasoning, abstraction, or critical thinking. This concept of "tabula rasa", or a, “supposed condition that empiricists have attributed to the human mind before ideas have been imprinted on it by the reaction of the senses to the external world of objects” (Britannica, 1998),

conceptualizes newborns as having minds comparable to blank canvases. In this early phase, children rely extensively on external influences to formulate their understanding of the world. Parents, caregivers, and the immediate environment assume crucial roles in this developmental process. Through keen observation and emulation, children gradually internalize the behaviors, attitudes, and values exhibited by those around them and so as they mature, their beliefs become increasingly shaped by the prevailing social and cultural norms practised by their community.

The early stages of socialization and education exert a profound influence on the development of a child's beliefs about and outlooks on the world, given their limited cognitive abilities to form independent opinions. A notable demonstration of this can be observed in American schools, in which the practice of reciting the Pledge of Allegiance every morning serves as a vehicle for instilling patriotism; many children become accustomed to this ritual and engage in it without much thought. However, as children age, they begin to question the reason for reciting the Pledge, while some refuse to do so, due to the fear of condemnation by their classmates and even instructors (Etzler, 2016). This decision to discontinue the recital of the Pledge reflects an individual's evolving beliefs and values, as they learn to critically evaluate the notions that were previously instilled in them involuntarily. Similarly, few children actively select which religion to follow; in fact, religion is frequently imposed upon the members of the younger generations as a way to perpetuate religious practice and traditions.

These early belief formations can be further attributed to the hierarchical nature of children's relationships with their elders or caretakers. In the context of the family unit, parents or caregivers typically occupy the positions of power, assuming care-taking responsibilities, and children's survival depends entirely on these parental figures to take care of them. This hierarchical dynamic generates a power imbalance whereby children are obligated to comply with the beliefs, values, and rules established by their elders.

Children are also embedded in broader social communities such as schools, religious institutions, and cultural groups, in which teachers, religious leaders, or community elders assume positions of authority and influence. On the other hand, the children, being in a position of relative powerlessness, tend to obey the authority of their elders due to the fear of actual or perceived punishment. Children seek acceptance, validation, and a sense of belonging from and within their families and communities (Young, 2021); thus as a result, they are more inclined to adopt the beliefs and values upheld by those in power as a means of conforming and fitting into their social environment.

Extending upon this, the power held by adults in these hierarchical relationships can also manifest in the form of social rewards and punishments. Children receive praise, approval, or materialistic rewards when they accord themselves with the beliefs and behaviors endorsed by their elders. Conversely, they receive criticism, disapproval, or reprimands when their beliefs deviate from those of their caregivers or community leaders. Such external factors are what further strengthen the influence of power dynamics on children's internalization of beliefs.

It is important to note that power dynamics within these hierarchical relationships are not to be seen as purely negative or coercive. Adults, as figures of authority, have the responsibility to guide and protect children, imparting the knowledge and skills essential in the service of healthy development. The transmission of beliefs within such hierarchical relationships is not necessarily an act of manipulation, but rather a means of socializing children into their culture, providing them with a framework of understanding the world to assist their integration into society.

Adulthood

As children enter adulthood, the extent to which individuals can voluntarily adopt beliefs increases. With the maturation of cognitive abilities, individuals gain the capacity to engage in independent introspection and critical analysis of their own belief systems and moral compass. This is evidenced through Jean Piaget's theory of cognitive development, which theorizes that children pass through distinct stages, acquiring new cognitive abilities and ways of comprehending the world (Main, 2022), where, as they transition from concrete to abstract thinking, the resolution of their beliefs become more intricate and nuanced. Similarly, Lawrence Kohlberg's theory of moral development posits that individuals progress through stages of moral reasoning, beginning with a focus on obedience and gradually incorporating principles of justice and ethical considerations (McLeod, 2008). This progression suggests that our beliefs about moral dilemmas evolve as we mature and become capable of engaging in more complex reasoning.

As we age, not only do our cognitive capabilities increase, but so does our accumulation of both tangible and intangible resources. The acquisition of physical assets like wealth and /or non-physical assets like experiences and skills further enables individuals to assert their autonomy over shaping their own belief systems. This is particularly the case as access to a diverse range of resources, including educational opportunities and new information, empowers individuals with a broader worldview and thus the privilege to exercise educated judgements in determining their beliefs. For example, while children in religious households often adopt their parents' religion initially, a notable number of them gradually transition to agnosticism or atheism as they mature. In the United States, of the young American adults who actively participated in Protestant churches for at least a year during their teenage years, 66% reported a subsequent period of disengagement lasting at least a year between ages 18 and 22 (Earls, 2021). The cognitive transformations that come with age are highlighted here, as individuals pursue education, gaining access to various sources of knowledge, and adopt a more open-minded attitude. It reinforces the notion that as individuals expand their understanding of the world and encounter diverse perspectives, they may reassess their convictions and adopt alternative philosophical or even non-religious worldviews.

As such, it is evident that children face inherent challenges when it comes to the formation and assertion of their beliefs, which are at this stage completely involuntary, mainly due to factors such as their basic cognitive development and limited autonomy. With their cognitive abilities still undergoing development, children possess neither the necessary knowledge nor experience to critically assess and evaluate complex belief systems. As we

mature into adulthood, however, humans develop volition, gaining the capacity to seek out diverse sources of information, engage in critical thinking, and explore different belief systems. This accumulation of knowledge and resources enables adults to weigh perceived evidence and make informed decisions about their values.

Power

While it is certainly the case that adults have more volition over their beliefs than children, there are nevertheless limits on the ability of even adults to form and express their own views as a result of power dynamics. French philosopher Michel Foucault argued that power operates through social institutions and discourses to shape and regulate individuals beliefs (Gaventa, 2003). Foucault emphasizes the role of power-knowledge, proposing that knowledge and power are inseparable (Fruhling, 2022). Power structures shape what is deemed legitimate knowledge, while knowledge, in turn, reinforces and justifies existing power structures. This interplay between power and knowledge perpetuates dominant belief systems and limits alternative perspectives from emerging. Consider the highly restrictive context of North Korea, for instance, where a tightly controlled political system and pervasive ideological indoctrination severely limit individual freedoms to shape the social landscape. Despite potential advantages such as age or resources, the overarching dictatorial political structure impedes the formulation and expression of independent beliefs. The regime's stranglehold on information, freedom of speech, and constant surveillance curtail opportunities for dissent and open expression of diverse viewpoints and discourses.

The social and political context in North Korea promotes a collectivist mindset that prioritizes conformity and loyalty to the ruling regime. Dissent or expression of independent beliefs is viewed as a threat to the stability and power of the regime, making it virtually impossible for individuals to openly assert their own convictions. Under such power imbalances, even adults find themselves unable to exercise autonomy in shaping their beliefs.

Conclusion

Ultimately, the formation of beliefs is a complex, multi-faceted process which is involuntary throughout childhood, due to cognitive limitations and social dynamics. The formation of beliefs, however, becomes progressively more voluntary with age as we possess greater volition to critically evaluate and shape our own beliefs. However, it is essential to recognize that complete autonomy in belief formation may be impossible due to the pervasiveness of power dynamics. In understanding the complexities of the voluntary or involuntary nature of beliefs and limitations imposed by external factors, our understanding of the nuanced nature of human belief formation is heightened.

Works Cited

- Audi, R. (2015). Doxastic voluntarism and the ethics of belief. *Rational Belief*, 27-46.
<https://doi.org/10.1093/acprof:oso/9780190221843.003.0003>
- Becchio, C., Manera, V., Sartori, L., Cavallo, A., & Castiello, U. (2012). Grasping intentions: From thought experiments to empirical evidence. *Frontiers in Human Neuroscience*, 6.
<https://doi.org/10.3389/fnhum.2012.00117>
- Bennett, O. (1990). Why is belief involuntary? *Analysis*, 50(2), 87-107.
<https://doi.org/10.1093/analys/50.2.87>
- Booth, A. R. (2015). Belief is contingently involuntary. *Ratio*, 30(2), 107-121.
<https://doi.org/10.1111/rati.12126>
- Britannica. (1998, July 20). Tabula rasa | Definition, history, & facts. *Encyclopedia Britannica*.
[https://www.britannica.com/topic/tabula-rasa#:~:text=tabula%20rasa%2C%20\(Latin%3A%20%E2%80%9C,John%20Locke](https://www.britannica.com/topic/tabula-rasa#:~:text=tabula%20rasa%2C%20(Latin%3A%20%E2%80%9C,John%20Locke)
- Clarke, M. (1986). Doxastic voluntarism and forced belief. *Philosophical Studies*, 50(1), 39-51.
<https://doi.org/10.1007/bf00355159>
- Cusimano, C., & Goodwin, G. P. (2020). People judge others to have more voluntary control over beliefs than they themselves do. *Journal of Personality and Social Psychology*, 119(5), 999-1029. <https://doi.org/10.1037/pspa0000198>
- Earls, A. (2021, August 11). Most teenagers drop out of church when they become young adults. *Lifeway Research*.
<https://research.lifeway.com/2019/01/15/most-teenagers-drop-out-of-church-as-young-ults/>
- Etzler, S. (2016). Why I don't stand for the pledge of allegiance. *The Lance*.
<https://lhslance.org/2016/opinion/why-i-dont-stand-for-the-pledge-of-allegiance/>
- Fruhling, Z. (2022). Foucault's concept of "Power/Knowledge" Explained — Zachary Fruhling. Zachary Fruhling.
<https://www.zacharyfruhling.com/philosophy-blog/foucaults-concept-of-power-knowledge-explained>
- Gaventa, J. (2003). Foucault: Power is everywhere. Understanding power for social change | the powercube | IDS, Sussex University.
<https://www.powercube.net/other-forms-of-power/foucault-power-is-everywhere/>
- Gurteen, D. (2009). What are beliefs? *Conversational Leadership*.
<https://conversational-leadership.net/beliefs/>
- Ionescu, I. (2022, March 11). How our beliefs and values shape our behavior. Iulian Ionescu.
<https://iulianionescu.com/blog/how-our-beliefs-and-values-shape-our-behavior/>
- Hansson, S. O. (2006). Logic of belief revision (Stanford encyclopedia of philosophy). *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/logic-belief-revision/>
- Main, P. (2022, December 16). Jean Piagets theory of cognitive development. *Structural Learning | Advancing Metacognitive Classroom Practice*.
<https://www.structural-learning.com/post/jean-piagets-theory-of-cognitive-development-a>

- nd-active-classrooms#:~:text=The%20Theory%20of%20Cognitive%20Development,world%20(Miller%2C%202011)
- Mcleod, S. (2008, February 5). Kohlberg's stages of moral development. Study Guides for Psychology Students - Simply Psychology.
<https://www.simplypsychology.org/kohlberg.html>
- Naylor, M. B. (1985). Voluntary belief. *Philosophy and Phenomenological Research*, 45(3), 427.
<https://doi.org/10.2307/2107274>
- Nickel, P. J. (2010). Voluntary belief on a reasonable basis. *Philosophy and Phenomenological Research*, 81(2), 312-334. <https://doi.org/10.1111/j.1933-1592.2010.00380.x>
- Nikolopoulou, K. (2022, November 18). What is confirmation bias? | Definition & examples. Scribbr. <https://www.scribbr.com/research-bias/confirmation-bias/>
- Waters, S. (2021, December 10). Family dynamics can lift you up (or drag you down). BetterUp: The People Experience Platform. <https://www.betterup.com/blog/family-dynamics>
- Wyers, R. S., & Albarracín, D. (2014). Belief formation, organization, and change: Cognitive and motivational influences. *The Handbook of Attitudes*, 287-336.
<https://doi.org/10.4324/9781410612823-15>
- Young, A. (2021). What is validation and why is it so crucial for our kids?. Results that make a difference.
<https://southshoetherapies.com/b/validation-why-is-it-so-crucial-for-our-kids#:~:text=Validation%3A%20Why%20Is%20It%20So,%2C%20emotional%20intelligence%2C%20and%20growth.>

How Do Traditional Dietary Practices, Cultural Perceptions, and the Role of the Familial Traditions Specific to Vietnam Collectively Influence the Nutrition Awareness and Dietary Habits of Children (Under 18) in Vietnam? By Bao Anh Pham

Abstract

Dietary practices and traditions have been a significant part of Vietnamese culture. This literature review sought to examine how traditional dietary practices, cultural perceptions, and familial traditions can influence nutrition awareness and dietary habits in Vietnamese children under the age of 18. The results of this literature review revealed that these components significantly contributed to the nutritional development of children in Vietnam. Traditional dietary practices promoted the consumption of certain foods over others, like a preference for rice over dairy, leaving kids lacking essential nutrients. Cultural beliefs in Vietnam also played a significant role in childhood nutritional awareness. Education is highly valued in Vietnam, so parents often prioritize academics over health, making these children more susceptible to a sedentary lifestyle and obesity. In light of this, it would be beneficial for the Vietnamese government to educate students about the importance of making healthy food and lifestyle choices. Furthermore, increased research focused on spreading nutritional awareness in Vietnam could have a great impact on the dietary choices of parents and children, alike.

Introduction

Vietnam has a diverse food culture, encompassing many traditional dietary practices, cultural perceptions, and familial etiquette. This literature review will explore how these dynamics develop within Vietnamese food culture and potentially shape Vietnamese children's nutrition awareness and eating habits.

The Dietary Guidelines Advisory Committee claims that dietary practices revolve around specific eating habits and food choices (Dietary Guidelines Advisory Committee 2020). Cultural dietary practices and traditions do not always align with good nutritional choices, which can lead to poor nutrition in populations that heavily depend on cultural perceptions to make their food decisions.

In addition, not only do traditional dietary practices affect the nutritional development of children, but cultural perceptions are also considered a crucial impact on shaping their eating habits. Cultural perceptions often refer to the beliefs regarding traditional aspects like the food culture. In Vietnamese food culture, these beliefs encompass foods' symbolisms, cultural taboos, or conventional usage of specific ingredients. As children adapt to these influential cultural aspects in their households, it is natural for them to be familiar with these rich cultural values, and gradually develop certain nutritional habits.

Moreover, familial traditions are a considerable factor in impacting the dietary habits of Vietnamese children. Familial customs discuss the principles associated with the food culture that remain from one generation to the next. For instance, the customs of family meals or the cultural importance of food-related events may shape children to have certain habits and gain

more knowledge of the customs and foods of their own country, leading them to lean towards a more specific eating culture.

Hence, the purpose of this literature review is to summarize how traditional dietary practices, cultural perceptions, and family traditions affect children's nutritional awareness and dietary habits. As we delve into this review, we will also evaluate the significance and the consequences of the impact on shaping Vietnamese children's understanding of nutrition.

Traditional Dietary Practices

Traditional dietary practices in Vietnam display Vietnamese-specific eating habits and food choices by how people choose, prepare, and consume their food. As Vietnamese children are exposed to these practices from an early age and surrounded by the Vietnamese culture, they will learn and absorb the nutritional aspects of their culture and the development of certain eating habits.

Raneri et. al. claims that white rice stands out as the major staple food consumed by nearly everyone daily (Raneri et al., 2019). Additionally, the website "Vietnam Online" portrayed how rice plays a pivotal part in Vietnamese food culture because rice is widely planted across most Asian countries, indicating that rice cultivation and consumption are popular in the region (Vietnam Online). This dietary pattern is deeply involved with Vietnamese culture, having a profound impact on the choices made by Vietnamese children surrounding their carb intake. Chaudhari et. al. proves that white rice is a perfect stable carb source as it is a complex carbohydrate, which must be broken down into glucose before the body can use it (Chaudhari et al., 2018). Furthermore, rice also has other elements like tocotrienol which can prevent or reverse blood clots and lesions that may lead to strokes and thrombosis, and manganese which aids in the normal functioning of the brain (Chaudhari et al., 2018). However, Krittanawong et. al. proclaims that white rice consumption is positively correlated with an increased risk of metabolic syndrome (MetS) (Krittanawong et al., 2017). Operating within this rice culture, Vietnamese children have sufficient daily carbohydrate intake that can foster growth and help prepare them to make healthy decisions. However, according to Phan et. al., rice can pose potential health threats for Vietnamese children (Phan et al., 2023). According to a research study that evaluated the health risks of mycotoxin contamination (particularly aflatoxins (AFs) and fumonisins (FBs)) caused by rice consumption in the Mekong Delta, Vietnam, children are one of the groups that have the highest exposure to AFs and FBs (Phan et al., 2023). The study attributed this to children's increased consumption of rice relative to their body weight, with children consuming 7.00g/kg bw/ day compared to adults' rice consumption which only reaches 6.43g/ kg bw/ day (Phan et al., 2023). Mycotoxin production is created by certain fungi that are prominently fostered in areas within the Mekong Delta due to the ideal environment with heavy frequent rainfalls during crop seasons (Phan et al., 2023). Rice can be contaminated with these intoxications due to the high humidity, which fosters fungal growth, poor crop management, and inefficient drying methods that support fungal growth and contamination of rice during transportation (Phan et al., 2023). According to the study, AF exposure in children can lead to

significant health impacts like an increased risk of liver cancer and malignant (necrotizing) otitis externa (MoE) (Phan et al., 2023). Therefore, while rice significantly contributes to Vietnamese children's necessary dietary needs and provides these children with numerous nutritious components, we also need to acknowledge the health threats of rice due to the lack of safety during production and transport. Unsanitary conditions in rice production lead to the immense growth of fungus which potentially spreads diseases among children and other populations, posing serious health risks.

However, in contrast to the popular consumption of the preferred staple like rice, the consumption rate of milk and dairy among the Vietnamese is comparatively low due to cost and food safety concerns (Raneri et al., 2019). This can be attributed to the high cost of dairy products, which serve as barriers for many Vietnamese households (Raneri et al., 2019). Reduced access to dairy can indirectly influence children's development and future dietary choices, as a lack of access to dairy during childhood limits exposure to this food source, reducing the chance that a child would choose to consume dairy in the future. An additional factor that contributes to the low consumption of dairy in Vietnam is that milk and dairy products are not a traditional part of the Vietnamese diet, leading to insufficient calcium levels in children's nutrition intakes. According to the study, the national survey showed that the dairy consumption of an average Vietnamese child aged 2 to 5 was 138 grams/day – only 1/3 of the national recommendation (Raneri et al., 2019). Furthermore, calcium intake in Vietnamese adults was 500- 540 mg/per/day, which only accounts for 50% to 60% of the WHO dietary recommendation (Raneri et al., 2019). Dairy is important during childhood development, as higher dairy intake is correlated with higher absorbance of nutrients, like vitamin B2, calcium, vitamin D, fat, vitamin B1, protein, and vitamin A (Khanh Le et al., 2016). Therefore, dairy products like milk could add nutrition and boost children's absorption of vitamins found in other foods. However, according to the study recorded by Khanh Le et. al., a rice-heavy diet often leaves young children with a dearth of nutritional components like iron, vitamin A, vitamin B1, vitamin C, energy, and protein (Khanh Le et al., 2016). The research concludes that low intake of dairy can be explained by a lack of awareness of the full benefits of dairy and the dietary patterns involved within Vietnamese culture. Hence, it is recommended that education programs be provided to raise awareness about the benefits of all nutrition, particularly dairy, as dairy provides numerous benefits to children's health.

Dietary patterns also involve the development of dietary habits among Vietnamese adolescents. Mai et. al. provide the analyses of dietary patterns: three different patterns applied to the tested cohort of Vietnamese adolescents which are traditional, discretionary, industrialized, and categorized based on certain types of food groups (Mai et al., 2023). The research was conducted on a group of Vietnamese children 9-11 years old in Ho Chi Minh City. The traditional dietary pattern included minimally processed food such as whole grains, fresh vegetables, and meat sources (Mai et al., 2023). The discretionary dietary pattern encapsulates food groups with a high content of sugar, fat products, or foods with low nutritional value like sweets and sweetened beverages (Mai et al., 2023). The industrialized dietary pattern includes

processed foods like fast food and processed meat (Mai et al., 2023). The study found that traditional dietary pattern was associated with obesity, discretionary dietary patterns correlated with being overweight, and industrialized dietary patterns led to elevated waist-to-height ratio (WHR) (Mai et al., 2023). According to the National Heart, Lung, and Blood Institute, overweight is a BMI between the 85th percentile and the 95th percentile while obesity is defined as a BMI in the 95th percentile and above (National Heart, Lung and Blood Institute 2022). The WHR is calculated by dividing the waist circumference by height to determine individuals with increased cardiometabolic risk (Yoo 2016). WHR of 0.5 or higher most often represents short people or those with a higher prevalence of cardiometabolic risk factors, meanwhile WHR of 0.5 or lower would be more representative of taller people or those with lower cardiometabolic risk factors (Yoo 2016). Mai et. al. concludes that these dietary patterns do have a positive correlation with the physical status of children's well being like obesity, overweight, and increased WHR. According to Mai et. al, the discretionary food pattern is positively correlated with an increase in the rate of being overweight since food items like sweetened beverages are more easily accessed by the majority of the population with low-cost markets (Mai et. al., 2023). Due to the low costs of these foods, children may be more likely to access them with allowances from their parents because parents tend to work full-time, leaving children to make poor dietary choices due to lack of supervision over their nutrient intake. Hence, it is crucial to encourage balanced nutrition for children's development and healthier eating habits to prevent health risks.

Dietary patterns significantly influence the development of nutritional habits. Analysis of the data generated by Mai et. al. shows that dietary patterns affect the health status of Vietnamese children. Certain food choices within this culture shape the nutritional choices of Vietnamese children. These specific food choices can lead to a lack of adequate nutritional intake, hindering children's development and posing additional health risks. Therefore, it is necessary to provide certain educational programs in schools to instruct students on how to balance their meals with all the necessary nutrients. Not only that, it is essential for governments to implement greater caution in food production to prevent the spread of diseases and ensure that food products contain the most nutritious components for optimal growth of the children population.

Cultural perceptions

Vietnamese food's cultural perceptions embrace various qualities like specific food restrictions or table etiquette. These deep values attached to cultural perceptions can guide the development of children's dietary habits. Essentially, cultural perceptions have a profound impact on filtering children's understanding of food culture, impacting their overall nutritional well-being.

Mai et. al. showed that besides the independent factors of dietary patterns that affect overweight and obesity rates in Vietnamese youth, there are also other factors contributing to these diseases like gender and parents' under-estimated perception of a child's weight status (Mai et al., 2023). According to the research paper, one of the most common factors impacting obesity, overweight, and elevated WHR is gender perceptions in Vietnam. Vietnamese boys are

encouraged to be big and strong, which can influence them to make dietary choices that support their weight gain, ultimately helping them to better fit with societal expectations. It is clear that cultural norms contribute to the rising prevalence of obesity between sexes (Sano et al., 2008). In support of this, a study by Sano et al., 2008 saw that when Vietnamese boys and Japanese boys were compared based on perceptions of their body images, Vietnamese boys tended to worry about their body image more and wanted a bigger and more muscular body compared to Japanese boys (Sano et al., 2008). Additionally, according to Sano et. al., parental perception is also identified as a contributing factor to childhood obesity as it showed around 30% to 90% of Asian parents tend to underestimate their children's weight due to constant inadequate food access (Sano et al., 2008). Many Asian parents correlate a child's weight with wealth status, so in their eyes, an overweight child is a wealthy one. Mai et. al. found parents who were unable to accurately categorize the weight status of their own children were statistically more likely to have children that were overweight and obese (Mai et al., 2023). They concluded that parental underestimation of a child's weight status contributed to rising rates of obesity, overweight, and elevated WHR (Mai et al., 2023). Mai et. al. highlights a high proportion (>75%) of weight status under-estimation by parents was found among children who were either overweight or obese (Mai et al., 2023). This underestimation is likely caused by the rapid nutrition transition that has occurred in Vietnam over the past decade. Previously, the majority of Vietnamese children were underweight, which impacted what was perceived as a healthy child. When undernutrition is endemic to a region, an overweight child is viewed as healthy and well-nourished. Furthermore, in Vietnam, extended family, such as grandparents, take care of the children in addition to their parents, which leads one to consider extended family member's perception of the child's weight status as well. Hence, it is significant to adjust the perspective of children's weight status to prevent the high possibility of obesity, overweight, and elevated WHR in the future as well as secure their health to be in a well-taken care mode.

Academic pressure may also contribute to high obesity rates in Vietnamese children. In the majority of Vietnamese households, parents have certain academic expectations of their children, so children mostly spend their time learning, both in school as well as by taking outside classes in order to compete in competitive academic environments (Mai et al., 2023). This results in limited time for other activities like exercising, leading to a decrease in physical activity among Vietnamese children. Additionally, due to the stress from academic expectations, some children might cope with their stress by eating energy-dense, poorly nutritious food as a form of emotional support leading to the excessive intake of calories, resulting in weight gain and obesity. Unfortunately, certain cultural perceptions and expectations can have an indirect, negative influence on children's dietary intake.

It is also important to consider a family's socioeconomic status when discussing cultural perceptions. Mai et. al. separated families into high-income families with parents who had higher education and low-income families with parents who did not have the opportunity to achieve higher education to analyze how socioeconomic and education status impacted childhood obesity. According to the research paper, parents with higher education and higher income

tended to have higher self-efficacy to help children foster healthier food choices and more often guided their children to perform outdoor physical activities (Mai et al., 2023). Families with higher incomes had a higher capability to support their children with after-school outdoor programs and provided healthier and more organic food to keep children healthy and fit. However, parents with lower education were more prone to hold fast to traditional cultural norms and often viewed a chubby child as a symbol of great health (Mai et al., 2023). This resulted in a lower self-efficacy to guide children to make healthier choices and increased the risk of developing obese and overweight children. Overall, they concluded that large gaps in socioeconomic groups could contribute to parental perceptions that impact children's development, where wealthier families with higher education can provide more balanced nutrition for their children, while poorer families with parents who did not have access to higher education might not have the resources to guide children into making healthy eating choices.

Cultural perceptions are a crucial element influencing children's dietary choices in Vietnam. Since Mai et. al. found a positive correlation between parental underestimation of child weight and BMI, it will be important to adjust this perspective in Vietnamese adults so that parents can help their children to make better dietary choices. It is clear that in the rich Vietnamese culture, dietary preferences are more prone to be influenced by cultural norms. Having a deep understanding of these perceptions can enhance our understanding of the culture and can help predict dietary preferences and possibly even health status among different populations of Vietnamese children.

Role of Familial Customs

The role of familial customs in Vietnamese culture influences Vietnamese children's interaction with food, which influences their habits within the food culture. These familial traditions involve things like family structure and table etiquette. These components can have a consequential impact on the development of children's relationships with nutritional awareness.

Mai and Le et. al. assert that filial piety represents the respect, love, and care that younger generations provide for their elders, carrying out their responsibilities towards those who raised them (Mai and Le et. al., 2023). The idea of filial piety is ingrained in Vietnamese culture as the recorded data by Mai and Le et. al. shows a positive correlation between family happiness and filial piety in Vietnam (Mai and Le et. al., 2023). This means filial piety is essential in maintaining the family structure, ensuring not only order but also the overall happiness within families. This value is widely promoted throughout Vietnamese culture and one expression of filial piety can be shown through table etiquette at family meals.

According to a travel blog by Chi Nguyen from Paradise Travel, one of the Vietnamese table manners is to always wait for the oldest person to sit down and start the meal first, as in Vietnamese customs, it is important to pay respect to elders (Chi Nguyen 2020). Consequently, the youngest ones in the families would be the last ones to eat (Chi Nguyen 2020). Not only that, it shows that Vietnamese people are taught to show love and respect for others starting from a young age, so they often choose the best dishes for their elders like grandparents or their parents

(Chi Nguyen 2020). This widespread practice leads to elders typically receiving the best portions of the meal, with first access to the most nutritious food, while young children do not receive the best portions, potentially resulting in a lower intake of nutrients than adults. Furthermore, according to an article on Travel Sense Asia, it is disrespectful not to finish the food that your elders or your hosts prepared (Travel Sense Asia 2022). Therefore, it is common for young Vietnamese children to finish their food to show that they are appreciative of the prepared food and respect the hard work of their elders. If the children are malnourished, having to finish everything on their plate can help their bodies absorb the essential nourishment they lack, improving both their health and body weight. On the other hand, this practice can have negative impacts on overweight and obese children's health, as it could encourage overeating. In cases of nutritional excess, food is stored as fat, which exacerbates obesity. Therefore, it is critical to balance meals for each type of body because different body types accommodate different nutrient intakes.

Chopsticks are the primary utensil used in Vietnam. Vietnamese dishes often include rice with side dishes like vegetables, fish, etc. and these dishes are often cut into bite-sized pieces, so that family members can pick up their food from shared dishes and put them into personal food bowls, paying respect to everyone. This culture emphasizes the importance of everybody having an equal amount of food, ensuring they do not take too much or have too little. This tradition highlights mindfulness and respect for everyone in the family, which indicates the family's qualities in Vietnamese tradition. According to Thuy Nguyen et. al., chopsticks have many functions and replace the use of many tools to accommodate Vietnamese customs as they can pick up food, slice fish, or slice while eating (Thuy Nguyen 2020). Thuy Nguyen et. al. states that chopsticks are significant to Vietnamese tradition as they represent love and unity, because chopsticks always exist in pairs, signifying husbands and wives, as well as solidarity (Thuy Nguyen 2020).

Chopsticks are the prominent utensils used in Vietnamese households. When using chopsticks, Vietnamese children learn how to pick food up gently and carefully, and how to not drop food on the table, as this is how people can judge your personality during eating customs. Therefore, it's harder to pick up food quickly which makes it harder for people to eat freely, influencing Vietnamese children to have more controlled, moderate eating habits and a lower likelihood of overeating. This reduced probability of overeating can prevent rapid weight gain. From this information, it is clear that familial customs are a contributing factor to the development of certain eating etiquette in Vietnamese children. Based on these studies, Vietnamese familial practices like filial piety and the use of chopsticks affect how Vietnamese children treat food as they grow up. While the act of showing respect by waiting for elders to receive the best cuts of food may limit Vietnamese children's access to the healthiest portions, the use of chopsticks encourages a controlled eating process and the culture of sharing ensures that these children can still receive enough nourishment alongside their elders. Therefore, it seems that familial customs can have negative and positive impacts on children's nutrition.

Discussion

In conclusion, traditional dietary practices, cultural perceptions and the role of familial traditions in Vietnam collectively influence the nutritional awareness and dietary practices of children. This literature review identified that traditional dietary practices influence Vietnamese children's food choices, like eating rice as a primary source of carbohydrates while disregarding the importance of milk consumption, leaving children to lack the sufficient energy required to reach their growth potential. Additionally, Vietnamese cultural perceptions impact these choices. For example, the perception that a chubby person is successful or wealthy has a negative impact on childhood nutrition. Additionally, parental prioritization of children's academic performance over physical activity leads to a lack of exercise and an increased prevalence of being obese or overweight. Lastly, familial customs are also shown to have a vital influence on dietary habits and nutritional awareness. For example, filial piety can prevent children from consuming the best cuts of meat, which can hinder efficient nutrition intake. However, the Vietnamese cultural practice of using chopsticks helps children to practice moderation and reduces the likelihood of overeating. Therefore, familial customs can positively and negatively impact childhood nutrition. It is clear that many factors influence Vietnamese children's nutritional awareness and dietary practices. To improve nutritional awareness in this population, children should be taught how to balance their meals with necessary nutrition through the implementation of an educational program. Governments also need to take serious measures while producing food to stop spreading diseases that harm children's health. In addition, it is crucial to increase parental awareness of nutrition and the prevalence of obesity in Vietnamese adolescents, as parents play a major role in developing children's eating habits. This can be achieved by implementing nutritional guidelines and tips that are shared through social media and national media to provide parents with the essential knowledge to care for their children in the best possible way. However, throughout this research, the lack of resources regarding this topic limits the analysis's potential. As a result, it is advised that more research be done on these nutrition-related scopes, particularly in developing countries, to disseminate vital health information and guarantee that children raised in these areas can grow healthily and equally like other kids.

How do traditional dietary practices, cultural perceptions, and the role of the familial traditions specific to Vietnam collectively influence the nutrition awareness and dietary habits of children (under 18) in Vietnam?

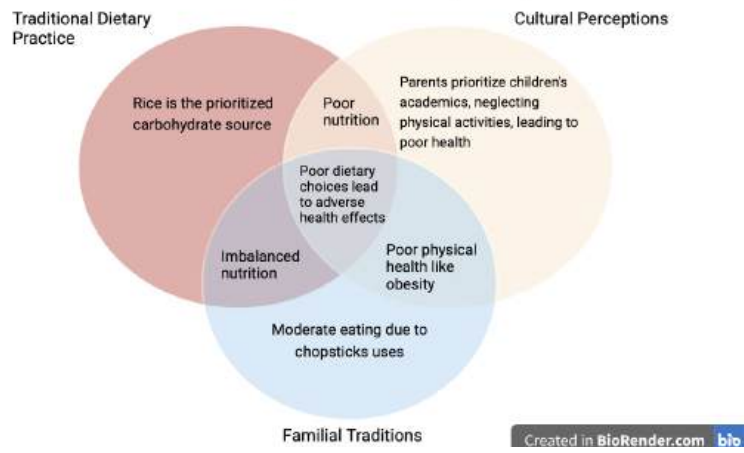


Figure 1: Impact of how traditional dietary practices, cultural perceptions and familial traditions in Vietnam collectively influence the nutrition awareness and dietary habits of children (under 18) in Vietnam

Works Cited

- Assessing Your Weight and Health Risk.
www.nhlbi.nih.gov/health/educational/lose_wt/risk.htm.
- Khanh, Le Nguyen Bao, et al. "Nutrient Intake in Vietnamese Preschool and School-Aged Children is Not Adequate." *Food and Nutrition Bulletin*, vol. 37, no. 1, Mar. 2016, pp. 100–11. <https://doi.org/10.1177/0379572116631642>.
- Krittanawong, Chayakrit, et al. "Is white rice consumption a risk for metabolic and cardiovascular outcomes? A systematic review and meta-analysis." *Heart Asia*, vol. 9, no. 2, Aug. 2017, p. e010909. <https://doi.org/10.1136/heartasia-2017-010909>.
- Mai, Thi My Thien, et al. "Dietary patterns and child, parental, and societal factors associated with being overweight and obesity in Vietnamese children living in Ho Chi Minh city." *Maternal and Child Nutrition*, vol. 20, no. S2, Apr. 2023, <https://doi.org/10.1111/mcn.13514>.
- Nguyen, Thuy Thi Thu. VIETNAMESE CULTURE IN CHOPSTICK USE COMPARE TO WESTERN CULTURE IN UTENSIL USE. no. 2, Jan. 2020, <https://doi.org/10.34768/rl.2020.v462.07>.
- Phan, Lien Thi Kim, et al. "Public health risk due to aflatoxin and fumonisin contamination in rice in the Mekong Delta, Vietnam." *Food Safety and Risk*, vol. 10, no. 1, Mar. 2023, <https://doi.org/10.1186/s40550-023-00104-0>.
- Sano, Ayami, et al. "Study on Factors of Body Image in Japanese and Vietnamese Adolescents." *Journal of Nutritional Science and Vitaminology*, vol. 54, no. 2, Jan. 2008, pp. 169–75. <https://doi.org/10.3177/jnsv.54.169>.
- Van Mai, Hai, and Hao Van Le. "The interdependence of happiness and filial piety within the family: a study in Vietnam." *Health Psychology Report*, Oct. 2023, <https://doi.org/10.5114/hpr/172091>.
- Vietnamese Culture: Rice. www.vietnamonline.com/culture/rice.html.
- Yoo, Eun-Gyong. "Waist-to-height ratio as a screening tool for obesity and cardiometabolic risk." *Korean Journal of Pediatrics*, vol. 59, no. 11, Jan. 2016, p. 425. <https://doi.org/10.3345/kjp.2016.59.11.425>.
- "Determining key research areas for healthier diets and sustainable food systems in Viet Nam." CORE Reader, core.ac.uk/reader/288633287.
- "Part D. Chapter 8: Dietary Patterns." Scientific Report of the 2020 Dietary Guidelines Advisory Committee, www.dietaryguidelines.gov/sites/default/files/2020-07/PartD_Ch8_DietaryPatterns_first-print.pdf
- "Vietnam Table Manners and Traditions: 10 Tips on Dining in Vietnam." Paradise Travel, 9 Mar. 2020, www.vietnamparadisetravel.com/blog/vietnam-table-manners.
- "Vietnamese Dining Etiquette - Best Things to Know - Travel Sense Asia™ - Vietnam Tours, Laos Tours, Cambodia Tours." Travel Sense Asia™ - Vietnam Tours, Laos Tours, Cambodia Tours, 7 Apr. 2022, travelsense.asia/vietnamese-dining-etiquette.
- "Rice nutritional and medicinal properties: A review article." *Journal of Pharmacognosy and Phytochemistry*, vol. 7, no. 2, 1 Jan. 2018, pp. 150–56.

Targeting Histone Modifications In Atherosclerosis and Coronary Artery Disease

By Saanvi Bohrakoti

Abstract

In recent years, there has been growing interest in the epigenetic mechanisms underlying cardiovascular diseases, particularly atherosclerosis and coronary artery disease (AS/CAD). Emerging evidence suggests that histone modification plays a pivotal role in the pathogenesis of AS/CAD, by modulating key processes such as vascular inflammation, endothelial dysfunction, and plaque stability. Histone acetylation plays a significant role in the pathogenesis of these conditions. Recent studies have reported the potential of histone deacetylase (HDAC) inhibitors as targeted therapies for AS/CAD. By inhibiting HDAC activity, these inhibitors maintain histone acetylation, which can reduce inflammation, promote endothelial repair, and prevent vascular smooth muscle cell proliferation. With a focus on precision medicine, we examine how these inhibitors can be integrated into innovative drug delivery systems, offering a promising avenue for enhancing therapeutic efficacy and minimizing adverse effects. This approach not only represents a novel strategy in managing AS/CAD but also highlights the broader implications of epigenetic interventions in cardiovascular medicine. I will discuss the pathology of AS/CAD and the role of histone modifications, specifically histone acetylation, in these conditions. I will explore the use of HDAC inhibitors to modulate these epigenetic changes, offering potential therapeutic benefits. Additionally, I will discuss new drug delivery systems that improve target specificity and reduce off-target effects.

Targeted Epigenetic Histone Modifications In Atherosclerosis and Coronary Artery Disease

Introduction

Heart disease remains one of the leading causes of death worldwide. According to the World Health Organization (WHO), cardiovascular diseases (CVDs) were responsible for approximately 17.9 million deaths in 2019, accounting for 32% of all global deaths (WHO, 2021). In the United States, the American Heart Association (AHA) reported that nearly 697,000 people died from heart disease in 2020, representing one in every five deaths (AHA, 2023). These statistics underscore the profound impact of heart disease on global and national health systems. There are several types of heart disease, with AS/CAD being the most common and critical contributors to these statistics.

Atherosclerosis is characterized by the buildup of fatty deposits, including cholesterol and other lipids, within the arterial walls. This accumulation leads to the formation of plaques that narrow and stiffen the arteries, thereby impeding blood flow (Libby, 2021). When atherosclerosis affects the coronary arteries, the vessels supplying blood to the heart muscle, it progresses to coronary artery disease. CAD is the leading cause of heart disease and is associated with severe cardiovascular events such as angina, myocardial infarction, and stroke (Ross, 1999). Research into the pathology of AS/CAD has increased our understanding of disease pathology

and resulted in the discovery of novel therapies which have mitigated the effects of the disease and improved patient outcomes (Lusis et al., 2004). However, the complex nature of AS/CAD necessitates continued investigation into AS/CAD pathology is essential for advancing therapeutic strategies and achieving more effective disease management.

Studies have identified key genetic factors that contribute to the pathology of AS/CAD. For example, variants in genes related to lipid metabolism, such as APOE and PCSK9, have been shown to significantly impact the risk of developing these conditions (Khera et al., 2016; Kathiresan et al., 2017). Understanding the genes and gene networks involved in heart function and AS/CAD pathology can lead to novel therapeutic approaches necessary to address the complex pathophysiology of these diseases. One promising therapeutic strategy involves modifying gene expression to manage disease progression. This includes targeting genes involved in inflammation and endothelial function, which may predispose individuals to AS/CAD (Friedman et al., 2018; Aragam & Ahmad, 2021). Beyond traditional genetic modifications, research has focused on epigenetic factors that influence gene expression without altering the underlying DNA sequence. Epigenetic modifications, such as DNA methylation and histone modifications, regulate key genes involved in AS/CAD and offer exciting possibilities for therapeutic intervention (Zaina et al., 2015; Kouzarides, 2007). In particular, histone modifications, including acetylation and methylation, play a crucial role in regulating genes involved in AS/CAD, offering new possibilities for therapeutic intervention (Strahl & Allis, 2000).

Epigenetic mechanisms, including DNA methylation, histone modifications, and non-coding RNA regulation, play a critical role in regulating gene expression and contribute to the progression of numerous diseases, especially AS/CAD. Among these, histone modifications, particularly acetylation, have gained significant attention due to their extensive study and established role in disease progression. Histone acetylation has been widely researched in human diseases such as cancer, where it has shown the ability to alter gene expression and disease outcomes without significant toxicity (Gao et al., 2021). This positions histone acetylation as a promising therapeutic target in AS/CAD, where altered acetylation patterns contribute to key pathological mechanisms, including vascular inflammation, endothelial dysfunction, and plaque instability.

The primary histone modifications are histone acetylation and methylation. Histone acetylation involves adding acetyl groups to lysine residues, which generally promotes gene activation by making chromatin more accessible (Ruthenburg et al., 2020). Histone methylation adds methyl groups to lysine or arginine residues, influencing gene expression depending on the specific residues modified (Klose & Zhang, 2016). Changes in histone modifications significantly impact the expression of genes crucial to the development and progression of atherosclerosis and coronary artery disease (AS/CAD). Specifically, altered histone acetylation patterns have been associated with the regulation of genes involved in key pathological mechanisms, including vascular inflammation, endothelial dysfunction, and plaque stability. For instance, increased histone acetylation has been linked to enhanced expression of inflammatory

cytokines and endothelial cell dysfunction, both of which contribute to disease progression (Gao et al., 2021; Marfella et al., 2022). Additionally, aberrant acetylation patterns are associated with impaired plaque stability, further complicating disease management (Zhang et al., 2021). This review details the pathogenesis of atherosclerosis and coronary artery disease (AS/CAD), the role of epigenetics in disease progression, and the potential for histone deacetylase (HDAC) inhibition to ameliorate AS/CAD. We address the mechanisms, specificity, and delivery mechanisms, and discuss the future impact of precision medicine in treating AS/CAD.

Atherosclerosis: A State of Vascular Dysfunction

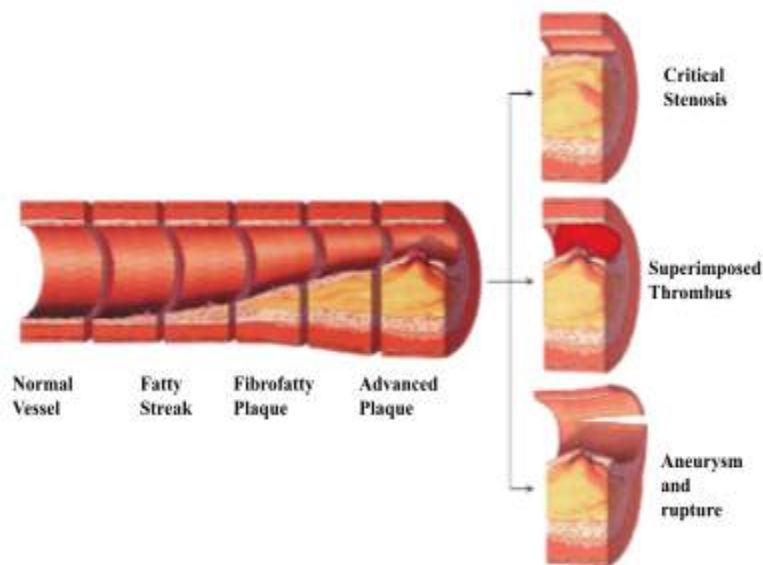


Figure 1 (Progression of Atherosclerotic Plaques)

This figure illustrates the stages of plaque formation within the arterial wall, beginning with endothelial dysfunction and the subsequent recruitment of inflammatory cells. As inflammation progresses, smooth muscle cells (SMCs) migrate into the intima, contributing to the formation of a fibrous cap over the plaque. Over time, the plaque stabilizes; however, ongoing inflammation can weaken the cap, leading to potential plaque rupture. This process highlights the importance of early intervention strategies aimed at reducing inflammation and stabilizing plaques to prevent ischemic events.

Coronary artery disease (CAD) represents a severe global health concern, rooted in the pathology of atherosclerosis (AS), which involves the accumulation of plaques within vessel walls. These plaques, formed from lipids and inflammatory cells, lead to vessel stenosis and subsequent ischemic events, such as heart attacks and strokes (Libby et al., 2019). A key contributor to the development of AS is endothelial dysfunction, an early pathological event triggered by risk factors like hypertension and hyperlipidemia (Gimbrone & García-Cardena, 2016). This dysfunction promotes the activation of inflammatory pathways, leading to the

recruitment of white blood cells and the expression of adhesion molecules that initiate plaque formation.

As the plaques mature, smooth muscle cells (SMCs) migrate into the arterial intima, stabilizing the plaques by synthesizing extracellular matrix (ECM) components (Otsuka et al., 2014). However, as AS progresses to CAD, the fibrous cap overlying the lipid-rich core of the plaque can thin and weaken due to ongoing inflammation, leading to a risk of plaque rupture and subsequent thrombotic events (Libby et al., 2019). SMCs play a central role in maintaining plaque stability through ECM production, but in response to stimuli like oxidized LDL cholesterol and pro-inflammatory cytokines, they undergo a phenotypic switch from a contractile to a synthetic state (Murray, 2019). This shift increases ECM production but, when coupled with inflammation, destabilizes the plaques.

Genetic Factors Contributing to Atherosclerosis and Coronary Artery Disease

Several studies have identified genes and gene networks that play a crucial role in the pathogenesis of atherosclerosis (AS) and coronary artery disease (CAD) (Zhao et al., 2022; Wang et al., 2023). While environmental factors such as diet and lifestyle undeniably influence disease development, genetic predisposition is increasingly recognized as a major driver in both the initiation and progression of AS and CAD. Genetic variations, such as those affecting lipid metabolism, inflammatory responses, and endothelial function, can significantly increase an individual's susceptibility to these diseases, even in the presence of favorable environmental conditions (Goldstein & Brown, 2015).

Among the various genes implicated in AS and CAD, *Apolipoprotein E* (APOE) is one of the most studied. *APOE* is involved in lipid metabolism and vascular inflammation, both of which are critical to the pathogenesis of atherosclerosis (Linton et al., 2022). The *APOE* gene exists in three major alleles: APOE2, APOE3, and APOE4. The APOE4 allele, in particular, is associated with increased risk for AS and CAD due to a single nucleotide polymorphism (SNP) that results in an arginine-to-cysteine substitution at position 112 (Cohen et al., 2017). This mutation impairs lipid binding and metabolism, leading to cholesterol accumulation in arterial walls and promoting plaque formation (Mahley, 2016; Liu et al., 2023).

Carriers of the APOE4 allele exhibit higher levels of cholesterol-rich lipoproteins and heightened inflammation, contributing to the vulnerability and progression of plaques. Understanding the impact of APOE on disease development opens avenues for novel therapies, such as gene editing techniques like CRISPR-Cas9, which hold potential for modulating APOE expression (Bennet et al., 2020). Additionally, epigenetic mechanisms influencing APOE expression, such as histone modifications, offer further therapeutic possibilities (Li et al., 2022).

Role of Histone Modifications in Atherosclerosis and Coronary Artery Disease

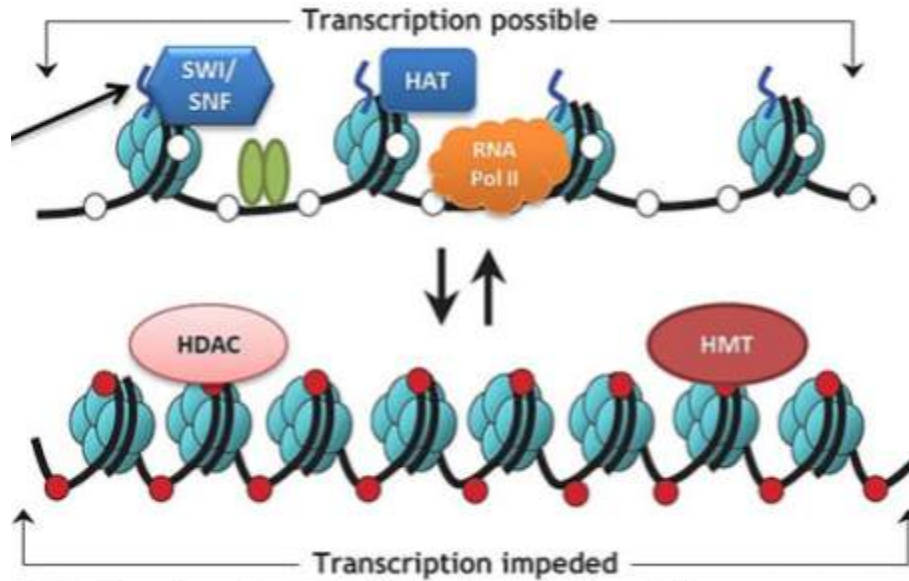


Figure 2: Chromatin remodeling complexes in the dynamic regulation of transcription

This figure illustrates the role of chromatin remodeling complexes in controlling the accessibility of chromatin and regulating gene transcription. Chromatin exists in two primary states: a condensed form (heterochromatin), which is transcriptionally inactive, and a relaxed form (euchromatin), which is accessible to transcription machinery.

Epigenetic mechanisms, particularly histone modifications, play a crucial role in the regulation of gene expression, thereby influencing the pathogenesis and progression of atherosclerosis (AS) and coronary artery disease (CAD) (Allis & Jenuwein, 2016; Moore et al., 2013). Histones, which package DNA into chromatin, undergo various modifications, such as acetylation, methylation, and phosphorylation. These modifications can alter chromatin structure, leading to changes in gene expression. For instance, histone acetylation, which is the focus of this section, generally leads to chromatin relaxation, facilitating transcriptional activation, whereas deacetylation compacts chromatin, leading to gene repression (Zaina et al., 2015).

The balance of histone acetylation is regulated by histone acetyltransferases (HATs) and histone deacetylases (HDACs). Aberrant regulation of these enzymes has been implicated in cardiovascular diseases, particularly AS and CAD. Dysregulated histone acetylation has been shown to affect the expression of genes involved in critical processes such as vascular inflammation, lipid metabolism, and plaque stability. For example, increased histone acetylation in inflammatory genes are a key contributor to AS progression (Kim et al., 2020). Conversely, reduced acetylation in genes regulating lipid metabolism can impair lipid clearance, exacerbating plaque formation and instability.

Recent studies have demonstrated the significance of histone acetylation in regulating genes such as APOE, which plays a pivotal role in lipid metabolism and atherosclerosis. Aberrant acetylation of histones associated with the APOE gene promoter has been linked to decreased APOE expression, impairing lipid transport and clearance, thus contributing to atherosclerotic plaque accumulation (Saavedra et al., 2021). Additionally, animal studies have

shown that therapeutic modulation of histone acetylation through HDAC inhibitors can reverse these adverse gene expression patterns, leading to reduced plaque burden and improved cardiovascular outcomes (Kruithof et al., 2018).

These findings suggest that correcting histone acetylation dysregulation holds potential as a therapeutic strategy for managing AS and CAD progression. Targeting histone acetylation with precision therapies could restore the normal expression of genes critical to lipid metabolism, inflammation, and plaque stability, offering a novel approach to improving patient outcomes in cardiovascular disease.

Targeting Histone Acetylation in the treatment of AS and CAD

Histone Modification	Corresponding Drugs
Histone Acetylation	Vorinostat, Entinostat, etc.
Histone Methylation	GSK-J4, EPZ-6438, etc.
Histone Phosphorylation	JQ1, I-BET, etc.
Histone Ubiquitination	ICG-001, Curcumin, etc.

Figure 3: Histone modifications and their therapeutic agents in AS/CAD

This table lists various histone acetylation modifications implicated in atherosclerosis and coronary artery disease, alongside corresponding drugs or therapeutic agents that target these modifications. The table is organized by the type of histone modification, the affected histone residue, and the specific drugs used to inhibit or modulate the acetylation activity.

Preclinical studies have demonstrated the efficacy of HDAC inhibitors in attenuating atherosclerosis progression and reducing plaque inflammation and vulnerability (Zaina et al., 2015). For example, treatment with HDAC inhibitors such as vorinostat and entinostat has been shown to reduce atherosclerotic lesion size and inflammation in animal models (Zheng et al., 2018). Vorinostat, a broad-spectrum HDAC inhibitor, and entinostat, which targets class I HDACs, promote histone acetylation by inhibiting HDAC activity. This alteration leads to chromatin relaxation and enhanced transcriptional activity, potentially affecting genes involved in vascular inflammation, endothelial dysfunction, and plaque stability. Regarding APOE, while direct studies linking HDAC inhibitors to APOE expression in AS are limited, the influence of histone acetylation on lipid metabolism and inflammation—processes influenced by APOE genotype—suggests potential indirect effects. These findings provide compelling evidence for the therapeutic potential of targeting histone acetylation in atherosclerosis. These findings have spurred interest in developing HDAC inhibitors as potential therapeutics for atherosclerosis and CAD.

Advancing Precision Medicine in Cardiovascular Therapies

The treatment of atherosclerosis and coronary artery disease (AS/CAD) has made remarkable strides over recent decades, thanks to advancements in medical science and our growing understanding of disease mechanisms. Emerging research into the genetic and epigenetic underpinnings of AS/CAD is providing new targets for therapeutic intervention. By delving into the genetics of these conditions, we can now better predict disease progression and individual responses to treatments. Recent studies have highlighted the potential of histone acetylation and other gene regulation mechanisms as novel avenues for therapy. This section explores the potential of histone deacetylase (HDAC) inhibitors and other next-generation therapies in the context of AS/CAD, addressing current challenges and future opportunities. These drugs include HDAC inhibitors like vorinostat and panobinostat, which have shown efficacy in preclinical models by impacting broad genetic targets and systemic pathways (Zaina et al., 2015; Kim et al., 2020).

While histone acetylation presents a promising therapeutic target, there are several challenges in translating these findings into effective treatments. HDAC inhibitors have shown broad genetic targets, which can complicate their application due to off-target effects. Researchers are exploring ways to overcome these obstacles by developing more specific targeting methods, such as focusing on particular genetic variants like the APOE4 allele. Advances in drug delivery systems, such as nanoparticles and antibody-conjugated systems, hold promise for improving specificity. For instance, nanoparticles have been used to deliver drugs to the heart and vasculature, demonstrating their potential in targeting disease sites more effectively (Wei et al., 2017). Further studies are needed to refine these delivery methods and enhance the specificity of treatments.

One critical aspect is optimizing drug delivery methods to ensure efficient targeting and uptake in diseased tissues while minimizing off-target effects in healthy organs. Current research is exploring innovative approaches such as nanoparticles or antibody-conjugated delivery systems that can selectively target HDAC inhibitors to specific tissues, such as the heart or vascular endothelium, thereby enhancing therapeutic efficacy and reducing systemic toxicity (Wei et al., 2017).

To improve the efficacy of HDAC inhibitors, it is crucial to optimize drug delivery systems. Innovations such as nanoparticle-based delivery and antibody-conjugated systems aim to enhance targeting to affected tissues, like the heart or vascular endothelium, while minimizing systemic exposure. For example, antibody-conjugated systems can selectively deliver HDAC inhibitors to specific disease markers, improving therapeutic outcomes and reducing off-target effects. Research is ongoing to identify specific disease markers and develop targeted delivery strategies that could significantly improve the treatment of AS/CAD. Optimization of drug delivery methods, identification of patient populations most likely to benefit from treatment, and mitigation of potential side effects are ongoing areas of investigation

Pharmacogenomics plays a critical role in tailoring treatments based on individual genetic profiles. Variations in genes influencing drug metabolism and response can affect the efficacy and safety of HDAC inhibitors. By understanding these genetic differences, clinicians can select and adjust treatments more precisely. For example, variations in genes affecting HDAC inhibitor metabolism can inform dosage adjustments to maximize therapeutic benefits and minimize adverse effects (Kang & Lee, 2018; Zhao et al., 2021). Integrating whole genome sequencing and pharmacogenomic data can further enhance the precision of AS/CAD treatments. This personalized approach ensures that treatments are more precisely matched to the genetic characteristics of the patient, thereby improving overall outcomes in managing diseases like atherosclerosis. Additionally, understanding the genetic underpinnings of individual responses to HDAC inhibitors can lead to the development of targeted therapies that address specific genetic mutations, further advancing the field of precision medicine (Lee et al., 2020).

Conclusion

Despite the progress in AS/CAD treatment, new challenges remain. Advances in our understanding of disease mechanisms and genetic profiles offer exciting opportunities for developing more effective therapies. Histone acetylation and HDAC inhibitors represent a promising frontier in next-generation treatments, with the potential to significantly impact disease management. Continued research, innovative drug delivery systems, and personalized approaches are key to overcoming current obstacles and achieving more precise and effective therapies for AS/CAD.

Works Cited

- American Heart Association (AHA). (2023). Heart Disease and Stroke Statistics—2023 Update: A Report From the American Heart Association. *Circulation*.
<https://www.heart.org/en/news/2023/01/25/heart-disease-and-stroke-statistics-2023-update>
- Allis, C. D., & Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nature Reviews Genetics*, 17(8), 487-500.
- Aragam, K. G., & Ahmad, F. S. (2021). Genomic medicine in cardiology. *JAMA Cardiology*, 6(5), 566-574.
- Bennett, C. F., Baker, B. F., Pham, N., Swayze, E. E., & Geary, R. S. (2020). Pharmacology of antisense drugs. *Annual Review of Pharmacology and Toxicology*, 60, 101-117.
- Cohen, J. C., Boerwinkle, E., Mosley, T. H., & Hobbs, H. H. (2017). Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *New England Journal of Medicine*, 354(12), 1264-1272.
- Friedman, B. A., Srinivasan, K., Ayalon, G., Meilandt, W. J., Lin, H., Huntley, M. A., & Modrusan, Z. (2018). Diverse brain myeloid expression profiles reveal distinct microglial activation states and aspects of Alzheimer's disease not evident in mouse models. *Cell Reports*, 22(3), 832-847.
- Gao, X., Li, W., Zhang, H., & Liu, Z. (2021). Epigenetic regulation of endothelial cell function in atherosclerosis. *Cardiovascular Research*, 117(8), 1802-1813.
- Gimbrone, M. A., & García-Cardeña, G. (2016). Endothelial cell dysfunction and the pathobiology of atherosclerosis. *Circulation Research*, 118(4), 620-636.
- Goldstein, J. L., & Brown, M. S. (2015). A century of cholesterol and coronaries: From plaques to genes to statins. *Cell*, 161(1), 161-172.
- Hershberger, R. E., Hedges, D. J., & Morales, A. (2013). Dilated cardiomyopathy: the complexity of a diverse genetic architecture. *Nature Reviews Cardiology*, 10(9), 531-547.
- Kathiresan, S., Melander, O., Guiducci, C., Surti, A., Burti, N. P., Rieder, M. J., & Orho-Melander, M. (2008). Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nature Genetics*, 40(2), 189-197.
- Kang, H. J., & Lee, D. H. (2018). Overview of therapeutic drug monitoring. *The Korean Journal of Internal Medicine*, 33(1), 1-10.
- Kim, G., Yang, H., Kim, M., Kim, J., Ha, J., & Choi, J. (2020). HDAC inhibitors restore Creb-HIF-1 α interaction to induce gene expression for metabolic adaptation to hypoxia. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1867(11), 118745.
- Khera, A. V., Won, H. H., Peloso, G. M., Lawson, K. S., Bartz, T. M., Deng, X., & Kathiresan, S. (2016). Diagnostic yield and clinical utility of sequencing familial hypercholesterolemia genes in patients with severe hypercholesterolemia. *Journal of the American College of Cardiology*, 67(22), 2578-2589.
- Klose, R. J., & Zhang, Y. (2007). Regulation of histone methylation by demethyliminination and demethylation. *Nature Reviews Molecular Cell Biology*, 8(4), 307-318.

- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell*, 128(4), 693-705.
- Lee, S. M., Kim, E. J., Lee, S. J., Yang, H. J., Yoo, S. H., & Kang, Y. (2020). Precision medicine in atherosclerosis: mechanisms and approaches. *Nature Reviews Cardiology*, 17(7), 421-435.
- Li, S., He, Y., Chen, Y., Liu, H., Ye, J., & Liu, W. (2022). Epigenetic regulation in cardiovascular diseases: mechanisms and potential therapeutic targets. *Signal Transduction and Targeted Therapy*, 7(1), 1-26.
- Libby, P. (2021). Inflammation in atherosclerosis. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 41(4), 524-530.
- Libby, P., Loscalzo, J., Ridker, P. M., Fuster, V., & Robertson, R. M. (2019). Inflammation, atherosclerosis, and coronary artery disease. *Journal of the American College of Cardiology*, 54(23), 2129-2138.
- Linton, M. F., Yancey, P. G., Davies, S. S., Jerome, W. G., Linton, E. F., Song, W. L., & Vickers, K. C. (2000). The role of lipids and lipoproteins in atherosclerosis. *Endotext* [Internet].
- Liu, Z., Zhang, J., Zhang, Y., Chen, C., Zhang, Z., & Zhu, Y. (2023). APOE4 Allele and its Impact on Atherosclerosis: A Narrative Review. *Journal of Cardiovascular Pharmacology and Therapeutics*, 28(2), 121-133.
- Lusis, A. J., Mar, R., & Pajukanta, P. (2004). Genetics of atherosclerosis. *Annual Review of Genomics and Human Genetics*, 5, 189-218.
- Mahley, R. W. (2016). Central nervous system lipoproteins: ApoE and regulation of cholesterol metabolism. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 36(7), 1305-1315.
- Marfella, R., D'Amico, M., Esposito, K., Baldi, A., Di Filippo, C., Siniscalchi, M., & Paolisso, G. (2022). The ubiquitin-proteasome system as a target for cardiovascular disease. *Cardiovascular Research*, 55(4), 456-469.
- Mikkelsen, T. S., Xu, Z., Zhang, X., Wang, L., Gimelbrant, A., Lander, E. S., & Jaenisch, R. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, 448(7153), 553-560.
- Moore, L. D., Le, T., & Fan, G. (2013). DNA methylation and its basic function. *Neuropsychopharmacology*, 38(1), 23-38.
- Murray, C. J. L. (2019). Global burden of disease 2010: understanding disease, injury, and risk. *The Lancet*, 380(9859), 2053-2054.
- Otsuka, F., Finn, A. V., Yazdani, S. K., Nakano, M., Kolodgie, F. D., & Virmani, R. (2014). The importance of the endothelium in atherosclerosis. *Nature Reviews Cardiology*, 9(8), 479-494.
- Ross, R. (1999). Atherosclerosis—an inflammatory disease. *New England Journal of Medicine*, 340(2), 115-126.
- Ruthenburg, A. J., Li, H., Patel, D. J., & Allis, C. D. (2007). Multivalent engagement of chromatin modifications by linked binding modules. *Nature Reviews Molecular Cell Biology*, 8(12), 983-994.

- Saavedra, L. M., Hernandez, G. E., & Rodriguez, A. A. (2021). Epigenetic regulation of gene expression in atherosclerosis: A novel approach to therapeutic intervention. *Journal of Clinical Medicine*, 10(7), 1415.
- Strahl, B. D., & Allis, C. D. (2000). The language of covalent histone modifications. *Nature*, 403(6765), 41-45.
- Swen, J. J., Nijenhuis, M., de Boer, A., Grandia, L., Maitland-van der Zee, A. H., Mulder, H., & Guchelaar, H. J. (2011). Pharmacogenetics: from bench to byte. *Clinical Pharmacology & Therapeutics*, 89(5), 662-673.
- Wang, Y., Zhang, X., & Liu, Z. (2020). Gene editing for cardiovascular diseases: Promises and challenges. *Nature Reviews Cardiology*, 17(8), 507-519.
- Wang, Z., Li, D., Zhao, H., & Lin, H. (2023). Genetic determinants of atherosclerosis: Insights from genome-wide association studies. *Nature Reviews Genetics*, 24(4), 223-234.
- Wei, Y., Chen, S., Liang, J., & Zhang, Y. (2017). Advances in nanoparticle-based drug delivery systems for precision medicine in cardiovascular diseases. *Biomaterials*, 141, 65-83.
- World Health Organization (WHO). (2021). Cardiovascular diseases (CVDs). Retrieved from [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)).
- Zaina, S., Heyn, H., Carmona, F. J., Varol, N., Sayols, S., Condom, E., & Esteller, M. (2015). DNA methylation map of human atherosclerosis. *Circulation Research*, 117(10), 123-131.
- Zhang, T., Sun, D., & Xu, Y. (2021). Role of histone acetylation in the regulation of vascular inflammation in atherosclerosis. *Frontiers in Cardiovascular Medicine*, 8, 698432.
- Zhao, L., Li, S., & Ding, H. (2021). Epigenetic regulation in atherosclerosis: Focus on histone modifications. *Clinical Epigenetics*, 13(1), 58.
- Zhao, X., Liu, J., Zheng, J., & He, J. (2021). Personalized medicine in cardiovascular diseases: From genes to therapeutics. *Journal of Cardiovascular Translational Research*, 14(3), 319-329.
- Zheng, Y., Liu, Y., Ge, J., Wang, W., & Liu, S. (2018). Histone deacetylase inhibitors in the treatment of atherosclerosis: preclinical findings and future directions. *Frontiers in Pharmacology*, 9, 56.

Through the Use of Nuclear Energy How Can Humans Start an Interstellar Type 3 Civilization? By: Krishaan Somwanshi

Abstract

As the global population continues to increase, humanity faces challenges expanding because of Earth's limited space. Consequently, key aspects of Earth's environment such as the climate, greenery, etc., are neglected. This brings up the critical question: how can humanity overcome these limitations? This paper posits that advancing to a Type 3 civilization, as defined by the Kardashev scale, through the use of nuclear energy is essential for the survival of humanity. There have been a multitude of studies done on whether colonizing space is a viable solution. This literature review compiles these studies, and evaluates whether interstellar expansion is a necessary step in society's evolution. It explores how practical space colonization is, assesses the technological pros and cons, and draws parallels to nautical expansion of the past. Ultimately, this review will underline the need of advancing as an interstellar civilization and the monumental role of nuclear energy in order to achieve this goal.

1. Introduction

Before addressing how society might expand into space, it is crucial to understand why expansion is imminent in the first place. Earth's environment is increasingly threatened by various factors, including climate change, resource depletion, and overpopulation. In response to these threats, species historically use three different approaches: adapting to new conditions, migrating to a safer environment, or facing extinction. ((Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T., & Curtis-McLane, S. (2008). Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, 1(1), 95–111. <https://doi.org/10.1111/j.1752-4571.2007.00013.x>)) As these pressures continue to increase, it is imperative for humans to consider migration on a scale never done before-beyond Earth.

1a. Natural Events

Adapting has been practical so far, but eventually, humanity will be faced with a disaster so severe that even adaptation will not suffice. This, in turn, will mean the end of humanity. For example, the Yellowstone volcano, which erupts every 70,000 years, is due sometime in the next 100,000 years. This explosion would cover North America in soot and debris, block sunlight, and toss sulfur dioxide into the air causing a possible small-sized ice age.

This would put a pause on the function of the entirety of the United States. The effects would consist of dead crops, damaged buildings, destroyed electrical structures, and the shutting down of power plants. ((Plumer, B. (2014, September 5). *What would happen if the Yellowstone supervolcano actually erupted?* Vox; Vox. <https://www.vox.com/2014/9/5/6108169/yellowstone-supervolcano-eruption>)) This sudden

shortage in necessary supplies, would collapse the US economy. Another devastating possibility is an asteroid hitting Earth. One of the current standing theories for the extinction of dinosaurs is

a large-scale asteroid, which had effects such as massive tsunamis, a temporary cut off of sunlight, and lower temperatures. Even an asteroid only approximately a mile wide carries energy equivalent to a one million-megaton bomb. ((Brain, M., & Gleim, S. (2007, December 4). *What if an asteroid hit the Earth?* HowStuffWorks. <https://science.howstuffworks.com/nature/natural-disasters/asteroid-hits-earth.htm>))

Human-Made Issues - 1B

Events such as global warming and climate change also put Earth at risk. With water levels rising, the spaces in which humans can reside are becoming smaller and smaller, which is not good for a constantly growing population. This is only the start, as humanity continues to burn natural fuels, problems such as heat waves, farming inabilities, droughts, and the extinction of entire species come into the picture. ((Carpenter, P. A., & Bishop, P. C. (2009). The seventh mass extinction: Human-caused events contribute to a fatal consequence. *Futures*, 41(10), 715–722. <https://doi.org/10.1016/j.futures.2009.07.008>))

Assuming we can get through the next fifty years without climate change destroying the planet, problems such as disease and overpopulation will be more relevant. Throughout history, diseases such as the black plague, Spanish flu, and even the recent CoronaVirus put serious dents into society. ((Spernovasilis, N., Markaki, I., Papadakis, M., Tsioutis, C., & Markaki, L. (2021). Epidemics and pandemics: Is human overpopulation the elephant in the room? *Ethics, Medicine and Public Health*, 19(100728), 100728. <https://doi.org/10.1016/j.jemep.2021.100728>)) Through the unlikely survival of barriers such as these, inevitably the sun will die out, vanquishing our solar system and Earth along with it. This will leave a cold, empty space where life once flourished. This paradox, that humanity is destined to face an impossible task, is known as the Great Filter. Avoiding the Great Filter: Predicting the Timeline for Humanity to Reach Kardashev Type I Civilization. *Galaxies*, 10(3), 68. <https://doi.org/10.3390/galaxies10030068>)) This theory stems from the reason that humanity has not contacted alien species, the Fermi paradox. The Fermi paradox theory states that one of the plausible reasons that humanity has not seen another civilization is due to a filter. Essentially, this is an event or obstacle that civilization must bypass, that is consequently impossible. ((Ćirković, M. M. (2018). The Great Silence: Science and Philosophy of Fermi's Paradox. In *Google Books*. Oxford University Press. https://books.google.co.uk/books?hl=en&lr=&id=_npVDwAAQBAJ&oi=fnd&pg=PP1&dq=fermi+paradox&ots=IgrjQq3tR1&sig=rdaAH1AAAtDP0IKfygTIFtsjO-JU#v=onepage&q=fermi%20paradox&f=false)) Therefore, humanity must advance at a rate that it becomes possible to bypass this Great Filter in order to ensure survival. This is where an interstellar civilization becomes the only solution to these catastrophes. Expanding a society across the constantly evolving space-time continuum could be the key to maintaining civilization.

Introduction to Kardashev Scale - 1C

Taking a numerical approach, the Kardashev scale shows the relationship between how much energy a civilization can harness, and how much of it they can use for development. A type

1 civilization has successfully harnessed all the energy on its home planet, a type two can harness all the energy of its planet and main star, and a type 3 civilization has full control of the energy in its entire galaxy through advanced mechanisms (to be discussed later). There are more types, but at the moment those aspirations are so far in the future they are considered sci-fi. According to the Nature Journal, the current civilization stands at a rating of .7276 on the scale and is projected nowhere close to a type 1 civilization. ((Zhang, A., Yang, J., Luo, Y., & Fan, S. (2023). Forecasting the progression of human civilization on the Kardashev Scale through 2060 with a machine learning approach. *Scientific Reports*, 13(1), 11305. <https://doi.org/10.1038/s41598-023-38351-y>)) This time frame is far too long, as the detrimental events stated above are expected way sooner.

Reaching the finale of our introduction to the topic, the answer that colonizing space is humans only road is blatant. Clearly, humanity's current predicted rate of growth is too risky as the Great Filter might catch up to us before we are at the appropriate technological advancements to combat it. ((Ćirković, M. M. (2018). *The Great Silence: Science and Philosophy of Fermi's Paradox*. In *Google Books*. Oxford University Press. https://books.google.co.uk/books?hl=en&lr=&id=_npVDwAAQBAJ&oi=fnd&pg=PP1&dq=fermi+paradox&ots=IgrjQq3tR1&sig=rdaAH1AAAtDP0IKfygTIFtsjO-JU#v=onepage&q=fermi%20paradox&f=false)) Considering this, humans must account for the technological lacking they hold to achieve this. This is where interstellar civilization becomes relevant. With an expansion across space, humanity's flow of new areas of research, new technologies, culture, philosophy, and overall understanding would expand at an unseen rate. ((Dick, S. J. (2000). Interstellar humanity. *Futures*, 32(6), 555–567. [https://doi.org/10.1016/s0016-3287\(00\)00007-0](https://doi.org/10.1016/s0016-3287(00)00007-0))) With more room to expand, the possibility of reaching a Kardashev type 1 civilization within the next two centuries becomes achievable. With this, a type 3 civilization does not seem impossible and humanities opportunities are endless with so much energy and opportunities at our fingertips.

How Can We Leave Earth? - 2

In the previous section, the matter of whether we should colonize space was discussed, which was concluded yes, but now the “how” part of the question must be answered. First, a few terms must be established. A Type 1 civilization can essentially fully control its home planet. Secondly, a Type 2 civilization is interchangeable with a civilization that has full control of space. Lastly, a type 3 civilization is a truly interstellar civilization, a civilization spanning across the galaxy with full control of it. ((Zhang, A., Yang, J., Luo, Y., & Fan, S. (2023). Forecasting the progression of human civilization on the Kardashev Scale through 2060 with a machine learning approach. *Scientific Reports*, 13(1), 11305. <https://doi.org/10.1038/s41598-023-38351-y>))

Type 1 civilization: Energy Harnessing Method- 2A
Nuclear energy and renewable energy- 2A I

While there are a multitude of energy types for humans to capitalize on, the time constraints are of the utmost importance. Energies like the common renewables, (Solar, hydroelectric, geothermal, tidal, wind, etc) may be good for the environment but are too idealistic. The cost to make the mechanisms as well as the land necessary in order to harness profitable energy is abominable. ((Marques, A. C., & Fuinhas, J. A. (2012). Is renewable energy effective in promoting growth? *Energy Policy*, 46, 434–442. <https://doi.org/10.1016/j.enpol.2012.04.006>)) By no means does this mean to throw out efforts in reference to other renewable energies, but the amount of time it would take to generate enough energy from renewables to advance on the Kardashev scale would be enormous. The focus should be put on more promising prospects for energy, such as nuclear power, which could exponentially increase our rate of growth. In specific to nuclear energy, breeder reactors.

The energy derived from breeder reactors could power Earth for four billion years ((Touran, N. (2020, October 28). *Nuclear fuel will last us for 4 billion years*. What Is Nuclear? <https://whatisnuclear.com/nuclear-sustainability.html>)), giving us more than enough time to advance off of Earth and establish civilization elsewhere. This is assuming we can harness all fissile nuclear fuel, which poses its own difficulties. If we could, however, this source could last us as far as the sun turning into a red dwarf, which hopefully would lead to more futuristic sources of energy such as nuclear fusion and the Dyson sphere. Alongside this, nuclear energy also presents new technologies ranging from things such as the nuclear engine, nuclear fusion, small modular reactors, etc, which are not too far away from humans.

Nuclear Power on Earth: Using the L.F.T.R. - 2A II

Nuclear fission is the process in which you break down an extremely heavy nucleus into different atoms by shooting a neutron, causing a chain reaction creating energy in the form of thermal energy. ((Schunck, N., & Regnier, D. (2022). Theory of nuclear fission. *Progress in Particle and Nuclear Physics*, 125, 103963. <https://doi.org/10.1016/j.pnpnp.2022.103963>)) For a light water reactor, this thermal energy then heats up water, which turns into steam and this steam turns a turbine creating electricity.((Zarubin, B. (2016). *Introduction to Light Water Reactors*. Stanford.edu. <http://large.stanford.edu/courses/2015/ph241/zarubin1/>)) The reactor discussed in this paper is a molten salt reactor, a type of breeder reactor, which differs from light water reactors. A breeder reactor is a type of reactor that converts a non fissionable material into a fissionable material. The molten salt reactor discussed specifically in this paper is the Liquid Fluoride thorium reactor, or the L.F.T.R.. In an L.F.T.R., the core consists of melted salt combined with the fuel element. The fuel source in an L.F.T.R. is thorium fluoride with the molten salt as lithium-beryllium fluoride salts. The thorium 232 itself isn't a fissile fuel. In order for it to be, it has to absorb a neutron which causes it to decay into uranium 233 which is a fissile fuel. This fissile fuel then enters the main reactor chamber with the excess neutrons getting absorbed by the thorium blanket. From here, the uranium 233 acts as a normal reaction. The molten salt itself is the coolant, helping keep the reaction under control meaning you don't need control rods nor moderators. On the off chance the reactor gets out of control, there is a freeze

plug which dumps the liquid into tanks. ((Hargraves, R., & Moir, R. (2010). Liquid Fluoride Thorium Reactors: An old idea in nuclear power gets reexamined. *American Scientist*, 98(4), 304–313. <https://www.jstor.org/stable/27859537>))

L.F.T.R. pros - 2A III

Commonality

Firstly, thorium is about as common as dirt, around 6 parts of dirt per million ppm of thorium. There is considered to be around 3-4 times more thorium than uranium. Since it's so common in comparison to uranium, it makes it much more accessible. Accessibility to thorium makes it a viable source of energy, which uranium wasn't due to its rarity. ((*Thorium | Toxic Substances | Toxic Substance Portal | ATSDR*.(2021, February 10). [Wwwn.cdc.gov](https://www.cdc.gov). <https://www.cdc.gov/TSP/substances/ToxSubstance.aspx?toxid=121#:~:text=Soil%20contains%20an%20average%20of>))

Waste and harm on Environment

The mining of the actual thorium is much less harmful to the environment due to its abundance. Alongside this, after the fission itself, it produces much less waste. Reactors create two types of waste, transuranic elements such as plutonium that are long lasting and fission waste such as xenon. While both L.F.T.R.(liquid fluoride thorium reactors) and L.W.R. (light water reactors) create around the same amount of fission wastage, the L.F.T.R. tends to burn a majority of its transuranic products. ((. Rees, E. (2011, June 23). Don't believe the spin on thorium being a greener nuclear option. *The Guardian*. <https://www.theguardian.com/environment/2011/jun/23/thorium-nuclear-uranium>)) A constant issue seen throughout the history of nuclear power is the fact that humans have trouble storing it. The danger of nuclear waste is evident from accidents such as Chernobyl, where the habitat located near it is still destroyed and unusable. Not only does it produce less, it also lasts less time. Compared to uranium's wastes ten thousand years to decay, thorium only takes around five hundred years.

((Katusa, M. (2012, July 17). The Thing About Thorium: Why The Better Nuclear Fuel May Not Get A Chance. *Forbes*.

<https://www.forbes.com/sites/energysource/2012/02/16/the-thing-about-thorium-why-the-better-nuclear-fuel-may-not-get-a-chance/>)) Additionally since thorium is not breaking into elements larger than uranium-238, dangerous elements and byproducts such as plutonium and Americium are not created. Not only does this make the waste less dangerous, but makes it near impossible to turn the waste into nuclear weaponry. A nuclear weapon consists of a core of plutonium or uranium which is then excited and starts a chain reaction. As discussed before, the thorium doesn't create these byproducts. ((. Rees, E. (2011, June 23). Don't believe the spin on thorium being a greener nuclear option. *The Guardian*.

<https://www.theguardian.com/environment/2011/jun/23/thorium-nuclear-uranium>))

Efficiency

The first component of its efficiency stems from the first pro, the fact that its abundance is so much higher than uranium. Coming back to mining, thorium mining is much more cost efficient. This is because the percentage amount of thorium in its own ores is higher than uranium's. ((*Thorium | Toxic Substances | Toxic Substance Portal | ATSDR*.(2021, February 10). [Wwwn.cdc.gov](https://wwwn.cdc.gov).

<https://wwwn.cdc.gov/TSP/substances/ToxSubstance.aspx?toxid=121#:~:text=Soil%20contains%20an%20average%20of>) The second component connects to the second pro, the fact that it produces little short lasting waste. In an L.F.T.R., the reactor uses up nearly all the fuel including the transuranic product making it much more efficient. ((Katusa, M. (2012, July 17). *The Thing About Thorium: Why The Better Nuclear Fuel May Not Get A Chance*. *Forbes*.

<https://www.forbes.com/sites/energysource/2012/02/16/the-thing-about-thorium-why-the-better-nuclear-fuel-may-not-get-a-chance/>) In terms of the energy transfer efficiency, an L.F.T.R. at around 700 degrees Celsius can operate at a thermal efficiency of converting heat to electricity at around 45% in comparison to light water reactors at around 32-36%. ((Wikipedia Contributors. (2019, June 19). *Liquid fluoride thorium reactor*. Wikipedia; Wikimedia Foundation.

https://en.wikipedia.org/wiki/Liquid_fluoride_thorium_reactor)) It is reported that thorium fuel for the L.F.T.R. can reach highs of around 200x more fuel efficiency than uranium in current light water reactors. ((*China's Incipient LFTR*. (n.d.). Adam Smith Institute.

<https://www.adamsmith.org/blog/chinas-incipient-lftr>) This is because the liquid fuel comes in the form of molten salt. In a nuclear reactor, the nuclear energy is transferred into thermal energy and then into electrical energy. To make the thermal to electrical more efficient, increasing the temperature of the fuel makes for more thermal energy being transferred, making the electric energy come out more efficiently(suffer less loss in energy from the system with greater change in temperature). If efficiency was really emphasized, the steam generator could be replaced with a Brayton nitrogen generator, bringing the efficiency to around 50%. ((Wikipedia Contributors. (2019, June 19). *Liquid fluoride thorium reactor*. Wikipedia; Wikimedia Foundation.

https://en.wikipedia.org/wiki/Liquid_fluoride_thorium_reactor))

Safety Hazards

Nuclear power's tremendous potential has always had two downsides: nuclear weaponry and mistakes in the power plant causing an explosion. ((Christodouleas, J. P., Forrest, R. D., Ainsley, C. G., Tochner, Z., Hahn, S. M., & Glatstein, E. (2011). Short-Term and Long-Term Health Risks of Nuclear-Power-Plant Accidents. *New England Journal of Medicine*, 364(24), 2334–2341. <https://doi.org/10.1056/nejmra1103676>) In the last 50 years, research in nuclear energy and its possibilities have drastically gone down, due to accidents such as Chernobyl and Fukushima. The L.F.T.R. has a few inherent safety features, in comparison to its volatile brother the L.W.R.. The first feature being the pressure that it operates at. Light water reactors, since they use water as the coolant and need to boil it to turn the turbine, are required to work at a high pressure as well to increase efficiency. This makes the reactor much more unstable since it is

working in this fixed volume at high temperature and high pressures making a chance of explosion much higher(energy increase inside reactor itself). In contrast, the L.F.T.R. operates at atmospheric pressure. The lower pressure of the L.F.T.R. not only means lowered chance of explosion, but also less containment material. ((Wikipedia Contributors. (2019, June 19). *Liquid fluoride thorium reactor*. Wikipedia; Wikimedia Foundation.

https://en.wikipedia.org/wiki/Liquid_fluoride_thorium_reactor))

The second part of its safety is the fact that the fuel comes in a liquid form. L.F.T.R. has a negative temperature coefficient of reactivity, a tendency to move towards subcriticality. Since the L.F.T.R. uses molten salt instead of water, it doesn't require any cooling or control rod system since the salt will just expand as the temperature of the core increases. As the molten salt expands, the fuel is spread throughout it and this causes the fission reaction to slow down. Therefore, in the unlikely event that the reactor begins to overheat the molten salt will compensate by expanding and reducing the nuclear reaction rate, preventing a meltdown scenario.((Wikipedia Contributors. (2019, June 19). *Liquid fluoride thorium reactor*. Wikipedia; Wikimedia Foundation. https://en.wikipedia.org/wiki/Liquid_fluoride_thorium_reactor))

Fossil Fuels Vs Nuclear Energy - 2A IV

Currently, about 80% of our energy supply is made from fossil fuels. ((Denchak, M. (2022, June 1). *Fossil Fuels: The Dirty Facts*. Natural Resources Defense Council. <https://www.nrdc.org/stories/fossil-fuels-dirty-facts>)) As discussed earlier in the paper, fossil fuels are the main source of global warming causing further issues such as rising sea levels, droughts, loss of biodiversity, drought, and crop loss. In order to prevent these issues from furthering, the usage of fossil fuels needs to be halted. This would require a gradual move away from fossil fuels, which is where nuclear energy comes into play. Nuclear energy is a very low carbon emission, and in comparison to fossil fuels and renewable energy is able to lower carbon emissions both long and short term. Not only this, it was shown to be more energy efficient than fossil fuels, being around 8000 more efficient. From all this, it can be concluded that nuclear power is a superior choice to fossil fuels, but how to shift towards it? If fossil fuels make up 80% of the world's current energy supply, to shift to this suddenly would cause a collapse in the world. This would be a gradual process, taking about 30 years. ((Biello, D. (n.d.). *The World Really Could Go Nuclear*. Scientific American. <https://www.scientificamerican.com/article/the-world-really-could-go-nuclear/#:~:text=And%20if%20the%20world%20as>))

Nuclear Engine for interstellar travels - 2B

Nuclear science, although may be extremely well researched, has very few applications because of disasters such as chernobyl and fukushima. Things such as the cost of reactors have all been a factor in why it's been left alone recently. Civilization fully comprehending the depth of nuclear power ranging from nuclear fusion to the dyson sphere will enhance our ability to advance as a society. Other more reachable technologies currently, such as nuclear engines could

be the key to powering humanity exploring our solar system. ((McLAFFERTY, G. H. (1970). Gas-core nuclear rocket engine technology status. *Journal of Spacecraft and Rockets*, 7(12), 1391–1396. <https://doi.org/10.2514/3.30179>)) These engines will aid humanity in achieving higher degrees of technological nuclear advancement such as nuclear fusion based designs, marking the peak of human energy harnessing ability.

The Nuclear Engine - 2B I

The nuclear engine is essentially a very small nuclear reactor. The nuclear engine is more of a large-scale technology, with the main use for this paper to be to power human exploration by fueling rockets for long term missions. ((McLAFFERTY, G. H. (1970). Gas-core nuclear rocket engine technology status. *Journal of Spacecraft and Rockets*, 7(12), 1391–1396. <https://doi.org/10.2514/3.30179>))

Chemical Rocket Engine

First, a basic understanding of the regular common chemical propulsion system. For all rockets, the essential understanding is to create a thrust under the rocket causing it to rise with a great enough speed to escape earth's gravitational pull. The chemical engine of the rocket creates the required thrust by combusting a fuel using an oxidizer (together to make propellant), and this exhaust gas from the reaction is spewed out of the rocket nozzle to operate as the thrust. This gas is referred to as the working fluid. The two types of chemical propulsion systems are liquid and solid. As they suggest, liquid is where the fuel and oxidizer are stored separately and then pumped into the burning chamber where the reaction occurs. In a solid system, the propellant is already one and doesn't require any mixing. Naturally, rockets are not so simplistic but the basic understanding of the engine is sufficient for this paper. ((Price, E. W., & Biblarz, O. (2024, July 19). *Rocket - Chemical rockets*. Encyclopedia Britannica. <https://www.britannica.com/technology/rocket-jet-propulsion-device-and-vehicle/Chemical-rockets>))

Nuclear engine use in the real world

Nuclear engines have already shown success in areas such as use for submarines and aircraft carriers, so the concept is proven already. ((US EPA, O. (2018, November 30). *Nuclear Submarines and Aircraft Carriers*. US EPA. <https://www.epa.gov/radtown/nuclear-submarines-and-aircraft-carriers#:~:text=Nuclear%20submarines%20and%20aircraft%20carriers%20are%20powered%20by%20onboard%20nuclear>)) The nuclear engine shows pros such as lowering of greenhouse gas emissions, no refueling, and even higher speeds. ((*Nuclear propulsion systems* | ANSTO. (n.d.). www.ansto.gov.au. <https://www.ansto.gov.au/our-science/nuclear-technologies/reactor-systems/nuclear-propulsion-systems#:~:text=Nuclear%20propulsion%20provides%20several%20advantages>)) These pros (specifically no refueling), are extremely helpful when it comes to interstellar travel. As

discussed before, a chemical engine's journey to Mars would require refueling midair, which would be difficult to calculate and dangerous for the people inside the actual rocket.

Nuclear Thermal Propulsion(N.T.P.)

The first type of engine too using nuclear energy, is the nuclear thermal propulsion engine. Essentially, a low mass fuel such as hydrogen flows into the reactor starting the chemical reaction. which can generate heat from the nuclear fission causing the thrust. Hydrogen is the desired particle since the lower the mass the higher the specific impulse/rocket efficiency.

Hydrogen would need to be stored cryogenically for it to be stored in a liquid form.((Gabrielli, R. A., & Herdrich, G. (2015). Review of Nuclear Thermal Propulsion Systems. *Progress in Aerospace Sciences*, 79, 92–113. <https://doi.org/10.1016/j.paerosci.2015.09.001>)) For the fuel itself, the isotope uranium 235 will come into play, as we saved it before by using thorium for ground reactors. Like a regular reaction, the U235 is hit with a neutron causing a chain reaction inside the fuel. More neutrons break off the U235 atoms, and more of them are hit by these broken off neutrons. ((. Massachusetts Institute of Technology. (2021). *The Fission Process | MIT Nuclear Reactor Laboratory*. Mit.edu.

<https://nrl.mit.edu/reactor/fission-process#:~:text=When%20a%20U%2D235%20nucleus>)) This reaction creates thermal energy, heating the hydrogen fuel. This creates hydrogen gas, coming out the nozzle extremely fast creating a thrust.((Gabrielli, R. A., & Herdrich, G. (2015).

Review of Nuclear Thermal Propulsion Systems. *Progress in Aerospace Sciences*, 79, 92–113. <https://doi.org/10.1016/j.paerosci.2015.09.001>)) This type of engine could reduce travel time to 25% of the original time, has a rocket efficiency of around 2x more, and also has a higher energy density meaning heavier payloads can be delivered. ((Office of Nuclear Energy. (2020, January 21). *6 Things You Should Know About Nuclear Thermal Propulsion*. Energy.gov.

<https://www.energy.gov/ne/articles/6-things-you-should-know-about-nuclear-thermal-propulsion>))

Nuclear Electric Propulsion System(N.E.P.)

The second engine discussed using nuclear energy is the nuclear electric propulsion system. This is essentially nuclear energy used to power electrical spacecraft propulsion. An electric spacecraft propulsion system is one in which electrostatic or electromagnetic fields are used in order to accelerate mass at high speeds in order to achieve thrust. The fluid being used to thrust the rocket (gas, ion, plasma) is accelerated via an electrical field or magnetic field. In specific to this paper, we will discuss ion propulsion as it is most efficient. First, a heavy propellant that is generally a noble gas, such as Xenon is picked to be the propellant. The propellant is then bombarded with electrons giving the propellant atom a positive charge. The overall charge of this mixture still has zero net charge. This produces plasma, which has similar properties to a gas except that it is affected by electrical and magnetic fields. These new positively charged ions then flow towards the ion thruster which consists of two grids. The first grid being a screen grid, holding a positive charge but is configured in a way such that it will force the plasma through the tiny holes of the grid. On the other side, there is a negatively charged screen, known as the accelerator grid. This grid attracts the plasma, pulling it towards it

at velocities of ninety thousand miles per hour being thrust out of the rocket as an ion beam. For a nuclear electric rocket, the electrical energy to charge the accelerator grid for thrust is derived from the N.E.P.. The nuclear electrical engine has the reactor create thermal energy the same as the N.T.P.. But instead of heating the fuel, this energy is then used in an electrical generator from the H2 steam turning the turbine. The electrical energy then is used in the acceleration grid in order to accelerate the ions. ((Office of Nuclear Energy. (2020, January 21). *6 Things You Should Know About Nuclear Thermal Propulsion*. Energy.gov. <https://www.energy.gov/ne/articles/6-things-you-should-know-about-nuclear-thermal-propulsion>))

Pros of the nuclear engine

The further human exploration gets, the more fuel we need, which in turn weighs the rocket down more and more. The first human base outside of earth should be established on Mars(discussed later), which takes around one thousand to four thousand metric tons of fuel. This means the engine would either need to be powerful enough to carry this much fuel which is essentially impossible, or need to be refueled around 4 times. Alongside being extremely difficult to pull this mission off, a rocket engine creates black carbon as result of the burning of the hydrocarbon inside the rocket engine. Black carbon is even worse than regular carbon, essentially being a stronger version of carbon doing all the things carbon does but at a faster rate. In general, regular rocket engines are unappealing due to their inefficiency, lower payload weight carrying ability, and longer travel times in comparison to N.T.P.'s for long term missions like the Mars journey. ((Office of Nuclear Energy. (2020, January 21). *6 Things You Should Know About Nuclear Thermal Propulsion*. Energy.gov. <https://www.energy.gov/ne/articles/6-things-you-should-know-about-nuclear-thermal-propulsion>))

The road to the Type 2 civilization - 3

In the previous section, we discussed how to harness nuclear energy as the main source for humanity in order to exponentially increase growth on the Kardashev scale in order to escape the Great Filter. This next section will outline how to become an interstellar civilization and what paths can be taken in order to harness the energy within our solar system through technology like the dyson sphere, the stellar engine, terraforming, and even the penrose mechanism.

Start of exploration - 3A

For humans to become an interstellar civilization, beginning civilizations must be created. This means humans must find a place outside of earth that fulfills the three basic components for human survival: food, water, and air.((Medicine, C. for F. (2019, September 19). *4 Things Your Body Needs To Survive*. Center for Family Medicine. <https://centerforfamilymedicine.com/general-health/4-things-your-body-needs-to-survive/>)) Out of all the planets within our solar system, the only planet fulfilling these requirements is Mars.((Shing, B., & Yim, Y. (n.d.). *Why Mars?* Retrieved July 20, 2024, from <https://marspapers.org/Essays/E19.pdf>))

Establishing a moon base - 3A I

Before colonizing Mars, a hub in which humans can travel from and forth is necessary. This hub would function as a lab for humans to collect important information that could be important in our travels to Mars, similar to the international space station. A lunar base is more so the beginning step to further space exploration such as Mars colonization, as a lunar base is much simpler and would lay the basic groundwork down for interstellar colonization. This would also help humanity understand how the human body reacts to space like atmosphere, such as the low gravity of the moon or the extreme colds. All in all, a lunar base is essential to eventually colonizing Mars. ((Dinkin, S. (2004, September 7). *The Space Review: Colonize the Moon before Mars*. [www.thespacereview.com](https://www.thespacereview.com/article/221/1). <https://www.thespacereview.com/article/221/1>))

For further understanding, a parallel of humanity past colonization methods can be made. For nautical expansion, humans would scope the land out, and do research. They would check factors such as soil fertility, types of biological life, and more. They would come back to their mother country with all the research they acquired, and show it to them. This first step is considered the first phase. Then, a group of people would be sent to begin the outpost. Once the outpost was established, the land was further researched and from here more colonization could occur. This is the second phase. Lastly, more humans were sent over, but now instead of for researching purposes it was to establish a society. Traders, merchants, children, families, governmental figures, and other necessary people were sent. Slowly, this society starts to grow, and the process repeats itself. This is the third and final phase. ((American History Central Staff. (2024, May 27). *Colonial America, Facts, Phases, Characteristics, APUSH*. American History Central. <https://www.americanhistorycentral.com/entries/colonial-america/>)) This same method can be applied to interstellar colonization, with the moon serving as our first colonization and the mars colony stemming off of this.

How to create a Lunar Base

In terms of the first phase, this has already been completed. Humanity has proven that the moon can be reached, and we've researched the geography and composition surface of the moon through rovers.

Here, we enter the second phase. This is where humans will send astronauts in small crews, to establish the first moonbase. This moonbase would consist of technicians, engineers, and other stem based individuals. ((American History Central Staff. (2024, May 27). *Colonial America, Facts, Phases, Characteristics, APUSH*. American History Central. <https://www.americanhistorycentral.com/entries/colonial-america/>)) These bases are short term, just to prove humans can live on the moon and see the effects on the humans from living here. Since there are no energy sources on the moon, solar energy will be important. These solar panels should be employed near the south pole, as it is well illuminated and will be the best location to generate the maximum amount of energy. ((. Ross, A. K., Ruppert, S., Gläser, P., & Elvis, M. (2023). Preliminary quantification of the available solar power near the lunar South Pole. *Acta Astronautica*, 211, 616–630. <https://doi.org/10.1016/j.actaastro.2023.06.040>)) While

the scientists reside on the Moon, they will further research the composition of the moon and research how to use the moon's material to create the first interstellar longterm outpost. The harnessing of lunar ice will be essential in reaching phase three. Lunar ice is not only the key to drinking water ((Villaver, E. (2023, September 5). *Water on the Moon and the civilizations of the future*. EL PAÍS English.

<https://english.elpais.com/science-tech/2023-09-04/water-on-the-moon-and-the-civilizations-of-the-future.html#>)), but also crop farming, hydrogen powered energy, and rocket fuel ((Kuthunur, S. (2023, November 22). *Scientists find hydrogen in Apollo moon rocks, suggesting astronauts can harvest lunar water*. Space.com.

<https://www.space.com/hydrogen-moon-rocks-apollo-astronauts-samples#:~:text=A%20fresh%20analysis%20of%20moon>)). Due to the moon's low gravity, it requires less fuel to get things out of the moon's gravitational field making missions back to earth and missions to Mars cheaper.

In the case of the Moon, it will never enter its third phase. The moon's base function is to act as a hub and there is no practical reason for normal citizens to start living on the moon. With the innovation of the nuclear engines, travel from earth to Mars is feasible in one shot, so only people who are directly involved in the interstellar colonization of our solar system would need to be on the moon. This demographic consists of key individuals focused on building and supporting the base.

The moon has a few cons in comparison to Mars. Although the moon as a hub to Mars would work well, the moon as the main interstellar colony is unviable. To begin, Mars shows more proof to have more possibility of ice for drinking, alongside other rich material. Mars has a rich atmosphere as well, filled with necessary biological chemical components like hydrogen, nitrogen, oxygen and carbon. The amount of these on the moon are little to none. ((Zubrin, R. (n.d.). *The Case for Colonizing Mars, by Robert Zubrin – NSS*. Nss,Org.

<https://nss.org/the-case-for-colonizing-mars-by-robert-zubrin/#:~:text=In%20contrast%20to%20the%20Moon>)) Alongside this, Mars has terraforming potential. While terraforming will be discussed later, the basic concept is that a planet can be transformed into conditions similar to earth. Terraforming the moon is not an option, while Mars can be terraformed.((*Terraforming*. (2024, July 3). Wikipedia.

<https://en.wikipedia.org/wiki/Terraforming#:~:text=While%20Venus%2C%20Earth%2C%20Mars%2C>))

Road to Mars - 3B

At this point of time, humanity would have achieved its first space colony on the moon. This colony on Mars would consist of contractors looking for resources, scientists researching the moon, and engineers working on technology to make the Moon a self-sufficient hub. At this point, the same three phases can be applied to Mars with the Moon acting as the mother planet. The difference between the Mars colonization and the Moon colonization, however, is that colonizing Mars would end in being a true society as all of humanity could live here for a long time in contrast to the Moon's limited demographic. People would be shipped directly to Mars

using the nuclear engines discussed earlier, and humans would have a new home in the event Earth fails us. The colonization of Mars would also be the markings of technological advancements on the Dyson sphere through the colonization of Venus, the Penrose mechanism, and the stellar engine.

Phase one: Scouting - 3B I

The first step to colonization as discussed before, is researching the land. ((American History Central Staff. (2024, May 27). *Colonial America, Facts, Phases, Characteristics, APUSH*. American History Central.

<https://www.americanhistorycentral.com/entries/colonial-america/>) From the moon, using fuel artificially produced by splitting water into hydrogen and oxygen, rovers and satellites can be created. A key theme throughout the path of Mars exploration will be the favorability of automation over humans. ((Campa, R., Szocik, K., & Braddock, M. (2019). Why space colonization will be fully automated. *Technological Forecasting and Social Change, 143*, 162–171. <https://doi.org/10.1016/j.techfore.2019.03.021>))

Past life and habitable environment

These automated technologies will investigate possible past life and research habitable environments. It would mark these spots, and relay this information back to the moon hub and earth. The reason signs of life are searched for, is that if one civilization could do it then this proves the concept of living on Mars. ((Weintraub, D. (2018, September 6). *Consider the Martians: Scientists need to confirm whether life exists on Mars before sending humans*. Vanderbilt University.

<https://news.vanderbilt.edu/2018/09/06/consider-the-martians-scientists-need-to-confirm-whether-life-exists-on-mars-before-sending-humans/>) In terms of habitat, due to the weaker atmosphere and lack of radiation, places with overhead shielding will function best.

Oxygen Production on Mars

Alongside this, the automation will also test out how to generate oxygen. NASA has a technology known as the Mars Oxygen In-situ Resource Utilization Experiment (MOXIE), which can produce oxygen in the Martian atmosphere. ((Hoffman, J. A., Hecht, M. H., Rapp, D., Hartvigsen, J. J., SooHoo, J. G., Aboobaker, A. M., McClean, J. B., Liu, A. M., Hinterman, E. D., Nasr, M., Hariharan, S., Horn, K. J., Meyen, F. E., Okkels, H., Steen, P., Elangovan, S., Graves, C. R., Khopkar, P., Madsen, M. B., & Voeks, G. E. (2022). Mars Oxygen ISRU Experiment (MOXIE)—Preparing for human Mars exploration. *Science Advances, 8*(35). <https://doi.org/10.1126/sciadv.abp8636>) Human survival requires pure oxygen gas, which Mars has a lack of. On the other hand, Mars has around a 95% carbon dioxide atmosphere. The MOXIE takes the carbon dioxide from the martian environment, and splits it into carbon monoxide and oxygen gas. First, the carbon dioxide gas is pulled from the atmosphere and pressurized to half an atmosphere. Then a solid oxide electrolyzer (SOXE) absorbs the carbon

dioxide. The point of the SOXE is to split the carbon dioxide, and create pure oxygen. This oxygen is then measured for purity and if it passes is passed through the exhaust along with the byproducts of this process. ((Hoffman, J. A., Hecht, M. H., Rapp, D., Hartvigsen, J. J., SooHoo, J. G., Aboobaker, A. M., McClean, J. B., Liu, A. M., Hinterman, E. D., Nasr, M., Hariharan, S., Horn, K. J., Meyen, F. E., Okkels, H., Steen, P., Elangovan, S., Graves, C. R., Khopkar, P., Madsen, M. B., & Voecks, G. E. (2022). Mars Oxygen ISRU Experiment (MOXIE)—Preparing for human Mars exploration. *Science Advances*, 8(35). <https://doi.org/10.1126/sciadv.abp8636>))

Martian Ice: Water and further

Mars has around 5 million cubic kilometers worth of ice. Water from Mars would be most efficiently harvested through mining.((Arizona State University. (n.d.). *Mars Education | Developing the Next Generation of Explorers*. Marsed.asu.edu. <https://marsed.asu.edu/mep/ice#:~:text=More%20than%205%20million%20cubic>)) Rovers that specifically work in mining would be sent. Although there is a lot of water to be harvested, the water is more than likely to be contaminated with perchlorate. ((Rothschild, L. (2024, January 4). *Detoxifying Mars: the biocatalytic elimination of omnipresent perchlorates* - NASA. Nasa.gov. <https://www.nasa.gov/general/detoxifying-mars/#:~:text=Upon%20arrival%20at%20Mars%2C%20spores>)) This means this water would need to be desalinated before consumption can occur, which humanity has already achieved. Since the concentration of perchlorate is likely stronger on Mars in the water than earth, the energy required to desalinate it is greater. Ion exchange seems to be the most obvious answer. Ion exchange is a process in which the water is de-perchlorate by replacing the perchlorate with an ion of the same charge which is harmless. Then, reverse osmosis is used. Pressure is built up on the water and perchlorate which are now separated, pushing the water through and leaving the perchlorate separated from it. ((Anthony', S., Jolle, S., & Captain, J. (n.d.). *Advanced Water Purification System for In Situ Resource Utilization*. Retrieved August 5, 2023, from <https://ntrs.nasa.gov/api/citations/20140002803/downloads/20140002803.pdf>)) Although this method is efficient, Mars would become dirty from the perchlorate left over. A better, more viable solution is to use bacteria. This bacteria would consume the perchlorate in the water leaving it clean. These types of bacteria are not new, but aren't suitable for the Mars environment. We would need to synthetically produce these same bacteria, but instead to be adapted to the Mars environment. ((Szondy, D. (2024, January 9). *NASA may use bacteria to make Martian water drinkable*. New Atlas. <https://newatlas.com/space/nasa-bacteria-drinkable-martian/>)) Water has further use though, than just for drinking. The splitting of water can be used to produce hydrogen fuel. Not only can this power the colony later in the second phase, but also creates oxygen for breathing purposes again. ((Romero-Calvo, A. (2024, January 4). *Magnetohydrodynamic Drive for Hydrogen and Oxygen Production in Mars Transfer* - NASA. Nasa.gov. <https://www.nasa.gov/general/magnetohydrodynamic-drive-for-hydrogen-and-oxygen-production/>))

Phase 2: Establishing base and the first group 3B II

At this time, humans through the use of rovers will have completed the scouting. The ability to use water for methods such as space farming, drinking and purification, rocket propellant fuel, and even oxygen on Mars through splitting will have been tested. Essentially, Mars' potential to be a habitable planet for humans, will have been established. Assuming that Mars has been determined to be colonizable, the process of sending humans will begin. The importance of the first group is that they will set precedent for those following them. They will conduct important research, setting up the stage for the colonization of Mars as a society. ((American History Central Staff. (2024, May 27). *Colonial America, Facts, Phases, Characteristics, APUSH*. American History Central. <https://www.americanhistorycentral.com/entries/colonial-america/>)) Things such as biodiversity, terraforming ability, ecosystem tendencies, and the geography will be checked by humans now, rather than rovers. The first group's job is to delve a little into all of this, making a deeper understanding on how humans can make this planet possible to live on by making it as similar to natural earth as feasible. A main key to remember throughout this entire section, is that due to how small this first group is, the humans will adapt to the habitat in which they live. This concept will not apply, when a full society is established on Mars, as these humans will transform the land around them to their own needs.

Geography

Although the geography of Mars may be similar to Earths, it is still unique compared to any other planet. Similar to our earth it is a terrestrial planet: core, mantle, and crust build. ((NASA's *InSight Reveals the Deep Interior of Mars* - NASA. (2021, July 22). Nasa.gov. <https://www.nasa.gov/centers-and-facilities/jpl/nasas-insight-reveals-the-deep-interior-of-mars/#:~:text=Like%20Earth%2C%20Mars%20heated%20up>)) Alongside this, Mars also holds water streams, volcanoes, and winds. However, while earth may have become a biological hub, Mars most likely is a cold and dead environment. ((NASA. (2023). *Mars: Facts - NASA Science*. Science.nasa.gov; NASA. <https://science.nasa.gov/mars/facts/>)) While they may seem a bad thing at firsthand, this only means Mars has potential to be turned into something similar to earth. Humans will still be looking at the three main components for life: food, breathable air, and water, ((Medicine, C. for F. (2019, September 19). *4 Things Your Body Needs To Survive*. Center for Family Medicine. <https://centerforfamilymedicine.com/general-health/4-things-your-body-needs-to-survive/>)), but at the same time will be looking at shelters to establish a long term base. Geographically wise, the most affected component here is shelter. Due to the rocky terrain with large volcanoes, a toxic Martian soil filled with perchlorates, high radiation levels, alongside low temperatures and pressures, bending all of these rules to our fitting would prove to be difficult. Such a task as terraforming, would only be applicable to a level 3 civilization with access to the entirety of their galaxy's energy.

Temperature

In regards to where to establish a base, the most desirable place would be the Valles Marineris, or the “Grand Canyon of Mars.” It is along the Martian equator, stretching 4000 kilometers and around 8 kilometers deep. This spot provides a few key benefits. Firstly, the temperature. With temperatures reaching a high of 70 degrees Fahrenheit in this valley, this is much higher in comparison to the usual temperature of around -64 degrees Fahrenheit. ((Lavars, N. (2016, September 30). *Three great places to live on Mars*. Newatlas.com; New Atlas. <https://newatlas.com/great-places-to-live-mars/45654/>))

Water

Alongside this, evidence points to the fact that there is groundwater in this habitat. This easy access to water will provide many benefits, such as easy drinking water and making space farms more feasible. Not only will it make life easier, but in such a warm and wet place in comparison to the rest of Mars barren land, this could possibly be a hub for Mars microbial activity. Research on possibilities for biosignatures and biodiversity would have the greatest chance of success in this canyon. ((Lavars, N. (2016, September 30). *Three great places to live on Mars*. Newatlas.com; New Atlas. <https://newatlas.com/great-places-to-live-mars/45654/>))

Pressure

Also, the low altitude of the canyon aids with pressure. One of the greatest flaws of Mars' potential is its lack of pressure, which is around one twenty fifth of earths. This lack of pressure connects to effects such as human blood boiling. Gasses in the blood that are usually dissolved, would suddenly start boiling since the lack of pressure is encouraging an inverse increase in volume. This would be fatal in a matter of seconds. ((Klotz, I. (2017, May 11). *Boiling Blood and Radiation: 5 Ways Mars Can Kill*. Space.com. <https://www.space.com/36800-five-ways-to-die-on-mars.html>)) Therefore, technology to solve the pressure differential between Mars and Earth must be utilized. At the moment, since it's only a small group of people researching Mars, the pressure problem can be solved by using a pressurized suit. This, unlike the previously bulky spacesuits used on the moon, is a skin tight suit wrapped with multiple muscle-like coils. These coils would form around the body of the user. This not only pressurizes the user, but also makes it easier for the astronaut to move. In a normal suit, the low pressure differential is usually fixed with gas inside the suit applying external force on the person inside. In the pressurized suit although, the suit applies the extra needed pressure on the human by tightening around them. This not only makes it more mobile for the suit user, but this leaves more room for oxygen to be carried by the user enabling longer trips on Mars. ((Chu, J. (2014, September 18). *Shrink-wrapping spacesuits*. MIT News | Massachusetts Institute of Technology. <https://news.mit.edu/2014/second-skin-spacesuits-0918>)) For the cabins themselves, using oxygen gas from the MOXIE, ((Hoffman, J. A., Hecht, M. H., Rapp, D., Hartvigsen, J. J., SooHoo, J. G., Aboobaker, A. M., McClean, J. B., Liu, A. M., Hinterman, E. D., Nasr, M., Hariharan, S., Horn, K. J., Meyen, F. E., Okkels, H., Steen, P., Elangovan, S., Graves, C. R., Khopkar, P., Madsen, M. B., & Voeks, G. E. (2022). Mars Oxygen

ISRU Experiment (MOXIE)—Preparing for human Mars exploration. *Science Advances*, 8(35). <https://doi.org/10.1126/sciadv.abp8636>) these cabins can be pressurized similar to how an airplane is pressurized. With the higher pressure of Valles Marineris in comparison to the rest of Mars, in addition to the pressurized suit, the pressure differential can theoretically be solved.

Radiation

The low altitude doesn't only help with pressure though, it also helps with radiation. Mars, due to its weak atmosphere, is very susceptible to cosmic radiation. Cosmic radiation has symptoms such as altering the cardiovascular system, heart damage, arteries hardening, cataracts, increased chances of cancer, and even sterility problems. ((Blanchett, A. (2017, September 19). *Space Radiation is Risky Business for the Human Body - NASA*. Nasa.gov.

<https://www.nasa.gov/humans-in-space/space-radiation-is-risky-business-for-the-human-body/>)

The main problem of cosmic radiation is the rate at which it works. Effects of cosmic radiation may cause gene mutations, but it is not necessary that this mutation shows up immediately. The affected offspring may end up with these symptoms instead, which is much worse, since the entirety of the posterity could be ruined, destroying the very thing we are working for.

((Blanchett, A. (2017, September 19). *Space Radiation is Risky Business for the Human Body - NASA*. Nasa.gov.

<https://www.nasa.gov/humans-in-space/space-radiation-is-risky-business-for-the-human-body/>)

Therefore, technology preventing cosmic radiation would be necessary. The obvious answer at first, is use the method passive shielding. Passive shielding is where a massive piece of shielding metal is placed between the group habitat, and the sky. Although this may seem desirable due to its ease at first, placing a massive piece of metal over the society as it begins to grow is unrealistic.((Case, D. ((2016, July 19). *Passive Radiation Shielding: Integrating Multilayer and Multipurpose Materials into Space Habitat Design - NASA*. Nasa.gov.

<https://www.nasa.gov/directorates/stmd/space-tech-research-grants/passive-radiation-shielding-integrating-multilayer-and-multipurpose-materials-into-space-habitat-design/#:~:text=One%20radiation%20protection%20method%20is>) A more desirable method is the AstroRad.((*AstroRad - What You Wear in Space Could Save Your Life*. ((n.d.). Lockheed Martin.

<https://www.lockheedmartin.com/en-us/products/astrorad---what-you-wear-in-space-could-save-your-life.html>) Since it's only a small group, the few members can be equipped with this radiation protecting vest. Similar to the pressure problem, through the use of technology helping with shielding, in addition to the natural protection of the low altitude leading to a thicker atmosphere, cosmic radiation can be prevented from affecting humans.

Supplying

Nuclear Energy

At this point in the paper, the theme of nuclear energy being the key to human growth seems pretty constant. This theme doesn't change, even if the planet has. Although nuclear fusion

falls more into the type III civilization level technology, at no point is nuclear fission restricted. Energy is necessary for life, and this energy will power the space technologies previously mentioned alongside with other things like space farming and housing. L.F.T.R's will begin to come into play. Luckily, there is in fact thorium on Mars meaning there is no need for constant supply from earth. ((Brandenburg, J. E. (n.d.). *EVIDENCE FOR A LARGE, NATURAL, PALEO-NUCLEAR REACTOR ON MARS*.

<https://www.lpi.usra.edu/meetings/lpsc2011/pdf/1097.pdf>) Other than the containment unit and mechanism itself, the fuel can be created on Mars. For the lithium beryllium, we'd need to separate the lithium-6 isotope from the lithium-7 isotope. This is where COLEX, or column exchange comes into play.

Since Lithium 6 has a greater affinity towards Mercury compared to Lithium 7, when combining a mixture of mercury and lithium with aqueous lithium hydroxide, the lithium 6 is more concentrated in mercury than before replacing the previous lithium 7, which moves to the hydroxide side. The lithium 6 and then mercury are then separated, sending the mercury to be reused and extracting the lithium 6. ((Giegerich, T., Battes, K., Schwenzer, J. C., & Day, C. (2019). Development of a viable route for lithium-6 supply of DEMO and future fusion power plants. *Fusion Engineering and Design*, 149, 111339.

<https://doi.org/10.1016/j.fusengdes.2019.111339>) From there, beryllium and fluoride can be shipped from Earth in large quantities, and since the LFTR doesn't need constant refilling, it will last a substantial amount till needing another shipment. Following the same rules as stated in previous sections, an L.F.T.R. can be constructed powering Mars till further technologies and a sufficient energy can be obtained.

Farming

Throughout the course of history, the moment a colony is self-sufficient it goes from just a colony to a full fledged society. Looking back at nautical expansion times, one of the colonies first goals was to establish farming in these colonies. ((Saenger, T. (2013). *Colonial Farming and Food: Famine to Prosperity* | *NCpedia*. www.ncpedia.org.

<https://www.ncpedia.org/colonial-farming-and-food-famine>) Similar to how nautical expansion referred to America as the New world, humans can consider Mars their New world. Therefore, we can apply similar principles of colonization learned from nautical expansion in the 1500's-1700's, to space expansion. Revolving back to the main point of this section, farming is important in order to ensure when a society is established, it will only need occasional transports from the mother country. Crops require 2 main things: water and sunlight. In order to create the ideal environment for farming, it is necessary to simulate earth like conditions. The most efficient method is expanding a current technology known as the Nutritional Closed-Loop Eco-Unit System, or the NUCLEUS. This has been able to create fresh microgreens, mushrooms, vegetables, and even insects that are edible. Within the capsules, the plants are grown similar to how vertical farming is done. The reason for this is since vertical farming is efficient at reducing the amount of water and fertilizer needed, and doesn't require soil. ((*Vertical farming could be*

the solution to growing food on Mars. (2023, March 24). CNN.

<https://cnn.com/2023/03/24/world/gallery/agriculture-on-mars-interstellar-lab-spc-scen-intl/index.html>) Both the concepts of the NUCLEUS and vertical farming, can be extrapolated on a scale suitable for Mars. Since vertical farming is indoors, it isn't dependent on temperature which is key for such a cold planet like Mars. Water from the Valles Marineres would be extracted, with previous methods such as nuclear energy fueling the artificial light in the capsule. Instead of a small capsule though, it'd be more of a massive box, similar to that of a greenhouse. Unlike vertical farming on earth, hydroponic systems to raise water would be more troublesome due to the low gravity. An alternative is aeroponics. This delivers the nutrient rich water on a large scale as a mist, enriching the plants. Mars is also already rich in CO₂, so plants would easily thrive by venting CO₂ into the system. By growing crops, humans receive an easy way to outsource oxygen, fresh food, and possibilities for medicines.

Dyson Sphere

The dyson sphere is the key mark of a true type 2 civilization. This means, this civilization is capable of gaining all the energy of its available solar system. The dyson sphere is a megastructure mechanism built around a star, which would absorb the star's energy. In this case, this would be the sun but an actual shell around the sun holds great problems such as freezing our solar system, so this isn't practical. Not only that, it would be easier for it to collapse when hit. A more practical solution is the dyson swarm. It is essentially a swarm of solar panels around the sun, taking in the energy of the sun and bringing it back to the home planet. (Wright, J. (2020). Dyson spheres. *Serbian Astronomical Journal*, 200, 1–18. <https://doi.org/10.2298/saj2000001w>)

The amount of material needed to create an entire megastructure is well past earth's resources, and would require the disassemblment of an entire planet. The most metal rich planet, which also happens to be closest to the sun, is Mercury. Alongside this, the low atmosphere plus the low gravity makes it easier to build from there. The basic idea is to make the solar panels on Mercury through automation from mined material, and send them out in a swarm around the sun. Firstly, beginning with the disassembly of the planet. Disassembling an entire planet would be a massive job, and would require amounts of energy comparable to the output of a dyson sphere itself, which is a conundrum. Therefore, Mercury will have solar panels absorbing the energy output from the sun. Since the planet is so close, the amount of energy absorbed by the solar panels is very high which is beneficial in speeding up the breakdown of mercury and its resources. This energy is then used to mine the planet, taking up tons of material which is later refined. This refined material is then made into solar panels. All this technology at the moment, is not out of human engineering capabilities, but the actual shipment of the solar panels is around the sun. To have to use a rocket every single time to bring a single solar panel would be inefficient and expensive, so it'd be much more valuable to shoot them into space, essentially a rail ejector. As these solar panels get shot up, they will be engineered to align themselves too parallel with the sun, and from there the process can be repeated. As more solar

panels are sent, the total energy output increases meaning the amount of solar panels that can be sent exponentially increases. ((Xu, V. (2022, June 9). *Should we destroy Mercury to create a Dyson sphere?* The Starfish Canada.

<https://thestarfish.ca/journal/2022/06/should-we-destroy-mercury-to-create-a-dyson-sphere>))

Concluding Phase 2:

The first group will have tested the ability of all stated above. The ability to successfully do space farming, use nuclear energy in space to sustain us, and adapting to the geography will have been completed. These humans will come back to Earth with their findings, and the process in which Mars will be fully colonized will begin. At this point, humans will have a choice. The first choice is that they continue to adapt to Mars, using technology such as advanced space suits. This will make it extremely difficult to colonize Mars as a whole society, and would be a very undesirable choice as it is unreasonable to have all humans adapt to every problem of Mars. A much more desirable option is using technology to alter Mars to human needs. This would require futuristic technology such as terraforming, the dyson sphere, the penrose mechanism, and nuclear fusion. Humans could recreate essentially a new earth, and this could be the mark of the beginnings of the phase 3 civilization as humanity gains all available energy within their galaxy.

Phase 3 - 3B III

At this point, human growth is hard to predict. The main essential of a phase 3 civilization, is that through advanced technology the galaxy is fully traversable to them, and all energy within this galaxy has been harnessed. Not only this, but humanity is in full motion of conquering the galaxy and spreading humanity across it.

Technology of Phase 3 civilization

Terraforming

A main key behind the idea of going across the galaxy and spreading humanity, is adapting planets to be habitable for humans. Terraforming, though, takes time and advanced technology. By recreating the specific scenario and requirements through which earth went onto another planet, ideally similar results should be yielded. For this section, we will use Mars as the main example. The usual method is to create a biosphere, then to shield it from outward interference and let time take its course. Not only is this time consuming, but it is also difficult to shield an entirely full planet. ((*Terraforming and the Future of Humans in Space*. (n.d.). <https://www.astro.umd.edu/~miller/teaching/astr380f09/lecture27.pdf>))

A more intelligent and cheaper planet, although radical, would be to melt Mars. At some point, Mars was oxygen rich and had a lot of water, but through the process of ultraviolet rays and solar winds, this planet became dry. This doesn't mean it is gone although, just buried away. There is still enough water on Mars for a shallow ocean, and the rocks of Mars are filled with oxygen and carbon dioxide trapped away. In order to free these gasses and unlock the water,

Mars needs to undergo thermolysis. ((*Exploring Space . Mars Terraforming Timeline . Mars as Home* | PBS. (n.d.). www.pbs.org.

<https://www.pbs.org/exploringspace/mars/terraforming/page7.html>)) Luckily, at this point humanity will have a dyson sphere or a black hole bomb or hopefully both, aiding in the path to terraform. Using a focused heat source such as a laser, powered by either one of the previously mentioned sources humanity could melt Mars surface unlocking these important components. This would lead to a similar start of what happened with earth, with Mars being extremely hot surrounded by thunder clouds. After thermolysis has taken place, the planet would cool down and the water would evaporate into clouds, raining down undesirable gasses such as chlorine away. Although the oxygen and water issue has been solved, it's still extremely hard to breathe due to such a low pressure. Due to this, we'd need nitrogen which Mars has a low amount of. This would call for an import of nitrogen, ideally from a nitrogen rich planet such as Venus. ((*Atmosphere of Venus*. (2024, July 17). Wikipedia.

https://en.wikipedia.org/wiki/Atmosphere_of_Venus#:~:text=Venus)) Similar to how the dyson sphere had solar panels shot from a rail jet, the imports of nitrogen would follow suit being launched into Mars's atmosphere.

After the atmosphere and pressure has been established, assuming there is already no life on Mars a biosphere must be created. By breaking the rocks down into sand again using thermolysis, and having the water combine with this, the basis for life on ground is created. Then, nitrogen consuming bacteria would be placed into the mud, which would consume this nitrogen and convert it into nitrates. These nitrates are essential for the start of planting. Slowly, this one enriched patch would grow into the rest of the barren land and soon the entirety of Mars would be covered in greenery. As more plants, animals, and insects are introduced the ecosystem will begin to become self-sustaining. ((Lutz, K. (2021). *Sustainable Nitrogen Cycle for Mars Settlement: Operations Analysis, Agriculture, and Soil Ecosystem*.)

At this point, Mars may be similar to Earth visually but it is still missing a magnetic field preventing cosmic radiation. This is where a superconductor could come into play, which would always be in line with the Sun and Mars and would be far ahead. This superconducting ring could be powered by nuclear fusion or energy from the sources of the dyson sphere, which would alter the ways away from Mars.

BLACK HOLES

Penrose mechanism

The penrose mechanism is a way in which we can harness energy from black holes, stemming from the fact that black holes can be spinning. The penrose mechanism came from basic physics, the idea of angular momentum specifically. Angular momentum is equal to the product given when mass is multiplied by angular velocity. The average angular velocity of a spinning black hole is around 90% of the speed of light, with speeds reaching that of 63 million miles per hour. Taking advantage of its high rotational kinetic energy, we can accelerate

spaceships for usually long missions, shortening the total travel time. First, a rocket is sent into the ergosphere of the blackhole. This rocket then drops some mass, having this mass be absorbed by the black hole in forms of energy. When the black hole absorbs this energy from the rocket's mass that was dropped, due to energy conservation laws the black hole energy in forms of mass will go up but the rotational kinetic energy will go down, slowing down the rocket. This loss in rotational kinetic energy will be accounted for by the spaceship, speeding up the spaceship in the form of kinetic energy. This tradeoff is extremely beneficial, due to the extremely high mass of the black hole and the low mass of the rocket. The black hole is basically unaffected by such a minimal mass, but since the rocket is so small this small fraction of the black hole energy transferred to the rocket is an exponential boost compared to before, helping it escape the ergosphere and leave it with more speed than it entered with. A desirable choice of this is using asteroids. A rocket that could essentially bring asteroids with them, and drop them into the black hole in return for its kinetic energy would be able to travel multifold than compared to without this mechanism. ((Tursunov, A., & Dadhich, N. (2019). Fifty Years of Energy Extraction from Rotating Black Hole: Revisiting Magnetic Penrose Process. *Universe*, 5(5), 125. <https://doi.org/10.3390/universe5050125>))

Black hole bomb

Although its name may be misleading, it may be the only way to a type 2 civilization, creating a near infinite energy source. The essential megastructure is a rigid body of multiple mirrors surrounding the black hole. In comparison to the dyson sphere, this would actually be easier to construct since the dyson energy output of available and black holes tends to be more compact than the stars. The building of this megastructure would follow similar to that of the dyson sphere, but in this case for a black hole the mass of the sun, this would be fully covered and would only require around a big asteroid's material. The mirrors would need to be able to remove in and out, unlike a dyson sphere, in order to shoot an electromagnetic wave into the structure at the black hole. The majority of this wave would enter the ergosphere being bounced around, with the rotational energy of the black hole energizing the wave and amplifying it. The waves continue to bounce between the mirror and the black hole, getting exponentially faster which is also known as super radiant scattering. If this process was continued, the structure would explode with similar magnitude to that of a supernova. By removing mirrors in intervals, energy from this scattering could be absorbed and used to power civilization for a near infinite amount of time. Since black holes have such a long life, and will most likely be the last thing left in the universe, civilization can last till the end of time when the last black hole evaporates. ((Wikipedia. (2020, March 27). *Black hole bomb*. Wikipedia. https://en.wikipedia.org/wiki/Black_hole_bomb))

Stellar engine

The stellar engine is a mechanism which is basically a way to move celestial bodies from one place to another. This is helpful in the case of something catastrophic such as asteroids

coming to hit earth. Not only this though, but this would be the main mechanism of which humanity would conquer the galaxy. By having a mobile solar system, humans could hop off onto different solar systems and spread humanity there, truly making an interstellar civilization. By moving the sun, due to gravitational laws, all the orbits will follow it. There are currently 2 main types of theoretical stellar engines, which are the passive method and the active method. ((Svoronos, A. A. (2020). The Star Tug: An active stellar engine capable of accelerating a star to relativistic velocities. *Acta Astronautica*, 176, 306–312. <https://doi.org/10.1016/j.actaastro.2020.07.005>))

Passive Method: Shkadov thruster

This method is slower and works well with our current technology that this type 3 civilization would have. It doesn't require an outside energy source beside the sun, and basically powers itself. Essentially, it is a massive 3 dimensional parabolically shaped mirror surrounding $\frac{1}{2}$ of the sun, which reflects the sun's light photons back. These photons carry momentum, which would carry the sun. This would take an extremely long time although it is basically free of energy cost. By building it over the sun's poles, not only does the mirror not orbit the sun but also won't accidentally burn a planet behind the sun from the reflecting photons. Due to this, we'd only be able to travel vertically, which would limit travel options, and also due to the low speeds bigger threats such as a supernova would catch up to the slow moving engine. ((Svoronos, A. A. (2020). The Star Tug: An active stellar engine capable of accelerating a star to relativistic velocities. *Acta Astronautica*, 176, 306–312. <https://doi.org/10.1016/j.actaastro.2020.07.005>))

Active Method:

The better method would be the Caplan thruster, designed by Matthew E Caplan. The engine works similar to that of a rocket thruster, using nuclear fusion at the core and requiring a dyson sphere to create the necessary energy input for this amount of thrust. It uses a thermonuclear driven jet which is collecting matter from the solar wind in order to obtain fuel for nuclear fusion. The jet shown on the left of the mechanism above, is the jet pushing the sun and keeping its distance from it, while the jet on the right pushes the mechanism and the sun. In order for so much mass to be moved, the amount of energy inputted must be tonnes, and in order to do this each cone on the bottom and top funnels hydrogen and helium using electromagnetic fields from solar winds into the engine. This solar wind, however, isn't sufficient and therefore the dyson sphere must come into play. The dyson sphere refocuses the light back on the sun at one point, burning it and causing much mass to come off, which is funneled into the cones. The helium is then burned in the core of the engine, a thermonuclear fusion reactor. This creates radioactive oxygen which is the source of thrust which is seen on the right in the image. The hydrogen collected from before is then accelerated by electromagnetic fields using particle accelerators making the jet stream seen on the left, preventing it from crashing into the sun. Not only is this more efficient at moving our solar system, but extends the sun's lifespan.

((Svoronos, A. A. (2020). The Star Tug: An active stellar engine capable of accelerating a star to relativistic velocities. *Acta Astronautica*, 176, 306–312.
<https://doi.org/10.1016/j.actaastro.2020.07.005>))

Conclusion - 4

As humanity stands on the edge of unseen technological and societal growth, the necessity of transitioning from a planetary to an interstellar civilization becomes more and more apparent. This paper has demonstrated that nuclear energy, with its remarkable efficiency and low environment impact, is the pivotal technology to enable this change. By phasing out fossil fuels and investing in nuclear propulsion systems, we can not only address the immediate challenges of energy consumption and climate change but also create an outlay for interstellar space colonization. The advancement to a currently theoretical Type 3 civilization, characterized by the ability to harness energy on an unprecedented scale, will require collective effort, innovation, and perseverance through the obstacles on the way. Ultimately, harnessing nuclear energy for interstellar travel is not just a scientific endeavor, but a critical step in ensuring the long term survival and prosperity of humanity.

Works Cited

1. Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T., & Curtis-McLane, S. (2008). Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, 1(1), 95–111. <https://doi.org/10.1111/j.1752-4571.2007.00013.x>
2. American History Central Staff. (2024, May 27). Colonial America, Facts, Phases, Characteristics, APUSH. American History Central. <https://www.americanhistorycentral.com/entries/colonial-america/>
3. Anthony', S., Jolle, S., & Captain, J. (n.d.). Advanced Water Purification System for In Situ Resource Utilization. Retrieved August 5, 2023, from <https://ntrs.nasa.gov/api/citations/20140002803/downloads/20140002803.pdf>
4. Arizona State University. (n.d.). Mars Education | Developing the Next Generation of Explorers. [Marsed.asu.edu](https://marsed.asu.edu). <https://marsed.asu.edu/mep/ice#:~:text=More%20than%205%20million%20cubic>
5. AstroRad - What You Wear in Space Could Save Your Life. (n.d.). Lockheed Martin. <https://www.lockheedmartin.com/en-us/products/astrorad---what-you-wear-in-space-could-save-your-life.html>
6. Atmosphere of Venus. (2024, July 17). Wikipedia. https://en.wikipedia.org/wiki/Atmosphere_of_Venus#:~:text=Venus
7. Biello, D. (n.d.). The World Really Could Go Nuclear. *Scientific American*. <https://www.scientificamerican.com/article/the-world-really-could-go-nuclear/#:~:text=And%20if%20the%20world%20as>
8. Blanchett, A. (2017, September 19). Space Radiation is Risky Business for the Human Body - NASA. [Nasa.gov](https://www.nasa.gov). <https://www.nasa.gov/humans-in-space/space-radiation-is-risky-business-for-the-human-body/>
9. Brain, M., & Gleim, S. (2007, December 4). What if an asteroid hit the Earth? *HowStuffWorks*. <https://science.howstuffworks.com/nature/natural-disasters/asteroid-hits-earth.htm>
10. Brandenburg, J. E. (n.d.). EVIDENCE FOR A LARGE, NATURAL, PALEO-NUCLEAR REACTOR ON MARS. <https://www.lpi.usra.edu/meetings/lpsc2011/pdf/1097.pdf>
11. breeder reactor | Description, History, & Types | Britannica. (2020). In *Encyclopædia Britannica*. <https://www.britannica.com/technology/breeder-reactor>
12. Campa, R., Szocik, K., & Braddock, M. (2019). Why space colonization will be fully automated. *Technological Forecasting and Social Change*, 143, 162–171. <https://doi.org/10.1016/j.techfore.2019.03.021>
13. Carpenter, P. A., & Bishop, P. C. (2009). The seventh mass extinction: Human-caused events contribute to a fatal consequence. *Futures*, 41(10), 715–722. <https://doi.org/10.1016/j.futures.2009.07.008>
14. Case, D. (2016, July 19). Passive Radiation Shielding: Integrating Multilayer and Multipurpose Materials into Space Habitat Design - NASA. [Nasa.gov](https://www.nasa.gov). <https://www.nasa.gov/directorates/stmd/space-tech-research-grants/passive-radiation-shie>

- lding-integrating-multilayer-and-multipurpose-materials-into-space-habitat-design/#:~:text=One%20radiation%20protection%20method%20is
15. China's Incipient LFTR. (n.d.). Adam Smith Institute.
<https://www.adamsmith.org/blog/chinas-incipient-lftr>
 16. Christodouleas, J. P., Forrest, R. D., Ainsley, C. G., Tochner, Z., Hahn, S. M., & Glatstein, E. (2011). Short-Term and Long-Term Health Risks of Nuclear-Power-Plant Accidents. *New England Journal of Medicine*, 364(24), 2334–2341.
<https://doi.org/10.1056/nejmra1103676>
 17. Chu, J. (2014, September 18). Shrink-wrapping spacesuits. MIT News | Massachusetts Institute of Technology. <https://news.mit.edu/2014/second-skin-spacesuits-0918>
 18. Ćirković, M. M. (2018). *The Great Silence: Science and Philosophy of Fermi's Paradox*. In Google Books. Oxford University Press.
https://books.google.co.uk/books?hl=en&lr=&id=_npVDwAAQBAJ&oi=fnd&pg=PP1&dq=fermi+paradox&ots=IgrjQq3tR1&sig=rdaAH1AAAtDP0IKfygTIFtsjO-JU#v=onepage&q=fermi%20paradox&f=false
 19. DEININGER, W., & NOCK, K. (1990). A review of nuclear electric propulsion spacecraft system concepts. 21st International Electric Propulsion Conference.
<https://doi.org/10.2514/6.1990-2553>
 20. Denchak, M. (2022, June 1). Fossil Fuels: The Dirty Facts. Natural Resources Defense Council. <https://www.nrdc.org/stories/fossil-fuels-dirty-facts>
 21. Dick, S. J. (2000). Interstellar humanity. *Futures*, 32(6), 555–567.
[https://doi.org/10.1016/s0016-3287\(00\)00007-0](https://doi.org/10.1016/s0016-3287(00)00007-0)
 22. Dinkin, S. (2004, September 7). The Space Review: Colonize the Moon before Mars. *Www.thespacereview.com*. <https://www.thespacereview.com/article/221/1>
 23. Exploring Space . Mars Terraforming Timeline . Mars as Home | PBS. (n.d.). *Www.pbs.org*.
<https://www.pbs.org/exploringspace/mars/terraforming/page7.html>
 24. Gabrielli, R. A., & Herdrich, G. (2015). Review of Nuclear Thermal Propulsion Systems. *Progress in Aerospace Sciences*, 79, 92–113.
<https://doi.org/10.1016/j.paerosci.2015.09.001>
 25. Giegerich, T., Battes, K., Schwenzler, J. C., & Day, C. (2019). Development of a viable route for lithium-6 supply of DEMO and future fusion power plants. *Fusion Engineering and Design*, 149, 111339. <https://doi.org/10.1016/j.fusengdes.2019.111339>
 26. Hargraves, R., & Moir, R. (2010). Liquid Fluoride Thorium Reactors: An old idea in nuclear power gets reexamined. *American Scientist*, 98(4), 304–313.
<https://www.jstor.org/stable/27859537>
 27. Hoffman, J. A., Hecht, M. H., Rapp, D., Hartvigsen, J. J., SooHoo, J. G., Aboobaker, A. M., McClean, J. B., Liu, A. M., Hinterman, E. D., Nasr, M., Hariharan, S., Horn, K. J., Meyen, F. E., Okkels, H., Steen, P., Elangovan, S., Graves, C. R., Khopkar, P., Madsen, M. B., & Voecks, G. E. (2022). Mars Oxygen ISRU Experiment (MOXIE)—Preparing

- for human Mars exploration. *Science Advances*, 8(35).
<https://doi.org/10.1126/sciadv.abp8636>
28. Jiang, J. H., Feng, F., Rosen, P. E., Fahy, K. A., Das, P., Obacz, P., Zhang, A., & Zhu, Z.-H. (2022). Avoiding the Great Filter: Predicting the Timeline for Humanity to Reach Kardashev Type I Civilization. *Galaxies*, 10(3), 68.
<https://doi.org/10.3390/galaxies10030068>
 29. Katusa, M. (2012, July 17). The Thing About Thorium: Why The Better Nuclear Fuel May Not Get A Chance. *Forbes*.
<https://www.forbes.com/sites/energysource/2012/02/16/the-thing-about-thorium-why-the-better-nuclear-fuel-may-not-get-a-chance/>
 30. Kuthunur, S. (2023, November 22). Scientists find hydrogen in Apollo moon rocks, suggesting astronauts can harvest lunar water. *Space.com*.
<https://www.space.com/hydrogen-moon-rocks-apollo-astronauts-samples#:~:text=A%20fresh%20analysis%20of%20moon>
 31. Lavars, N. (2016, September 30). Three great places to live on Mars. *Newatlas.com*; New Atlas. <https://newatlas.com/great-places-to-live-mars/45654/>
 32. Lutz, K. (2021). Sustainable Nitrogen Cycle for Mars Settlement: Operations Analysis, Agriculture, and Soil Ecosystem.
 33. Marques, A. C., & Fuinhas, J. A. (2012). Is renewable energy effective in promoting growth? *Energy Policy*, 46, 434–442. <https://doi.org/10.1016/j.enpol.2012.04.006>
 34. Massachusetts Institute of Technology. (2021). The Fission Process | MIT Nuclear Reactor Laboratory. *Mit.edu*.
<https://nrl.mit.edu/reactor/fission-process#:~:text=When%20a%20U%2D235%20nucleus>
 35. McLAFFERTY, G. H. (1970). Gas-core nuclear rocket engine technology status. *Journal of Spacecraft and Rockets*, 7(12), 1391–1396. <https://doi.org/10.2514/3.30179>
 36. Medicine, C. for F. (2019, September 19). 4 Things Your Body Needs To Survive. *Center for Family Medicine*.
<https://centerforfamilymedicine.com/general-health/4-things-your-body-needs-to-survive/>
 37. NASA. (2023). Mars: Facts - NASA Science. *Science.nasa.gov*; NASA.
<https://science.nasa.gov/mars/facts/>
 38. NASA's InSight Reveals the Deep Interior of Mars - NASA. (2021, July 22). *Nasa.gov*.
<https://www.nasa.gov/centers-and-facilities/jpl/nasas-insight-reveals-the-deep-interior-of-mars#:~:text=Like%20Earth%2C%20Mars%20heated%20up>
 39. Nuclear propulsion systems | ANSTO. (n.d.). *Www.ansto.gov.au*.
<https://www.ansto.gov.au/our-science/nuclear-technologies/reactor-systems/nuclear-propulsion-systems#:~:text=Nuclear%20propulsion%20provides%20several%20advantages>
 40. Office of Nuclear Energy. (2020, January 21). 6 Things You Should Know About Nuclear Thermal Propulsion. *Energy.gov*.
<https://www.energy.gov/ne/articles/6-things-you-should-know-about-nuclear-thermal-propulsion>

41. Plumer, B. (2014, September 5). What would happen if the Yellowstone supervolcano actually erupted? Vox; Vox.
<https://www.vox.com/2014/9/5/6108169/yellowstone-supervolcano-eruption>
42. Price, E. W., & Biblarz, O. (2024, July 19). Rocket - Chemical rockets. Encyclopedia Britannica.
<https://www.britannica.com/technology/rocket-jet-propulsion-device-and-vehicle/Chemical-rockets>
43. Rees, E. (2011, June 23). Don't believe the spin on thorium being a greener nuclear option. The Guardian.
<https://www.theguardian.com/environment/2011/jun/23/thorium-nuclear-uranium>
44. Romero-Calvo, A. (2024, January 4). Magnetohydrodynamic Drive for Hydrogen and Oxygen Production in Mars Transfer - NASA. Nasa.gov.
<https://www.nasa.gov/general/magnetohydrodynamic-drive-for-hydrogen-and-oxygen-production/>
45. Ross, A. K., Ruppert, S., Gläser, P., & Elvis, M. (2023). Preliminary quantification of the available solar power near the lunar South Pole. *Acta Astronautica*, 211, 616–630.
<https://doi.org/10.1016/j.actaastro.2023.06.040>
46. Rothschild, L. (2024, January 4). Detoxifying Mars: the biocatalytic elimination of omnipresent perchlorates - NASA. Nasa.gov.
<https://www.nasa.gov/general/detoxifying-mars/#:~:text=Upon%20arrival%20at%20Mars%2C%20spores>
47. Saenger, T. (2013). Colonial Farming and Food: Famine to Prosperity | NCPedia. Wwn.ncpedia.org. <https://www.ncpedia.org/colonial-farming-and-food-famine>
48. Schunck, N., & Regnier, D. (2022). Theory of nuclear fission. *Progress in Particle and Nuclear Physics*, 125, 103963. <https://doi.org/10.1016/j.pnpnp.2022.103963>
49. Shing, B., & Yim, Y. (n.d.). Why Mars? Retrieved July 20, 2024, from <https://marspapers.org/Essays/E19.pdf>
50. Spervovasilis, N., Markaki, I., Papadakis, M., Tsioutis, C., & Markaki, L. (2021). Epidemics and pandemics: Is human overpopulation the elephant in the room? *Ethics, Medicine and Public Health*, 19(100728), 100728. <https://doi.org/10.1016/j.jemep.2021.100728>
51. Svoronos, A. A. (2020). The Star Tug: An active stellar engine capable of accelerating a star to relativistic velocities. *Acta Astronautica*, 176, 306–312.
<https://doi.org/10.1016/j.actaastro.2020.07.005>
52. Szondy, D. (2024, January 9). NASA may use bacteria to make Martian water drinkable. New Atlas. <https://newatlas.com/space/nasa-bacteria-drinkable-martian/>
53. Terraforming. (2024, July 3). Wikipedia.
<https://en.wikipedia.org/wiki/Terraforming#:~:text=While%20Venus%2C%20Earth%2C%20Mars%2C>
54. Terraforming and the Future of Humans in Space. (n.d.).
<https://www.astro.umd.edu/~miller/teaching/astr380f09/lecture27.pdf>
55. Thorium | Toxic Substances | Toxic Substance Portal | ATSDR. (2021, February 10). Wwn.cdc.gov.

- <https://wwwn.cdc.gov/TSP/substances/ToxSubstance.aspx?toxid=121#:~:text=Soil%20contains%20an%20average%20of>
56. Touran, N. (2020, October 28). Nuclear fuel will last us for 4 billion years. What Is Nuclear? <https://whatisnuclear.com/nuclear-sustainability.html>
 57. Tursunov, A., & Dadhich, N. (2019). Fifty Years of Energy Extraction from Rotating Black Hole: Revisiting Magnetic Penrose Process. *Universe*, 5(5), 125. <https://doi.org/10.3390/universe5050125>
 58. US EPA, O. (2018, November 30). Nuclear Submarines and Aircraft Carriers. US EPA. <https://www.epa.gov/radtown/nuclear-submarines-and-aircraft-carriers#:~:text=Nuclear%20submarines%20and%20aircraft%20carriers%20are%20powered%20by%20onboard%20nuclear>
 59. Vertical farming could be the solution to growing food on Mars. (2023, March 24). CNN. <https://cnn.com/2023/03/24/world/gallery/agriculture-on-mars-interstellar-lab-spc-scen-intl/index.html>
 60. Villaver, E. (2023, September 5). Water on the Moon and the civilizations of the future. EL PAÍS English. <https://english.elpais.com/science-tech/2023-09-04/water-on-the-moon-and-the-civilizations-of-the-future.html#>
 61. Weintraub, D. (2018, September 6). Consider the Martians: Scientists need to confirm whether life exists on Mars before sending humans. Vanderbilt University. <https://news.vanderbilt.edu/2018/09/06/consider-the-martians-scientists-need-to-confirm-whether-life-exists-on-mars-before-sending-humans/>
 62. Wikipedia. (2020, March 27). Black hole bomb. Wikipedia. https://en.wikipedia.org/wiki/Black_hole_bomb
 63. Wikipedia Contributors. (2019, June 19). Liquid fluoride thorium reactor. Wikipedia; Wikimedia Foundation. https://en.wikipedia.org/wiki/Liquid_fluoride_thorium_reactor
 64. Wright, J. (2020). Dyson spheres. *Serbian Astronomical Journal*, 200, 1–18. <https://doi.org/10.2298/saj2000001w>
 65. Xu, V. (2022, June 9). Should we destroy Mercury to create a Dyson sphere? The Starfish Canada. <https://thestarfish.ca/journal/2022/06/should-we-destroy-mercury-to-create-a-dyson-sphere>
 66. Zarubin, B. (2016). Introduction to Light Water Reactors. Stanford.edu. <http://large.stanford.edu/courses/2015/ph241/zarubin1/>
 67. Zhang, A., Yang, J., Luo, Y., & Fan, S. (2023). Forecasting the progression of human civilization on the Kardashev Scale through 2060 with a machine learning approach. *Scientific Reports*, 13(1), 11305. <https://doi.org/10.1038/s41598-023-38351-y>
 68. Zubrin, R. (n.d.). The Case for Colonizing Mars, by Robert Zubrin – NSS. Nss,Org. <https://nss.org/the-case-for-colonizing-mars-by-robert-zubrin/#:~:text=In%20contrast%20to%20the%20Moon>
 69. Klotz, I. (2017, May 11). Boiling Blood and Radiation: 5 Ways Mars Can Kill. Space.com. <https://www.space.com/36800-five-ways-to-die-on-mars.html>

An Explorative Analysis into the Relationship between Age and Peak Player Performance in the NBA By Johnny Chen

Abstract

Historically the NBA has been filled with many prominent players and stars, but one of the longest standing questions has been when these specific NBA players reach their maximum player performance. Previous research suggests that player performance reaches its greatest point from the ages of 24 to 27, with a dropoff occurring around 30 years of age. Utilizing Python 3.10, an analytical study was performed, generating visuals from data collected on the top 100 NBA players from the 2013 to 2020 seasons and evaluating the trends of points per game (PPG) and player efficiency rating (PER) as players aged. Ultimately, the study concluded that from the ages of 24 to 27 peak player performance was observed in regards to offensive output and efficiency rating, as was concluded by previous studies. Further exploration into additional metrics in combination with PPG and PER could provide further insight into how performance changes as players age, while also aiding front office decision making within trades and other team adjustments.

Introduction

From old legends like Bill Russell and Wilt Chamberlain to the 90's greats of Jordan and Magic. From modern day superstars Stephen Curry and LeBron James to young phenoms of the likes of Victor Wembenyama and Anthony Edwards. Throughout the history of the National Basketball Association (NBA), the league has been littered with talented and skilled players. These players encapsulate viewers' attention, "oohs" and "ahhs" following their every action. They stuff the box score every game, seemingly dominating at ease. Yet, there comes a time when even the greats begin to see a decline in performance. Kobe Bryant dominated the association his first 17 years before a steep dropoff in performance his 18th season ("Kobe Bryant Stats", n.d). Michael Jordan averaged nearly 32 points per game over his career before his first retirement. When he returned, his average dropped to less than 23 a game ("Michael Jordan Stats", n.d).

The NBA showcases the meteoric rise and fall of these stars, and fans are often left watching the twilight years of legendary players. Why do we see player performance dropoff significantly? When do we see these top players perform at their prime? To what degree do players ascend and descend from their peak performance? To truly understand when and how the premier players in the NBA reach their athletic peaks, crucial metrics that best evaluate player performance must be analyzed throughout the years. By looking in depth at age trends and picking apart the numbers, we can allow for both viewers and front office decision makers to better understand the game in front of them. Such insight would allow for teams to make more efficient trades and better analyze player value, simultaneously aiding the casual viewer with key statistics. The necessity to analyze the trends of age and performance has increased in recent

years given the rise of ever-younger superstars. Examples of stars. Hypothesize that players are reaching their peak performance at a younger age. How best to evaluate performance.

In order to maximize the accuracy and efficiency of the study, the two metrics selected to analyze were points per game (PPG) and player efficiency rating (PER). PPG serves as a simple but glaring statistic that could be looked at. Wins and losses in the NBA are determined solely on which team scores more points. When the NBA was created in 1946, the stats kept were very basic- points scored, shots made and attempted, free throws made and attempted and fouls (“Basic/Traditional”, 2016) . As one of the core statistics in the NBA, points are one of the most accurate ways to analyze players. A player's contribution to team and individual performance can be most directly observed and evaluated through points per game. Thus, points per game is one of the most pivotal and important stats to analyze, and will be one of the bases of the study. From a more advanced metrics standpoint, player efficiency allows for an accurate all-in-one basketball rating. PER allows for all of a player's productivity and statistics to be measured, summarizing a player's statistical accomplishments into a single number. PER takes into account field goals, free throws, 3-pointers, assists, rebounds, blocks and steals, while also factoring missed shots, turnovers and personal fouls (“What Is Player Efficiency”. 2024) . Although PER isn't particularly indicative of a player's defensive contributions due to its emphasis on offensive contribution, it allows for a generally accurate representation of a player's worth, and condenses an expansive pool of stats to a manageable metric.

The aim of this report was to discover and further analyze the average age at which the top NBA players perform at their peak. For the 2013-2020 seasons, the top 100 NBA players were grouped and filtered by taking the highest ranked in PER and PPG, respectively. Two separate data pools were created for analysis based upon these two metrics. Basic and in-depth analysis was conducted on each of these data sets, including the gathering of general information, as well as a thorough visual analysis and the inspection of possible trends.

Methods

Data Acquisition

Player data was collected from basketball-reference.com. Overall, 51 player metrics were collected with a focus on two separate player output metrics – points per game (PPG) and player efficiency rating (PER). The remaining 49 player metrics and their accompanying variable names can be found in the data dictionary (Appendix A). The desired player pool was the top NBA players, thus only the top 100 players in the PPG and PER metrics were observed in the data sets. In addition, data was observed from 8 NBA seasons collectively, 2013-2014 to 2020-2021. This was to ensure there would be a large enough sample size to strengthen our model and mitigate possible data abnormalities. The desired data was not downloadable in a table format; as a result, the data was transposed into csv format, pasted into a google sheets document, and separated into columns. This process was then repeated twice for each statistic being measured (PPG and PER), and that process repeated 8 times for each season being accounted for.

The data collected was sorted into two separate Google Sheets files, PPG and PER. The files were then converted into .csv, and imported to a Google Colab file. The PPG data and PER data were assigned variable names, `ppg_file` and `per_file`, respectively. Using the command `pd.read_csv()`, the data files were displayed, and our output was the data file that was requested.

Data Cleaning

Once read into Colab, all analyses were performed using Python 3.10. After the acquisition of the two data sets, cleaning operations and basic data analyses were conducted. Null values were identified using the `print(ppg.isnull().sum())` and `print(per.isnull().sum())` commands to detect null values within the data sets. In the ppg data set, 11 null values for 3P% and 789 non null values were identified, representing that in 11 instances, players did not have a 3P%. This was the result of attempting and making zero 3 pointers, meaning there would be no applicable percentage. To fill out the null values, the command `ppg['3P%'].fillna(value=0)` was applied, replacing all na values with zero. This new version was re-saved as the value `ppg['3P%']=ppg['3P%'].fillna(value=0)` meaning that the ppg data set no longer had null values for 3P%. Within the PER data set, there were 2 columns in the original csv file with zero inputs. Therefore, when checking for null values, there were 800 null values under the titles, Unnamed: 20 and Unnamed: 25. To eliminate null values in this data set, the command `per.drop(columns=['Unnamed: 20', 'Unnamed: 25'])` was used to remove the “Unnamed: 20” and “Unnamed: 25 columns”.

In addition, the datasets took into account midseason player trades, calculating combined stats from players throughout the entire season, regardless of team. Furthermore, the datapool eliminated all players who did not meet the prerequisite of games played (60 games). This was in an attempt to take into account performances that might have been injury-influenced, allowing for a more accurate player pool.

General and Specific Analysis

To conduct a basic analysis on the data pool, general information on each data file was required. The PER data was analyzed first using the command, `per.describe()`. This gave generalized information about the PER data set, providing count, mean, standard deviation (std), minimum, lower quartile, median, upper quartile, and max value of each statistic in the file. Due to the goal of the report, the Age and PER columns were prioritized, thus attention was focused upon values in those two statistics specifically.

Among the given measures of central tendencies, mode was not provided. As a result, the command `per['Age'].mode()` was used to discover the most repeated age value within the PER player pool. With the desired basic PER data acquired, focus was shifted onto the PPG data set. Using the same describe command, `ppg.describe()` was used to provide measures of central tendency for each statistic in the PPG file. Likewise to the PER pool, PPG and Age were the two stats prioritized. The command `per['Age'].mode()` was used to determine the most common age within the PPG player pool. With the conclusion of basic analysis, the count, mean, standard

deviation, minimum, lower quartile, median, upper quartile, and max value were gathered for the PPG and PER files, taking into account all 8 seasons.

To perform a further in-depth analysis on the data pool, the two data sets were subset by season. This allowed for each season to be analyzed individually, providing more specific and in-depth breakdown of the numbers. Another reason for this action was to allow for the analysis of the possible trends throughout the seasons. To assess each season within the timeframe, both the PPG and PER data pools needed to be separated. To do so, the columns were broken up and separated by the yearly values from 2013 to 2020. Using the 'loc' command, columns were separated to isolate seasons. The PPG data set was separated first. The first season collected was 2013, so the command `ppg.loc[ppg['Yr'] == 2013,]` was used to take only data entries from the 2013 season. This command was saved as the variable name `ppg_2013` by setting the command equal to `ppg_2013`. The 'loc' command was used for each of the 8 seasons and set to a variable name of `ppg_(desired year)`. The end product was 8 values (`ppg_2013`, `ppg_2014`, `ppg_2015`...`ppg_2020`) each containing only PPG data from their respective season.

With each season separated, the ability to analyze each season specifically arose. By using the describe command, each season's individual descriptive statistics are given. Since mode is not given using `.describe()`, the command `ppg_2013['Age'].mode()` was used to determine the most common player age of the top 100 PPG players in 2013. The same command was used for the rest of the seasons, resulting in descriptive PPG statistics for each separate season in the NBA from 2013-2020. A similar process was followed for the PER player pool, with the loc command being applied to every season again. This time, the command followed the format `ppg.loc[ppg['Yr'] == (desired year),]`. These commands were saved as "`per_(desired year)`", giving 8 separate seasons of top 100 PER player statistics from 2013 to 2020. To analyze each of these new values, the command `.describe()` gathered mean, median, upper/lower quartiles, max/min, and count for PER data set, while `per_(year)['Age'].mode()` accounted for the mode of the age in a season. By the end of this process, there were 8 individual seasons worth of data, with each season from 2013 to 2020 having its own data breakdown from both PPG and PER statistics.

Visual Analysis

In-depth and visual analysis was conducted on the data pool using a multitude of commands from separate libraries. Models and visual graphs were generated using `matplotlib.pyplot` and `seaborn`, while `pandas` and `numpy` allowed for the restructuring and analysis of our data. To begin, a scatter plot for PER and PPG was necessary to visualize the general spread of player age in relation to player performance (Figure 4 and Figure 8). This also allowed for detecting obvious and clear trends throughout the data. Using `matplotlib`, the commands `plt.scatter(x=per['Age'],y=per['PER ▼'])` and `plt.scatter(x=ppg['Age'],y=ppg['PTS ▼'])` output a PER vs Age scatter plot and PPG vs Age scatter plot, respectively. The X axis for both graphs represent the age range from the player pool, from 17.5 to 40 years of age. The Y axis graphs represent the metrics being compared with age, PER and PPG.

In order to create a more visually organized graphic, box plots were utilized (Figure 3 and Figure 7). By creating age bins to bunch together players of the same age group, the data was compiled to be more easily interpreted. Furthermore, box plots presented a more direct way to decipher each data set, providing quartiles, medians, maximums, and outliers within each age range. A dotted red line representing the overall mean of PER or PPG (dependent on which graph) throughout the entire data set was also implemented to aid with analyzing the individual plots. The vertical axes were PER and PPG, respectively, while the horizontal axes represented separate age groups.

Finally, histograms were created that visualized age distributions for the PER and PPG data sets (Figure 2 and Figure 6) in addition to histograms that visualized the PER and PPG distributions themselves (Figure 1 and Figure 5). The x-axes specified a predetermined statistics range, while the y-axis represents the quantity of players within the range. The red line in each graph represented the mean value for the statistic being measured. These graphs allowed for a more general interpretation of the data pool, also helping to explain possible trends or outliers in previous charts and data.

Results

PER

Over 8 cumulative seasons of PER player data, the age ranged from 19 years to 39 years, with the median age of 26 years old (std: 3.81). The upper and lower quartiles of the data set were 29 years and 24 years, respectively. It can also be seen that from 2013 to 2020, 26.57 years old was the average age of a player ranked in the top 100 PER (Table 1). The most represented age within the top 100 PER of a season was 25 years old. As presented in Figure 1, the PER distribution histogram has a right skew. The Age vs PER histogram displays a normal or “bell curve” distribution (Figure 2).

Given the 800 entries of player seasons, the average PER rating was 19.19, with a minimum of 14.6, a lower quartile of 16.3, a median of 18.25, an upper quartile of 21.2, and a maximum of 31.9 (Table 1).

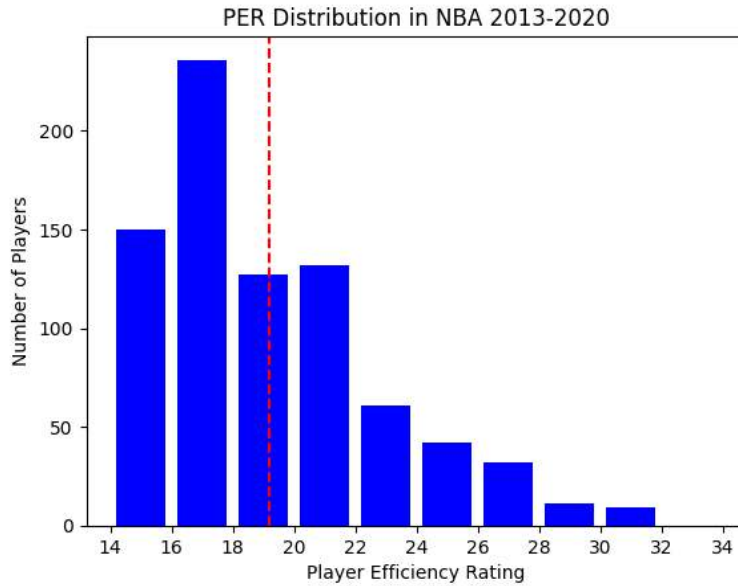


Figure 1. Histogram of Player Efficiency Rating count for the top PER NBA players from 2013-2020.

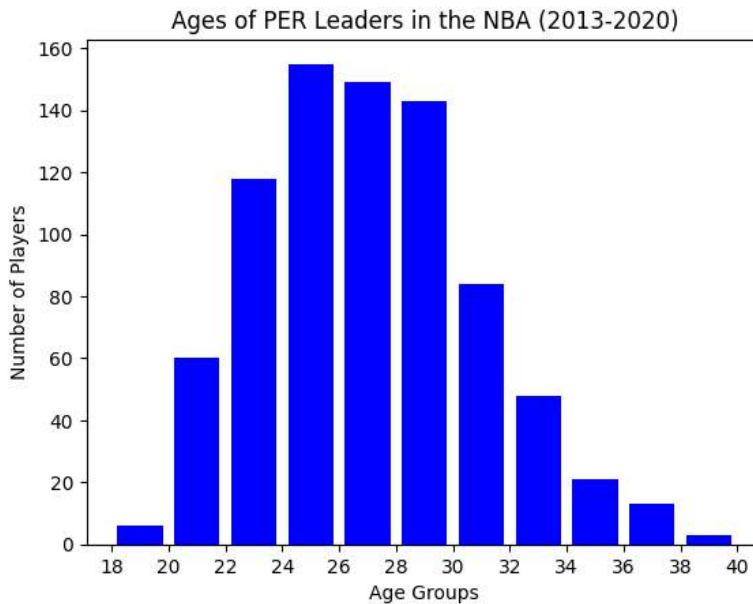


Figure 2. Histogram of Age groups for the top PER NBA players from 2013-2020.

When observing the 2013 season alone, the average age of a player was 26.66 years old with the oldest being 37 and youngest 20. The mode of the 2013 season was 3 years greater than the overall dataset, with 28 being the most common age of a player in the list. The 2014 season had a mean player age of 26.97, with the oldest and youngest players being 38 and 21. The mode of the top 100 PER ranked players was 25. The following season, 2015, the average age was

26.48. The oldest player in the PER pool was 39, and youngest 20. The average age for players in 2016 was 26.43 years, with 20 and 36 being the maximum and minimum. The 2013, 2014, 2015, and 2016 season all shared the same lower, middle, and upper quartiles as the general dataset. The 2017 season had a mean player age of 26.40, with a minimum of 19 and a maximum age of 39. The 2017 season had 3 modes, meaning 3 ages appeared most frequently an equal amount. These ages were 23, 25 and 27, with lower and upper quartiles 23 and 29 respectively. There was an increase in average player age of .16, resulting in a mean age of 26.56 years. The minimum of the 2018 season was the same as 2017, with a maximum of 2 years younger (37 years). The upper and lower quartiles of the 2018 season were 23.75 and 29.25. In the 2019 season, the mean, median and mod of the data set were 26.54, 26, and 27 respectively. The minimums and maximums of 2019 were 20 and 34, with quartiles of 23.75 and 29.25. The final season data was gathered for the 2020 season, with the average PER leader age being 26.52. The minimum was 19 years, while the maximum was 36. The entirety of the general and individual PER data is displayed in Table 1.

Table 1. Data breakdown and descriptive statistics of age vs season from 2013-2020 of the top 100 PER ranked players each season.

	Mean	Min	Lower Quart.	Median	Upper Quart.	Max	Mode
Overall	26.57	19	24	26	29	39	25
2013	26.66	20	24	26	29	37	28
2014	26.97	21	24	26	29	38	25
2015	26.48	20	24	26	29	39	25
2016	26.43	20	24	26	29	36	25
2017	26.40	19	23	26	29	39	23, 25, 27
2018	26.56	19	23.75	26	29.25	37	26
2019	26.54	20	23.75	26.5	29	35	27
2020	26.52	19	24	26	29	36	28

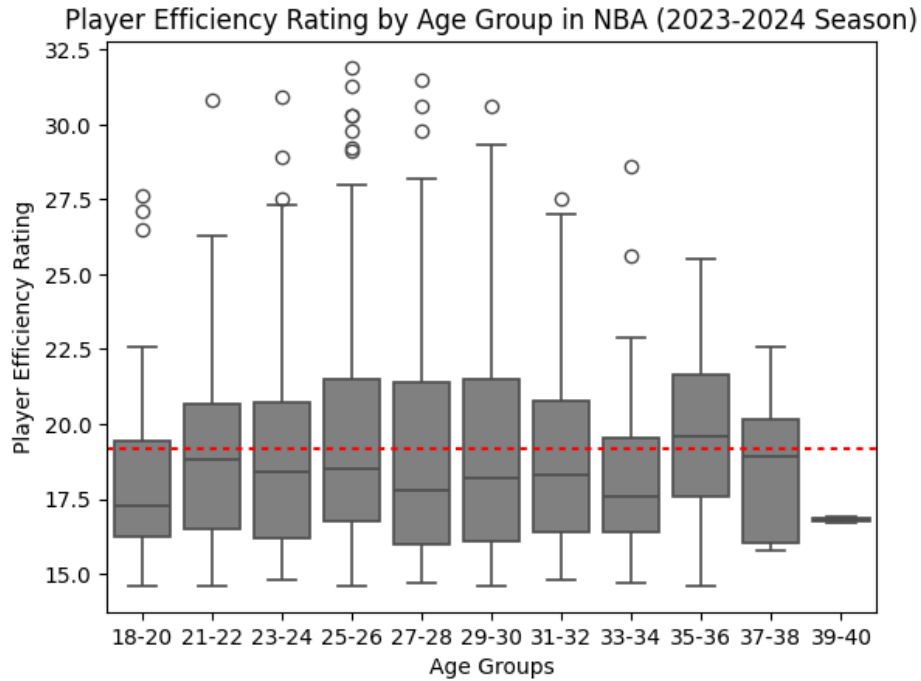


Figure 3. Box plot of NBA age groups compared to PER for players each year from 2013-2020

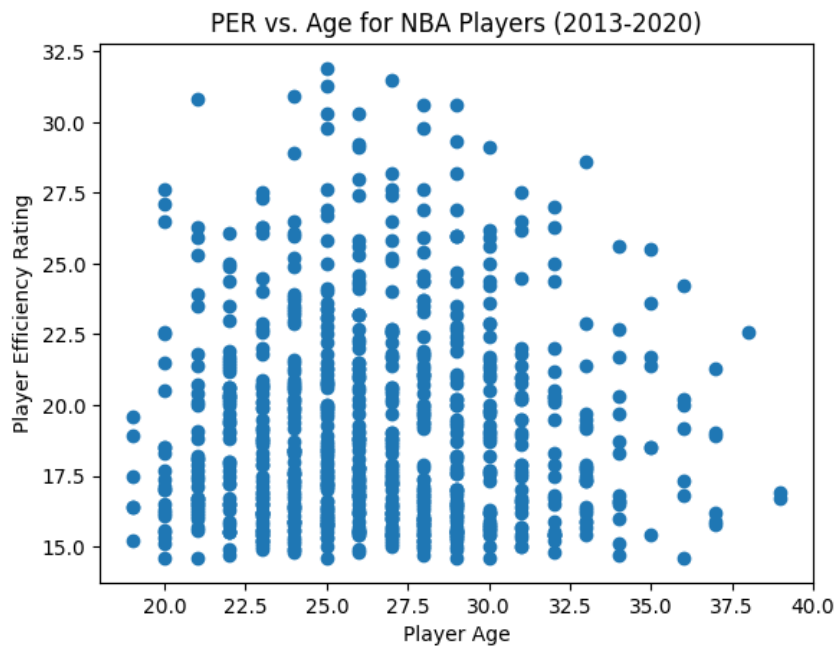


Figure 4. Cumulative scatter plots of top 100 PER rated players each year from 2013-2020 by Player Age vs. Player Efficiency Rating (right)

The PER scatter plot has a slight right-skew, with the highest recordings appearing to the left of 27.5 years (Figure 4). Furthermore, as you look at the right side of the plot, the players are more scarce and appear less in general, with only 3 players after 37.5 years. This compared to the very high frequency of players between 20 and 30. All age groups median, quartiles, maximums, minimums, and outliers in PER are displayed in Figure 3.

PPG

From the 8 years worth of PPG leader data, the age range spanned 20 years, a maximum age of 39 and minimum age of 19. The median was 26 years old (std: 3.87). The upper and lower quartiles over the 8 years were 23 and 29. From 2013 to 2020, 26.2 years old was the average age of a points per game leader. The most represented age within the top 100 PPG leaders of a season was 24 years old. As presented in Figure 3, the left PPG distribution histogram has a heavy right skew, representing that the highest quantity of players averaged between 14 to 16 points. Nearly all players average less than 24 points, with the quantity of players above that range diminishing significantly. The age variable in the PPG data set was observed to have a normal distribution with a range from 18 to 40 (Figure 3).

The average points per game (ppg) for PPG leaders over the 8 seasons was 17.2 points per game. Within the data set, the maximum ppg was 36.1 (registered in 2018), and the minimum 11.7 points (registered in 2013 & 2014). Lower and upper quartiles were 13.6 and 19.82 ppg respectively, with a median of 15.9 points.

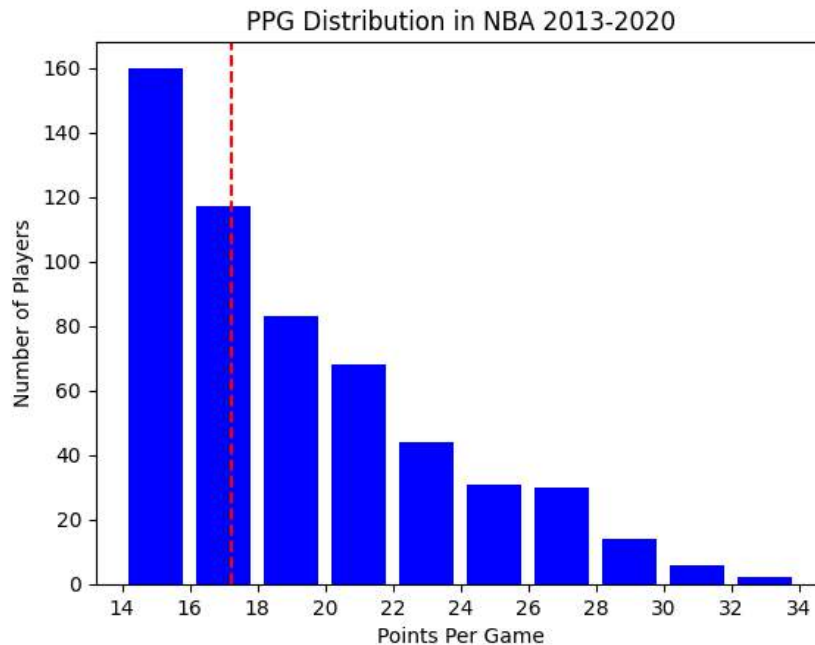


Figure 5. Histogram of Points Per Game count for the top PPG NBA players 2013-2020.

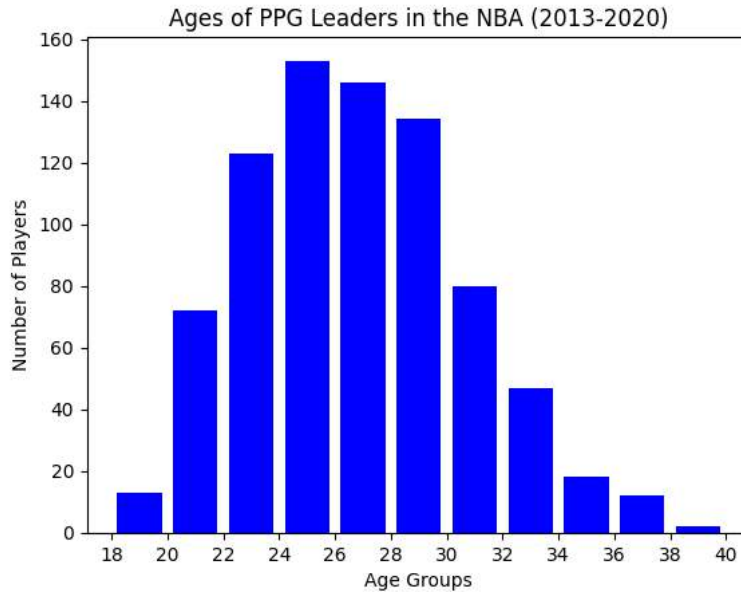


Figure 6. Histogram of Age groups for the top PPG NBA players 2013-2020.

In analyzing individual seasons, the 2013 season had an average PPG leader age of 26.58, with a minimum of 19 and maximum of 37 years old. The lower quartile and upper quartiles of the season were 24 and 29 years. The median age of 2013 PPG leaders was 26, and the mode 25. The 2014 season shared the same minimum, median, and upper quartile as 2013, with a higher lower quartile by 1 year (24 years), and a max of 1 year older (38). The mean of 2014 was 26.52, with a mode of 24. In 2015, the minimum was 20, and the maximum age was 37 years old. The median of the data set was 25.5, with a mode of 28 years old. The average age of a top 100 ppg scorer in 2015 was 26.08 years old. The mean, median, and mode ages for the 2016 season were 26.49, 26, and 26. The lower and upper quartiles were 24 and 28.25, with a max of 20 and 36. In 2017, the minimum and maximum ages of players in the top 100 PPG was 19 and 39, with an average age of 26.25. The lower, middle, and upper quartiles of the 2017 season were 23, 25, and 29 years old. The most common age in the player pool was 25 years old. In the 2018 season, the quartiles were 23, 26, and 29 respectively, with a minimum and maximum of 19 and 37. The average age was 26.31, with a mode of 25 years. The 2019 season featured a mean of 26.17, a minimum of 19 and a maximum of 35. The lower, middle, and upper quartiles were 23, 26, and 29. The mode of the 2019 season was 25 years old. In the final season of the study, the average age of the PPG leaders was 25.83. The 2020 season had a minimum and maximum age of 19 and 36 years. The lower and upper quartiles of the 2020 season were 22.75 and 28.25 years. The mode for 2020 was 22 and 28, with those two ages appearing most frequently.

Table 2. Data breakdown and descriptive statistics of age vs season from 2013-2020 of the top 100 PPG ranked players each season.

	Me an	Min	Low er Quar t.	Medi an	Up per Qu art.	Ma x	Mo de
Over all	26.2	19	23	26	29	39	24
2013	26.5 9	19	24	26	29	37	25
2014	26.5 2	19	24	26	29	38	24
2015	26.0 8	20	23	25.5	28	37	28
2016	26.4 9	20	24	26	28. 25	36	26
2017	26.2 5	19	23	25	29	39	25
2018	26.3 1	19	23	26	29	37	25
2019	26.1 7	19	23	26	29	35	25
2020	25.8 3	19	22.7 5	26	28. 25	36	22, 28

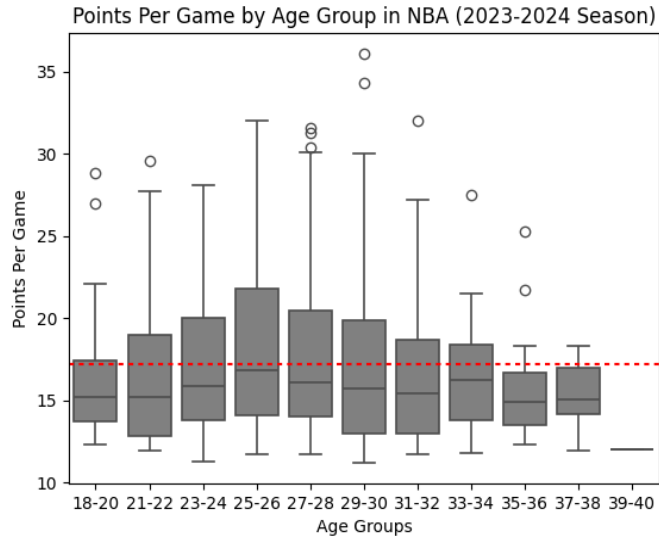


Figure 7. Box plot of NBA age groups compared to PPG for players each year from 2013-2020

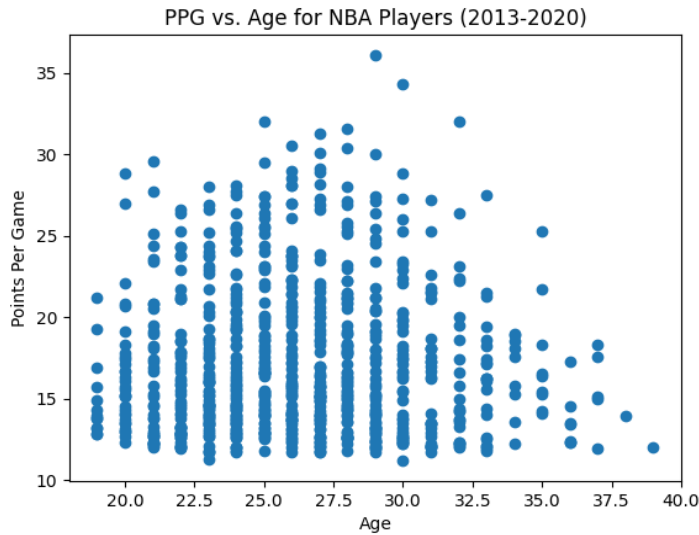


Figure 8. Cumulative scatter plots of top 100 PPG rated players each year from 2013-2020 by Player Age vs. Points Per Game

As can be seen in the PPG scatter plot in Figure 8, a normal distribution was observed, with the highest PPG recordings occurring between 27.5 to 30 years (Figure 8). In addition, the highest density of players occurs around 27.5 years old. All age groups median, quartiles, maximums/minimums, and outliers are displayed (Figure 7), as well as a red line drawn across to visualize the universal average points per game for all players in the data set.

Discussion

PER

As a result of the PER leader analysis, it can be seen that the most common age among top 100 PER leaders from 2013 to 2020 was 25 years old. This meant that within all 800 entries

of player seasons, the most frequent age was 25. However, in only 4 out of the 8 individual seasons was 25 the mode of the data set, and in one of those instances, 25 shared the same number of entries as ages 23 and 27 (Table 1). In addition, the 2013 and 2020 season had the highest mode of age, with 28 being the most common player age in PER leaders. However, when looking at the average player age of each season, the 2020 season had the 4th lowest mean age (26.52), and the 2013 season had a lower average age (26.66) than 2014 (26.97) which had a mode of 25 years old. This meant that even though the 2013 and 2020 season had the highest mode, the overall age spread of the two seasons resulted in a lower average age than 2014 (26.97).

The median age, or the middle player (in terms of PER score) of every season was 26 with the exception of the 2019 season, when the median was 26.5. This meant that despite the most common age amongst PER leaders being 25, if one were to take the 50th/51st player in age each season, they would have all been 26 years old barring one season. It can also be seen that the average age of a PER leader in every season, including the cumulative measurement, was between 26.4 and 27 years old (Table 1). Thus, despite the most common age being 25 throughout all PER leaders, if you were to take the average age or median age, you would be left with an age between 26 and 27 years old. From the table we can determine that the most frequent range of player age among the PER leaders was 25-27 years old.

There was no trend in terms of PER leader player age throughout the years, with the average age, mode, and median all showing no specific growth pattern.

The overall PER distribution by age can be seen to have a normal distribution, with a slight right skew (Figure 2). From this graph, you can determine that most players in the PER leaders are between the ages of 22 and 30. It can also be concluded that the largest population of PER leaders is within the 24-26 age range, with the 2nd and 3rd largest groups being 22-24 and 26-28. This supports the findings of our table, with the two age groups sharing similar age ranges. By evaluating PER vs age we can see a higher concentration of players within the ranges of 22.5 to 32.5 (Figure 4). We see a significant dropoff of players after the 32.5 year old mark. This shows that after 32.5 years old, the number of PER leaders decreased heavily. The highest PER average value was registered by a player that was 25 years old, and the highest PER averages in general were located between 25 and 27.5 years. The boxplot however (Figure 3) yields interesting results: The highest general “maximum” values were recorded by the 29-30 age group, with the 25-26 age group having the largest number of outliers. When assessing the mean vs median, we can see that the mean PER of all 800 entries - represented by the red line - was greater than the median PER for every age group except the 35-36 age group. This is the result of outliers increasing the mean PER value, while the median PER value for each group was simply the middle player’s PER rating. This also means that out of all the age groups, the 35-36 group was the only age group that had more than half of its players average greater than the mean PER of all PER leaders from 2013-2020. This is due to the fact that because the median was greater than the mean, the upper half of the age group would all have PER’s greater than the mean PER.

PPG

The PPG table concluded that the most common age among top 100 PPG leaders from 2013 to 2020 was 24 years old. When comparing the general mode to the mode of each individual season, it can be seen that 24 was the mode of a singular season in only 1 out of the 8 seasons. This meant that despite 24 not being the mode in 7 out of 8 seasons, when combining the number of 24 year olds each season over the 8 years, it emerged as the most common age throughout the time frame.

The 2015 and 2020 season had the highest mode of age, with 28 being the most common player age. Despite that, the average age of the two seasons represented the two lowest mean ages in the dataset. The 2020 season had a mean age of 25.83, the lowest in all 8 years. The 2015 season had a mean age of 26.08, the 2nd lowest. Within the 2020 PPG leaders, the mode age was shared between 28 and 22 years, explaining why the average was the lowest in the study. Because the mode of the 2020 season had both the highest and lowest ages, the overall mean averaged out to a lower value than the rest of the seasons. In the 2015 season, the low average can be attributed to the overall spread of PPG leader ages. Compared to the rest of the data, the 2015 season had the lowest upper quartile value (28 years). This meant that 75% of the players were aged 28 or less, resulting in a lower average age than all years except 2020.

The mean age of all seasons ranged from 25.83 to 26.59 years old (Table 2), meaning that the average age group of a PPG leader was between 25-27 years old. The median age ranged from 25 years to 26 years, proving that the mean and median shared a similar age group. This meant the middle aged player of the PPG leaders each season was an accurate representation of the average age of PPG leaders. This also showed that despite the most common age of a player being 24, the average age of a PPG leader was greater than the most common age. The player distribution also shows how most PPG leaders fell between the 24-26 range, with the 26-28 range and 28-30 range being the 2nd and 3rd most frequent age groups (Figure 6).

Overall, there was a slight downwards trend in age when evaluating mean player age of PPG leaders. The 2020 season yielded the lowest player average, and apart from the 2015 and 2018 season, the 8 year trend was one that had a consistent decrease in age.

The PPG vs Age scatter plot distribution has a right skew (Figure 8). Most PPG leaders are between the ages of 22.5 and 30. From the concentration of markers, we can determine that the greatest number of PPG leaders are within 24 to 30 years old. This supports the findings of Figure 6 and Table 2, with the two age groups sharing similar age ranges. There is also a significant dropoff of players after the 30 year old mark. However, the frequency of outliers after 30 years old seems to increase, with the highest PPG average registered by a 29 year old and the 2nd highest PPG average registered by a 30 year old.

The boxplot (Figure 7) supports our findings by displaying the highest overall maximum and upper quartile values in the 25-26 age group, with 29-30 yielding the highest singular value due to outliers. Furthermore, when comparing the mean vs median for PPG leaders, the mean PPG was greater than the median PER for every age group. This meant the average points scored universally was greater than the points scored of the middle PPG leader in an age group. This is

the result of the highest PPG players scoring an amount that was much greater than other players, thus the difference between any value and max PPG was great enough to raise the mean PPG value higher than the median.

Previous Studies

In a previous study conducted (Salameh, 2023) investigating the prime performance age of NBA players from the perspective of general managers, the age of peak performance was found to be between 24 and 27 with a slight decline in performance after the age of 29 (Salameh 10). This aligns with the results of this study: The age at which we see player efficiency decrease for PER was 32.5, and 30 for PPG. In addition, the general age where we saw peak performance was from 24-27 for PPG and 25-27 for PER, which is almost completely agreeable to Salameh's findings. Furthermore, the variables evaluated altered from study to study. In this study, PER and PPG were evaluated, while Salameh used PER, Box Plus Minus (BPM), and Games Played (GP). Despite the differences in stats measured, the findings of peak player performance remained the same.

Conclusion

From 2013-2020, it can be concluded that the peak performance age of a PER leader was between 25 and 27 years old. The overall "maximum" in PER production was between 29-30 years, but at 25-26 years we saw players reach the greatest heights in terms of rating. Each season, the average age of a PER leader was from 26.4-27 years old. It can also be deduced that 25 year olds most frequently output high PER rankings, as a result of them being most commonly represented among PER leaders. After around 32.5 years old, there was a noticeable dropoff in PER rating, as well as the quantity of PER leading players as well.

The prime performance of a PPG leader from 2013-2020 was between 24-27 years old. The highest output of points scored for players was between the ages of 24-26, with the age group representing both the largest number of players and highest "maximum" and upper quartile values. The 29-30 year old age group contained the singular highest scoring outputs, with multiple outlier seasons producing the greatest PPG values. The average age of a PPG leader each season was 25.8-26.6 years old. The most frequent appearance amongst PPG leaders came from 24 year olds, with a dip in PPG averages and quantity of players occurring after 30 years old.

This study can be evaluated and further tested to encompass more metrics and parameters, allowing for a full analysis of when NBA players reach their athletic primes. These results and conclusions can be used when evaluating a players market value, when the ideal time to trade for a player might be, as well as aiding in a plethora of other front office or fantasy decisionsn.

Works Cited

- admin. (2024, April 16). *What Is Player Efficiency Rating (PER) in NBA Advanced Stats*. The Midfield. <https://themidfield.com/nba-betting/player-efficiency-rating/Basic/Traditional Stats vs. Advanced Stats>. (2015, October 15). NBA.com: Jr. NBA. <https://jr.nba.com/basictraditional-stats-vs-advanced-stats/>
- Basketball Reference. (2000a). *Kobe Bryant Stats* | *Basketball-Reference.com*. Basketball-Reference.com. <https://www.basketball-reference.com/players/b/bryanko01.html>
- Basketball Reference. (2000b). *Michael Jordan Stats* | *Basketball-Reference.com*. Basketball-Reference.com. <https://www.basketball-reference.com/players/j/jordami01.html>
- Salameh, T. (2023). *An Empirical Analysis of Prime Performing Age of NBA Players; When Do They Reach Their Prime?* Bryant Digital Repository. <https://digitalcommons.bryant.edu/eeb/vol16/iss1/14/>

Impact of Media on Political Polarization By Tiana Settle-Bowser

According to the Carnegie Endowment for International Peace, from the 1970s until the mid-2010s, political violence remained an activity of extremist groups on the fringes of society. In the late 1970s, only 6% of Americans justified violence for political goals. By February 2021, 25% of Republicans and 17% of Democrats felt threats against the other party's leaders were justifiable, and one in five Republicans believed that political violence was justified "these days" (Kleinfeld, pg. 22). Through news sources and differing values, political polarization has greatly impacted voting, political parties, and political identities. Political polarization is destructive because it makes it harder to pass legislation in the government. Two groups that do not want to work together creates stunlock within the legislative process and leads to a less effective lawmaking process. Political polarization also affects human connection by making it more difficult to have positive relationships with people of different political identities. This leads to even more polarization due to the lack of human connection within the two-party system. Polarization leads to increased hostility and a lack of willingness to compromise or work together with people who hold different political views. This phenomenon can be seen in both the online world and in-person interactions. Factors such as race, ethnicity, and religion have the potential to shape individuals' core beliefs, values, and worldviews. Individuals may feel a stronger connection to those who share a similar racial, religious or ethnic background to them. This does not mean that diversity within Congress contributes to political polarization in itself. It is very beneficial for people to have representation of their identities through government representatives. Instead, news sources create political polarization in different areas of politics, such as voting, political parties, and political identity, which in turn makes it harder to pass legislation, especially in identity politics. The growth of political extremism within both parties does not unify Americans and increases political polarization. This in turn leads to even more separation along the political ideological axis.

An idea that transcends political boundaries of parties, and something that most people agree upon, is the fact that there is currently no common cause or collective national identity within the US as of now. Although this tension has been lingering in American politics since the country's beginnings, the intensity of it is unique to this time in history. This issue is not a global one; it seems to be particularly acute in America. These divisions have also fallen upon a singular political axis, with the Republican party on one side, and the Democratic party on the other.

Tribal mentality is the human tendency to seek out and connect with like-minded people who share common interests, beliefs, or habits. A tribal mentality creates behaviors and attitudes that stem from strong loyalty to one's own tribe and social groups. This mentality is very common in the culture of political parties. It oversimplifies political differences and distorts perspectives of others with different opinions. This tribal mentality also distorts "complex problems by dividing the world into an us and them," (Roscini, n.d, pg. 3). This mindset leads to the tendency to favor members of one's own group over other groups at an irrational level. This

in turn creates hostilities between political parties, since they work not only in their own interest, but against the other party's as well.

During the 1980s, a new wave of conservatism hit Great Britain when Margaret Thatcher became prime minister in 1979. Thatcher was the first women prime minister, and the longest serving of the 20th century. Thatcher's political and economic philosophy became known as "Thatcherism" and followed four main components; commitment to free enterprise, British nationalism, a plan to strengthen the state by improving efficiency, and a belief in traditional Victorian values, especially hard work and civic responsibility. In opposition to Thatcherism during this era was the Labour Party. This party proclaimed themselves a socialist movement whose principles included a guaranteed minimum standard of living for everyone, nationalism of industry, and heavy taxation of large incomes and of wealth. Despite the irreconcilable differences between these two parties during the 80s, the dispute did not tear down the walls of British democracy. This is because at the time, Britain had a strong government institution that was built to withstand these party divisions. Although, the same degree of ideological distance and societal division shown in England would create violent conflict in a society with weaker institutions. The established tradition that made sure England could withstand this conflict is a concept called "democratic pluralism." A democratic pluralist government allows individuals to achieve positions of political authority through electoral colleges. Both England and the United States are democratic pluralist countries.

Twenty years earlier, in the 1960s, in America, the Democrats began passing major civil-rights legislation. The Civil Rights Act of 1964 prohibited the discrimination on the basis of race, color, religion, sex or national origin. Before this legislation was passed, the Democrat and Republican party were both racially and gender mixed, and were not demographically different from one another. After the legislation passed, an event called "the great re-sorting" took place. Black voters began to gravitate towards the Democratic party because of their stance on civil rights issues. This caused white Democratic voters to flee to the Republican side. Over time, the effects of this sorting registered more broadly and began to shape political identity in America. The most noticeable effect of this sorting were voting patterns. People of the same race, ethnicity, gender, or neighborhood, tended to vote for the same candidates in the same party. This alongside a rather rigid party structure can lead to increased polarization between different groups.

Something unique to the American political system is how we collapse an extremely wide range of legitimate political and social debates onto a singular battle line. America's rigid, two-party electorate system can make our differences appear even larger than they may be. During election season, when the balance of support for both the parties is close enough for either of them to gain electoral advantage, the competition becomes extremely cutthroat and politics begin to feel less about resolving issues, but rather for personal gain. This has been the state the United States government has been in for more than a quarter of a century. One side's gain is not the gain of the entire country; it is always inherently the other side's loss. This obsession with a person's "side" winning has made it extremely difficult to find a common cause

that everyone will fight for. The coronavirus, for example, was a common enemy to both public health and the American economy. Despite this, political parties still found a way to fight over it, in terms of vaccines and lockdown protocol. With the Republican party encouraging anti-lockdown sentiment, and the Democratic party implementing mask protocols.

A 2016 Pew Research Center study confirmed that “64 percent of US adults say that fabricated news stories cause a great deal of confusion about the basic facts of current events” (Barthel, pg.2). Sixty-four percent of Americans also state that they are unable to identify internet bots as a source of misinformation, meaning that they cannot tell the difference between information from real people and computers. This means that they cannot detect or disregard fake news that these bots may spread either. Due to ideological polarization across the political axis, diversity of opinions and arguments may be compromised. This is because people feel the need to fit themselves into the categories supplied by the axis, which makes everything very black and white. The spread of online misinformation⁷ increases easily because of people refusing to listen to other perspectives. People in extremely homogeneous political groups may ignore the facts that would prove some of their arguments wrong.

A major source of misinformation and leading cause of political polarization are different news sources on both ends of the spectrum. As news channels are constantly vying for the attention of their target audience, they have to find new ways to keep people engaged and continuously watching, leading to more and more extreme ideologies being shared with the general public.

After the 2020 election, FOX news, a republican leaning news channel, had two choices to make. They could announce the loss of Donald Trump when the second results came in, or they could deny the election and give their audience false hope. FOX News decided to do the first option and announce to their audience that Trump did in fact lose the election. Newsmax, a rival Republican leaning news channel, did not call Biden a president-elect right away, meaning that in their audience's mind Trump was not a loser just yet. Newsmax's 7pm host Greg Kelly kept stating that he believed Trump could stay in office for more than four years after the election. The Newsmax channel's banner claimed “IT ISN'T OVER YET,” making Newsmax watchers truly believe that Trump was the rightful winner of the 2020 elections. Before this election, FOX News only lightly dabbled in election denial at first. Newsmax on the other hand went all in, and FOX News watchers went from FOX to Newsmax. Audiences suddenly had somewhere else to go where they were told that their desired outcome was in fact achieved.

FOX realized after this event that in order to stay a competitive news source, they had to give their viewers information that they wanted to hear rather than accurate news. After the 2020 election Trump was treated as a political genius on FOX news, not someone who failed at reelection. Some of the network's key shows waded even deeper into the voter fraud depths. So much so that the news source faced massive defamation lawsuits by voting machine companies Dominion and Smartmatic. FOX agreed to pay more than 787 million to settle the lawsuit with Dominion Voting Systems over the networks airing of false claims of voter fraud after the election (Bauder, 2023, par. 1). Cable news's power also comes from their direct connections

with the President himself. Trump constantly bashes FOX news cable rivals and communicates with the cable network regularly. This means that it is unlikely for the network to say anything negative about Trump, since he himself defends them and their journalism. Due to the network speaking on any of the negatives of Trump's presidency, viewers do not get the full story and have holes in their knowledge of current events.

Bias is not a problem on the just the Republican side of news, Democrat-leaning news sources have their fair share of controversy as well. MSNBC as a news source does not offer coverage of stories from an unbiased perspective, but rather counterprogramming against their right-wing rivals. The news source uses arch sarcasm and an extremely partisan style of reporting to engage their audience. This style of broadcasting does not have a good reputation among people who don't watch the source. This constrains NBC news, MSNBC's corporate sibling and gives them a bad reputation. Despite this, NBC still remains one of the country's number one news sources in the evening with 778,000 average viewers in April 2024 according to CNN.

Despite both MSNBC and FOX having very partisan audiences and extreme approaches, FOX news is significantly more dangerous because they go beyond encouraging partisan divisions. The news source misleads the public by not giving the full story of events, which in turn endangers the public by giving them false information. This leads to extremely differing opinions between FOX news viewers and the rest of the country. Which creates even more discord and intense quarrels between the two parties. The network also welcomes dangerous and irresponsible guests commonly onto their shows, who promote false evidence and stories. This is what differs FOX from MSNBC.

Trump's speech sparked both acclaim and controversy, as his rhetoric resonated with some parts of the population while igniting criticism from others, representing the deepening divide within American society. Trump made the speech at Mount Rushmore on July 4 in 2020. At this time, he was down in the polls and failing to control the raging coronavirus pandemic, so he shifted the focus to a waging battle he called "against a new far-left facism" (Nakamura, 2020, par. 2). Trump stated in this speech that a new radical ideology was attacking our country, following the protests that came after George Floyd's death in police custody. He said that these protests would "demolish both justice and society" (Nakamura, 2020, par. 3) While some called Trump's speech against the protests divisive, Tucker Carlson praised it as the best speech the President has made in his life (Bauder, 2020, par. 5). In his talk show on FOX news after this speech, Carlson played clips of cable TV hosts and analysts critical of some of America's actions historically. Carlson said that "these people hate America. There's no longer any question about that" (Bauder, 2020, par. 6). This idea that being critical of American institutions and historical actions is unamerican and unpatriotic encourages people not to question their current circumstances, and encourages a sense of passivity within citizens when it comes to injustices.

Geographic advantages play a large role in local elections, requiring candidates to appeal to a diverse array of voters, including those in rural areas. However, polarized media outlets complicate this task, hindering efforts to pass legislation. Most Republicans who lean the farthest

right are from rural counties. Due to homogeneity within political beliefs of these rural counties, they are more capable of moving further away from the center than metropolitan areas. This geographic advantage allows the GOP to run campaigns aimed at the rural voter well to the right of the median American. To win over these rural groups in local elections, a democratic candidate must appeal to both the voter in the middle and the voter well to the right of the middle. Furthermore, as media outlets continue to cater to these partisan audiences, their coverage reinforces the existing political divisions, making it harder for candidates to find common ground and pass legislation that addresses identity political issues.

While economic policy is extremely important in governmental policy, there needs to be equal attention towards policy against oppression and marginalization of minorities. Despite this, there tends to be an ultimatum between economic and identity issues that suggests an “either-or” scenario. Both of these issues can be addressed at the same time by politicians, despite this mentality. Research by economists Daron Acemoglu and James A. Robinson argues that inclusive economic institutions, which promote opportunities for all individuals regardless of background, can contribute to reducing identity-based conflicts. Political party polarization between Republicans and Democrats places limitations on making progress in both identity and economic policies due to the increasing divide between political ideals and values. Identity politics refers to the discussion of and politicking around issues pertaining to one's identity. According to a Vox article titled, “The battle over identity politics, explained”, focus typically falls on women, racial minorities, immigrants, LGBT people, and religious minorities, such as Muslim Americans. Social issues such as same-sex marriage, police shootings of unarmed black men, trans people in bathrooms, fluidity of gender, discussion of rape culture, safe spaces and trigger warnings, are typically the kind of issues people mean when they refer to identity politics.

The two political parties were not always as polarized as they are now; the two parties were similar around the 1950s in terms of social values. “The American Political Science Association issued a plea that Democrats and Republicans make more of an effort to distinguish themselves around this time” (Kolbert, 2021, par. 13). Division began with the Civil Rights Movement, the Women's Movement, Richard Nixon's Southern Strategy, and Roe v. Wade. These landmark movements instigated a social sorting that eventually led to the ideological division between the two parties that we see today.

In the wake of the previously mentioned movements, the GOP became demographically whiter, more religious, and more male than its counterparts (Kolbert, 2021, par. 15). These differences, already made significant by the early 1990s, had become even more pronounced by the 2010s (citation needed here). Due to lack of diversity within the Republican party, identity politics are not of the most importance to the party, since they are not necessarily pressing issues for the population represented within the party. People representing the Democratic party say that America needs to continue discussing and fighting on all issues relevant to identity politics. These issues include racism in the criminal justice system, if gay people should be protected against discrimination, and whether transgender people should be allowed to use the bathroom that aligns with their gender identity. With some progress made on these issues in the past few

years, advocates say that it's too late to abandon them now. Trump's election also created a new sense of urgency to discuss these issues. This is because he was so at odds with many of the ideals promoted by the left in discussions about identity politics.

Through news sources and identity politics, political polarization has greatly impacted voting, political parties, and political identities. Furthering understanding of how political polarization affects what is seen in politics is very beneficial because. Critically analyzing political news and media and deciding whether or not the source is reliable will better our understanding of politics. Looking at the political content consumed with a more centrist attitude, instead of a skewed one, can help individuals engage in more civil conversations in politics.

Works Cited

- Lopez, German. "The Battle over Identity Politics, Explained." *Vox*, Vox, 2 Dec. 2016, www.vox.com/identities/2016/12/2/13718770/identity-politics.
- Kolbert, E. (2021, December 23). *How Politics Got So Polarized*. The New Yorker. <https://www.newyorker.com/magazine/2022/01/03/how-politics-got-so-polarized>
Political polarization, identity politics and social media // The Observer. (2022, November 11). The Observer.
<https://ndsmcobserver.com/2022/11/political-polarization-identity-politics-and-social-media/>
- Grynbaum, M. M., Alba, D., & Epstein, R. J. (2021, March 1). How pro-Trump forces pushed a lie about Antifa at the capitol riot. *The New York Times*.
<https://www.nytimes.com/2021/03/01/us/politics/antifa-conspiracy-capitol-riot.html>
- DiResta, R. (2021, October 9). *It's not misinformation. It's amplified propaganda*. The Atlantic. <https://www.theatlantic.com/ideas/archive/2021/10/disinformation-propaganda-amplification-ampliganda/620334/>
- Molla, R. (2020, November 11). Social media is making a bad political situation worse. *Vox*. <https://www.vox.com/recode/21534345/polarization-election-social-media-filter-bubble>
- Isaac, M., & Ember, S. (2016, June 29). Facebook to change news feed to focus on friends and family. *The New York Times*.
<https://www.nytimes.com/2016/06/30/technology/facebook-to-change-news-feed-to-focus-on-friends-and-family.html>
- Levendusky, M. (2014, February 3). Are Fox and MSNBC polarizing America? *The Washington Post*.
<https://www.washingtonpost.com/news/monkey-cage/wp/2014/02/03/are-fox-and-msnbc-polarizing-america/>
- Emamzadeh, A. (2019). *The Psychology of "Us-vs-Them."* Psychology Today. <https://www.psychologytoday.com/us/blog/finding-new-home/201908/the-psychology-us-vs-them>
- POLITICO. (2013, December 9). *Is MSNBC worse than Fox News?* POLITICO. <https://www.politico.com/blogs/media/2013/12/is-msnbc-worse-than-fox-news-179175>
- Bauder, D. *Fox News' Tucker Carlson criticized for saying Democrats hate America*. (2020, July 7). Chicago Sun-Times. <https://chicago.suntimes.com/politics/2020/7/7/21316740/fox-news-tucker-carlson-demos-crats-hate-america>
- Jones, T. (2021, February 1). *No, Fox News and MSNBC are not the same thing*. Poynter. <https://www.poynter.org/newsletters/2021/no-fox-news-and-msnbc-are-not-the-same-thing/>
- Stanley, A. (2012, August 31). How MSNBC became Fox's liberal evil twin. *The New York Times*.

<https://www.nytimes.com/2012/08/31/us/politics/msnbc-as-foxs-liberal-evil-twin.html>
 Perspective | “Birthing centers for polarizing rhetoric”: The outsize influence of Fox, CNN and MSNBC. (n.d.). *Washington Post*. Retrieved January 4, 2022, from

https://www.washingtonpost.com/lifestyle/style/birthing-centers-for-polarizing-rhetoric-the-outsize-influence-of-fox-cnn-and-msnbc/2019/05/23/2bcc429a-7cbe-11e9-8ede-f4abf521ef17_story.html

Business, B. S., CNN. (2021, June 8). “*We turned so far right we went crazy:*” *How Fox News was radicalized by its own viewers*. CNN.

<https://www.cnn.com/2021/06/08/media/fox-news-hoax-paperback-book/index.html> Nast, C. (2020, March 11). *Ezra Klein’s “Why we’re polarized” and the drawbacks of explainer journalism*. *The New Yorker*.

<https://www.newyorker.com/books/under-review/ezra-kleins-why-were-polarized-and-the-drawbacks-of-explainer-journalism>

Carothers, T. (2019, September 24). *Democracies Divided: The Global Challenge of Political Polarization* (A. O’Donohue, Ed.) [Review of *Democracies Divided: The Global Challenge of Political Polarization*]. Brookings.edu.

https://www.brookings.edu/wp-content/uploads/2019/04/democracies-divided_introduction-1.pdf

Roscini, F. (n.d.). *How The American Media Landscape is Polarizing the Country*. Frederick S. Pardee School of Global Studies.

<https://sites.bu.edu/pardeeatlas/advancing-human-progress-initiative/back2school/how-the-american-media-landscape-is-polarizing-the-country/>

Dimock, M., Richard W. (2020, November 13). *America is exceptional in the nature of its political divide*. Pew Research Center.

<https://www.pewresearch.org/short-reads/2020/11/13/america-is-exceptional-in-the-nature-of-its-political-divide/>

<https://www.pewresearch.org/short-reads/2020/11/13/america-is-exceptional-in-the-nature-of-its-political-divide/>

Using Machine Learning to Predict the Success of Startup Businesses

By Ahmad Zayan Sadek

Abstract

Startups present a high-risk investment opportunity, but they also offer the potential for substantial profits and returns. Given the significant risk of failure, it would be valuable to develop a program that can predict a startup's success by comparing its financial statements with those of established, successful businesses. Startups are entrepreneurial ventures in their early stages, often created to solve real-world problems. Due to their potential to address societal needs, many startups attract investors and funders eager to capitalize on their growth opportunities. My research paper explores the potential success of startups based on financial metrics. Developing a model to predict and analyze the future success of these companies can provide useful guidance for investors, helping them decide whether to invest in startups, which are inherently riskier than established businesses.

The study synthesizes insights from various sources, including ResearchGate, 123Financials, and AIContentfy, to answer the research question: *How can we determine the potential success of a startup company based on its financial statements?* The paper begins with a literature review, presenting diverse perspectives on predicting startup success. ResearchGate highlights the use of specific financial ratios, 123Financials suggests a broader approach incorporating Key Performance Indicators (KPIs) and qualitative factors, and AIContentfy examines the role of Artificial Intelligence (AI) and Machine Learning (ML) in forecasting trends and startup growth.

The methodology involves collecting financial data from S&P 500 companies using Yfinance and calculating any additional, missing financial metrics. These metrics were added to our dataset and analyzed using various models to identify any correlations or relationships. The data was then processed through multiple classification models to assess accuracy, F1 and confusion matrix scores. After analyzing the performance of these models, Bagging and Boosting emerged as the most accurate and reliable.

The research concludes that while financial metrics are crucial, incorporating AI and ML can significantly enhance predictive accuracy. However, challenges such as reliance on historical data, sample size, model optimization, complexity, and geographical constraints must be addressed to improve the model's effectiveness. These factors limit the quality and accuracy of our model and overall investigation.

The findings provide valuable insights for investors, venture capitalists, and financial institutions, underscoring the importance of financial metrics and advanced modeling techniques in assessing the potential success of startups. Future research will focus on refining the model, addressing the identified limitations, and exploring practical applications for evaluating future startups.

Introduction

The core issue of our investigation is the inherent risk associated with investing in startups, which, despite their potential for high profits and returns, have a significant chance of failure. Therefore, it would be valuable to develop a program that can predict the potential success of a startup by comparing its financial statements with those of established, successful companies.

Although the tech startup sector has experienced a general downturn, AI-focused startups have been thriving since late 2022. Between April and June alone, investors funneled \$27.1 billion into U.S. AI startups, accounting for nearly half of all startup funding during that period. Notable funding rounds included \$1.1 billion for CoreWeave, \$1 billion for Scale AI, and \$6 billion for xAI. The surge in investment was sparked by OpenAI's release of ChatGPT, which ignited widespread interest in generative AI. While AI startups face high computing costs, their growth potential continues to attract substantial funding. However, competition from large tech companies may impact future funding opportunities.

To assess the fundamental strength of these companies, we compared their financial metrics with those of robust S&P 500 companies. We used Yfinance to download and import the necessary data into our Python code. From there, we identified key financial metrics essential for determining fundamentally strong stocks, calculated these metrics using the imported data, and added the results to our dataset. We then established thresholds for each metric that had to be met for a stock to be considered fundamentally strong. Graphs and plots were generated to analyze the relationships between different metrics, ultimately revealing little to no correlation among them.

The industry utilizes various machine learning models to predict the success of startups and companies, including logistic regression for binary classification, decision trees and random forests for clear decision paths and improved accuracy, support vector machines (SVM) for high-dimensional spaces, gradient boosting machines like XGBoost for high performance, and neural networks, including deep learning models, for identifying complex patterns in large datasets. These models help in pinpointing key features and making accurate predictions based on historical data.

In our research, we explored a range of classification models, including Logistic Regression, K-Nearest Neighbors (KNN), SVM, Decision Trees, and ensemble methods such as Bagging, Random Forest, and Boosting. We evaluated these models based on their F1 scores, accuracy scores, and confusion matrix scores to determine which were most suitable and effective for our investigation. Ultimately, Bagging and Boosting models demonstrated superior performance across all metrics, proving to be the most effective in predicting startup success.

Literature Review

Source 1: ResearchGate

The first source from ResearchGate supports our investigation by exploring how specific financial metrics can predict startup success, particularly in the context of Indonesian startups.

The study uses a quantitative approach with a cross-sectional design, focusing on three key financial ratios: debt-to-equity, activity, and EBITDA ratios. It hypothesizes that these ratios can indicate the level of financial distress risk in startups, with a particular emphasis on the debt-to-equity ratio. The research suggests that a higher debt-to-equity ratio increases the likelihood of financial distress, thus impacting the startup's success.

While this study provides valuable insights, its focus on Indonesian startups means it may not fully represent the global startup landscape, where economic conditions and industry standards can vary. Despite this limitation, the article strongly relates to our research question, demonstrating that financial metrics can indeed be used to predict the potential success of startups. It offers a real-world example of how financial ratios are directly linked to a startup's financial health, suggesting potential factors we could incorporate into our predictive model.

Source 2: 123Financials

The 123Financials source presents a viewpoint that contradicts our thesis by arguing that predicting the potential success of startups should not rely solely on financial metrics. According to this source, stakeholders and investors must consider a broader range of factors to fully understand a company's performance and potential. These factors include Key Performance Indicators (KPIs) for operational efficiency, market size, competition analysis, the intended use of funds, leadership team capabilities, and legal compliance. The source emphasizes that these elements are crucial for assessing a startup's vision and execution.

However, while the source focuses on positive indicators of success, it does not adequately address the factors that might signal potential failure or risks within a startup. This omission suggests that a startup could appear promising based on these indicators but still harbor significant risks that make it challenging to predict its success accurately.

By highlighting the importance of non-financial factors, this source directly challenges the methodology we are using to answer our research question. It suggests that a more comprehensive approach, beyond just financial statements and metrics, is necessary to evaluate a startup's potential success effectively. This source offers us a further lead in terms of future research.

Source 3: AIContentfy

AIContentfy explores how Artificial Intelligence (AI) and Machine Learning (ML) can be pivotal in predicting startup success, which aligns closely with our model's focus. The source argues that AI and ML can analyze vast datasets to predict future trends, customer behavior, and other critical factors, enabling startups to make more informed decisions and allocate resources more effectively. This technological approach not only aids in creating innovative products and services but also enhances customer experiences, driving long-term growth.

However, the source also highlights potential challenges. As AI and ML become more widely adopted, increased competition could drive innovation but also raise the bar for success. Additionally, startups must address ethical concerns, such as data privacy and security, to ensure

responsible use of these technologies. While AI and ML offer powerful tools for predicting success, the source emphasizes that these technologies are not without limitations, particularly regarding data quality, interpretability, and ethical considerations.

Overall, AIContentfy supports our investigation by showcasing the significant role AI and ML can play in predicting startup success based on historical and real-time data, while also pointing out areas that require careful attention to maximize the effectiveness of these tools.

Methodology

To collect data, we used Yfinance and imported the S&P 500 companies' financial statements, balance sheets and cash flow statements. We needed them to calculate certain financial metrics that we used to determine whether certain startup companies are financially or fundamentally strong compared to the S&P 500 companies.

Financial Metric	Formula	Description	Threshold
Net Profit Margin	Net Profit / Total Revenue	Percentage of total income kept after all expenses and taxes are paid.	Above 0.2
Debt-to-Equity Ratio	Total Debt / Total Equity.	Shows how much of a company is owned by creditors compared with how much shareholder equity is held by the company.	Below 2
Current ratio	Total current assets / total current liabilities	A liquidity ratio that measures a company's ability to pay short-term obligations or those due within one year.	Above 1
Cash ratio	Dividing cash by current liabilities	The metric calculates the ability of a company to repay its short-term debt with cash or near-cash resources.	Between 0.5 and 1
Return on equity	Net income / average shareholders' equity	A gauge of a corporation's profitability and how efficiently it generates those profits. The higher the ROE, the better a	Above 0.12

		company is at converting its equity financing into profits.	
Asset turnover ratio	Net revenue / average total assets	A measurement that shows how efficiently a company is using its owned resources to generate revenue or sales.	2.5 or greater
Long Term Debt to Net Income Ratio	Long Term Debt / Net Income	This enables lenders to measure borrowers' ability to manage monthly repayments.	Less than 5

We used Yfinance to import income statements, balance sheets and cash flow statements onto our Python code. We constructed our code to get the specific financial metrics within a period and print it into a table for our viewing. Consequently, to get some of our essential financial metrics (listed above in the table), we had to use formulas to calculate them, and input them into new columns, based on the information that was already given in the table, as these metrics weren't included in the data frame from YFinance. We then set equalities and thresholds in our code; certain financial metrics had to be higher or/and lower than certain numbers. If all requirements were met, then the stock was classified as 1, which means it's fundamentally strong and we can predict future success for the startup. After this, we created numerous graphs and models to analyze the data such as: scatterplots, histograms, boxplots and a correlation heatmap.

Next, we ran classification models to determine which companies were fundamentally strong, which would help us predict potential success. Consequently, analyzed the classification models to identify which model was the most accurate in identifying whether companies are fundamentally strong.

We ran classification models such as Logistic Regression, K-Nearest Neighbors (KNN), SVM, Decision Trees, and ensemble methods such as Bagging, Random Forest, and Boosting; Boosting and Bagging turned out to be the most accurate and effective classification models.

Listed below is a screenshot of the data frame. The 0s and 1s indicate whether each stock is fundamentally strong or not. If all the thresholds incorporated in our code were successfully met, you would see a 1 under the strength of the code; however, if just one or more of the thresholds for the financial metrics weren't met, it's classified as not fundamentally strong and has a 0.

	Net Profit Margin	Debt-to-Equity Ratio	Current Ratio	Cash Ratio	Return On Equity	Asset Turnover	Long Term Debt To Net Income Ratio	Strength
2023-12-31	-0.214039	3.402202	1.070733	0.331111	-1.45517	0.646125	-1.871051	0
2023-12-31	0.144466	0.084147	1.597115	0.632286	0.301778	1.198793	0.210744	1
2023-12-31	0.142666	0.402503	1.637867	0.365508	0.148253	0.547832	2.376201	1
2023-12-31	0.089528	5.711744	0.872123	0.583018	0.469402	0.403219	10.732661	0
2023-08-31	0.107181	0.118017	1.286344	0.489532	0.28745	1.251075	0.006271	1
...
2023-12-31	0.0827	0.234866	1.761451	0.256689	0.059306	0.457051	3.724138	0
2023-12-31	0.225692	-1.531051	1.250964	1.032106	-0.203232	1.135612	6.950532	0
2023-12-31	0.064572	0.795125	1.046308	-0.050946	0.097497	0.62743	6.915541	0
2023-12-31	0.138487	0.473459	1.61318	0.421397	0.082048	0.343966	4.895117	0
2023-12-31	0.274345	1.353436	3.357861	0.858126	0.469081	0.598068	2.800341	1

503 rows x 8 columns

Fig 1: Output of Code showing fundamental strength of stocks

Modeling and Analysis

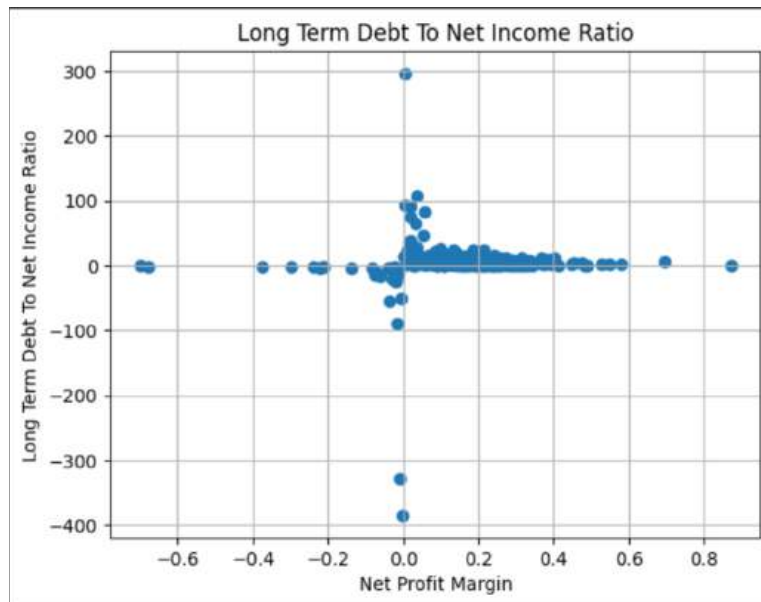


Fig 2: Scatter Plot output from code

Scatter plots are graphs in which the values of two variables are plotted along two axes. The y values (Long Term Debt to Net Income) don't increase or decrease as the x values increase (Net Profit Margin). We can therefore conclude that there is no relation between these variables, which is what we want for an accurate model. There are outliers which seem to be vastly spread from where the data is clustered on the top, bottom, left and right, and once again since the relationship isn't linear, it proves that there is no correlation.

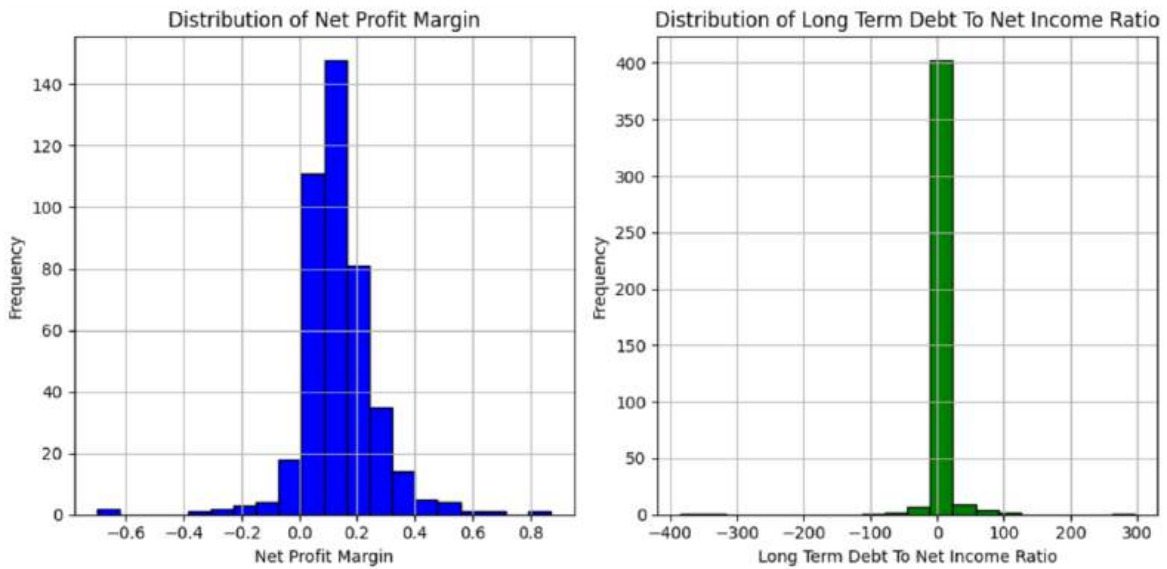


Fig 3: Histogram output from code

Histograms are graphical representations of data distributions. Our histograms didn't show correlation between more than one variable. They allowed us to measure and compare the center, spread, median and range. Clearly, the standard deviation of the graph on the right is much greater than the graph on the left. Both graphs seem to have outliers: the graph displaying Net Profit Margin has a center at around 0.1 and range of about 1.5, whereas the graph displaying Long Term Debt to Net Income Ratio has a center at around 10 and a much greater range of around 700.

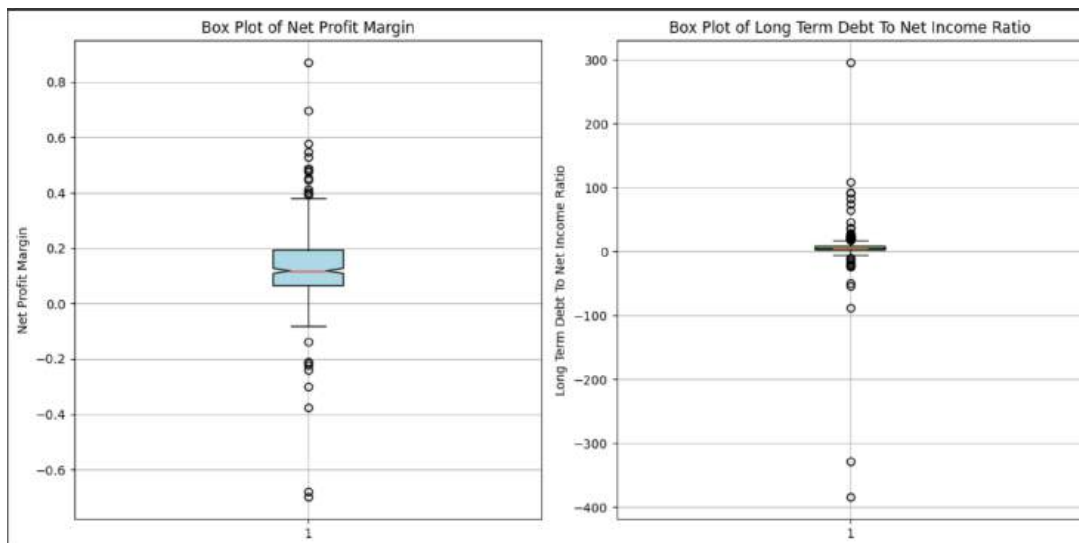


Fig 4: Boxplot output from code

Boxplot is a chart that is used to visualize how a given data (variable) is distributed using quartiles. As the data appears randomly scattered with no real association or correlation, this classification model further proves that there is little influence of the variables on each other, suggesting the accuracy and reliability of the results. Both graphs and their outliers are more visible in the box plot, but remain consistent with the mean, range and median displayed in the histogram. There is consistency in the frequency of outliers.

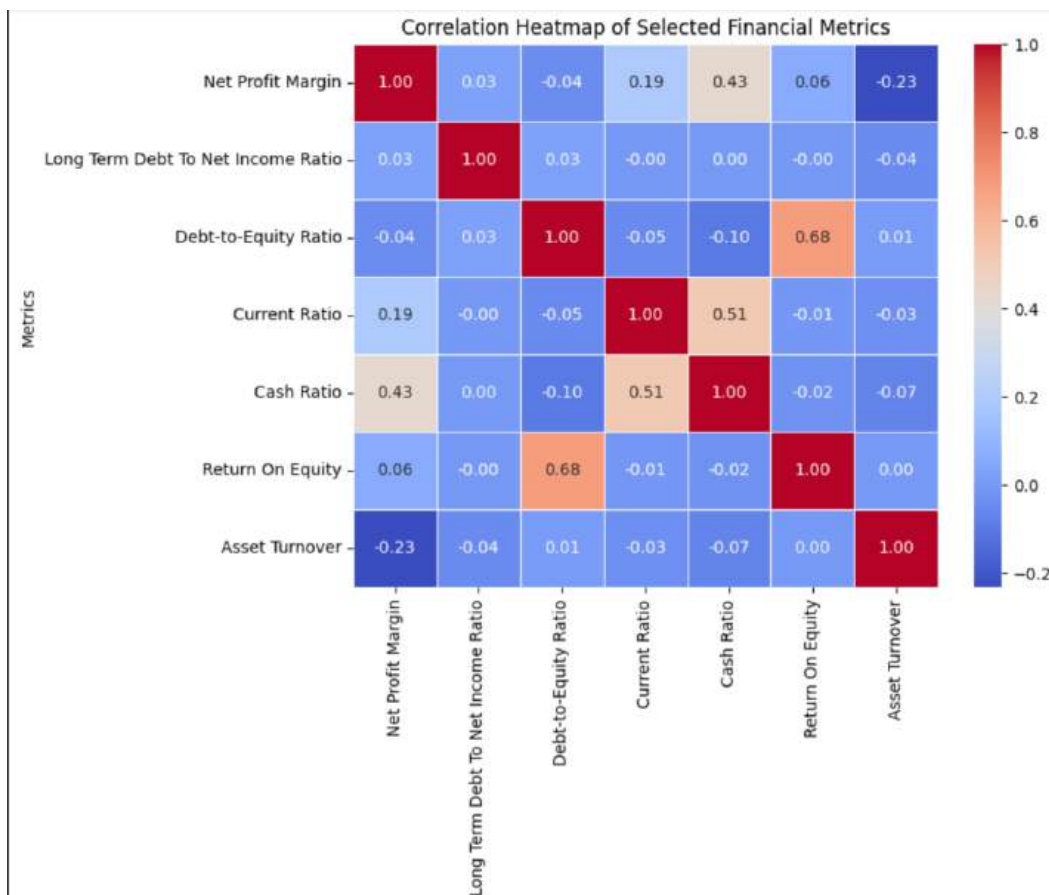


Fig 5: Correlation Heatmap output from code

In a correlation heatmap, each cell's color represents the strength of the correlation between two variables. Brighter colors (red) indicate a stronger positive correlation, while darker colors (blue) indicate a stronger negative correlation. Looking at this correlation heat map, we can clearly see that there is little to no correlation between the variables. The correlation between the variables is low and this is proven through the fact that the strongest correlation is 0.68 (a high threshold would be anything less than -0.8 or greater than 0.8), which is still relatively low, between return on equity and debt-to-equity ratio. Therefore, we can conclude that this classification model also illustrates that there is little to know association and connection between any of the financial metrics.

Missing Data

We removed the following S&P 500 companies' data to allow our model to successfully run and proceed with our investigation.

Standardize data

Standardizing data in Python is a crucial preprocessing step in machine learning and data analysis that ensures all features have a mean of zero and a standard deviation of one. This process helps improve the performance and convergence of optimization algorithms like gradient descent, ensures that all features contribute equally to distance calculations in models like KNN and clustering, and enhances the interpretability of linear model coefficients. Standardization is easily accomplished using the StandardScaler from the scikit-learn library, which scales the features by removing the mean and scaling to unit variance. We standardized the data of the financial metrics to get them all into the same units, which would help further interpretation and comparisons between our financial metrics when it came to analyzing the correlation between them in our classification models. It makes the data easier to compare and identify if there is any relationship or correlation between metrics.

Removing correlated data

We didn't have to remove any correlated data as there was little correlation between the financial metrics regardless. This is proved by our classification models, specifically our correlation heat map, which conveyed little to no correlation between any of the financial metrics included in our code.

Remove unimportant feature importance

All of our features were essential and therefore none were removed.

Class balance

Both classes had slightly more fundamentally strong than weak outcomes, which suggests that there is reduced bias. S&P 500 is widely regarded as the best single gauge of large-cap U.S. equities. The index includes 500 leading companies and covers approximately 80% of available market capitalization.

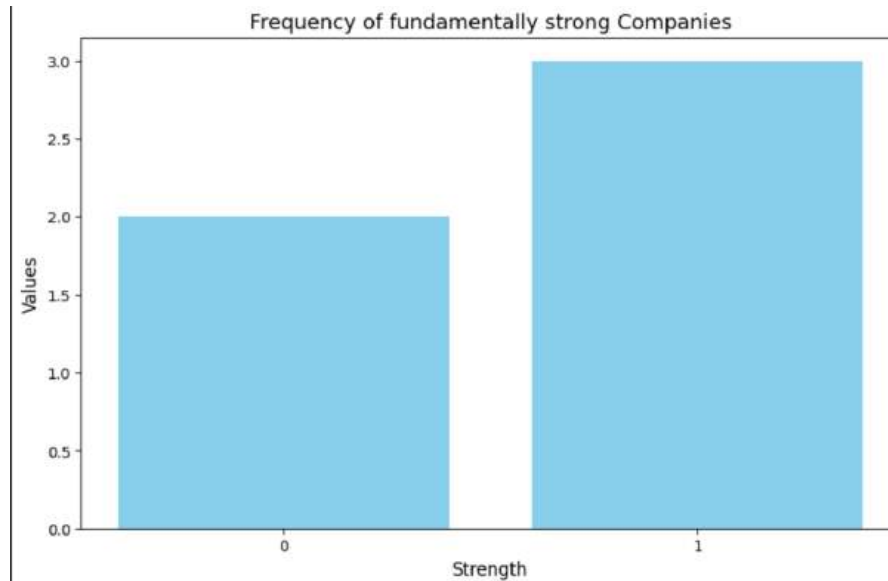


Fig 6: Class balance of strong (1) to weak (0) companies

F1 Score

Indicates the model's balance between precision and recall. Higher scores reflect better performance in distinguishing between successful and unsuccessful stocks.

Accuracy Score

Shows the proportion of correctly classified predictions. Higher accuracy means the model correctly predicts stock success more often.

Confusion Matrix

Provides a detailed view of true positives, false positives, true negatives, and false negatives. A higher number of accurate predictions relative to inaccuracies indicates better model performance

Logistic regression

Logistic regression is a binary classification technique that models the relationship between a binary dependent variable and one or more independent variables using the logistic function to map predicted values to probabilities. It starts with a linear combination of input features, then applies the logistic (sigmoid) function to produce a value between 0 and 1, representing the probability of belonging to one of the two classes. Advantages of logistic regression include its simplicity, interpretability, and effectiveness with linearly separable data. However, it may struggle with non-linear relationships and is sensitive to outliers and multicollinearity.

KNN

K-Nearest Neighbors (KNN) is a non-parametric, instance-based learning algorithm used for classification and regression. It classifies data points based on the majority class of their k-nearest neighbors in the feature space, where k is a user-defined parameter. The algorithm calculates the distance between the input data and the training samples using a distance metric, such as Euclidean distance, and assigns the most common class among the nearest neighbors. Advantages of KNN include its simplicity, ease of implementation, and effectiveness with small datasets. However, it can be computationally expensive, sensitive to irrelevant features and the choice of k, and struggles with high-dimensional data.

SVM

Support Vector Machine (SVM) is a powerful supervised learning algorithm used for classification and regression tasks. It works by finding the hyperplane that best separates data points of different classes in the feature space, maximizing the margin between the closest points of the classes (support vectors). SVM can efficiently perform non-linear classification using kernel functions to map data into higher dimensions. Advantages of SVM include its effectiveness in high-dimensional spaces, robustness to overfitting, and versatility with various kernel functions. However, it can be less effective with noisy data, requires careful parameter tuning, and can be computationally intensive for large datasets.

Decision Trees

Decision Trees are a versatile and intuitive supervised learning algorithm used for classification and regression tasks. They work by recursively splitting the data into subsets based on the feature that provides the highest information gain or lowest Gini impurity, creating a tree-like model of decisions. Each internal node represents a feature, each branch represents a decision rule, and each leaf node represents an outcome. Advantages of Decision Trees include their simplicity, interpretability, and ability to handle both numerical and categorical data. However, they can be prone to overfitting, sensitive to noisy data, and may not perform well with small datasets.

Ensemble Methods

Bagging: Bagging, short for Bootstrap Aggregating, is an ensemble learning technique that improves the stability and accuracy of machine learning algorithms, particularly for classification and regression tasks. It works by generating multiple subsets of the training data through random sampling with replacement and training a base model (e.g., decision trees) on each subset. The final prediction is made by aggregating the predictions of all base models, typically through majority voting for classification or averaging for regression. Advantages of bagging include reduced variance, improved accuracy, and robustness to overfitting. However, it can be computationally intensive and may not significantly improve performance if the base models are already strong.

Random Forest: Random Forest is an ensemble learning method that enhances the performance of decision trees by constructing a multitude of them during training and outputting the mode of the classes (classification) or mean prediction (regression) of the individual trees. It operates by using bagging (bootstrap aggregating) and randomly selecting subsets of features for each tree, ensuring diverse and uncorrelated trees. Advantages of Random Forest include high accuracy, robustness to overfitting, and the ability to handle large datasets and high-dimensional data. However, it can be computationally expensive and less interpretable than single decision trees.

Boosting: Boosting is an ensemble learning technique that improves the accuracy of weak learners by sequentially training them, with each new model focusing on the errors made by the previous ones. It combines the outputs of these weak models to create a strong predictive model. Common boosting algorithms include AdaBoost, Gradient Boosting, and XGBoost. Advantages of boosting include high predictive accuracy, the ability to handle complex data patterns, and reduced bias. However, it can be prone to overfitting, computationally intensive, and sensitive to noisy data and outliers.

Classification Models	What they're suitable for:
Logistic Regression	<ul style="list-style-type: none"> - Realistic for binary classification problems. - Suitable for problems with a linear decision boundary and where interpretability is important. - Works well with smaller datasets and when the relationship between features and the target variable is relatively simple.
KNN	<ul style="list-style-type: none"> - Realistic for small to moderately sized datasets. - Suitable when simplicity and ease of implementation are prioritized. - Can be computationally intensive for large datasets and may struggle with high-dimensional data.
SVM	<ul style="list-style-type: none"> - Realistic for both binary and multiclass classification. - Effective in high-dimensional spaces and for problems with a clear margin of separation. - Can be computationally intensive and require careful parameter tuning.

Decision Trees	<ul style="list-style-type: none"> - Realistic for both classification and regression tasks. - Suitable for problems where interpretability and simplicity are important. - Can overfit on small datasets but are quick to train and understand.
Bagging	<ul style="list-style-type: none"> - Realistic for improving the performance of unstable models like decision trees. - Suitable when variance reduction is needed. - Can be computationally intensive due to multiple model training.
Random Forest	<ul style="list-style-type: none"> - Realistic for both classification and regression tasks. - Suitable for large datasets and high-dimensional spaces. - Provides high accuracy and robustness to overfitting.
Boosting	<ul style="list-style-type: none"> - Realistic for high accuracy and complex data patterns. - Suitable when bias reduction is needed. - Can be computationally intensive and sensitive to noisy data.

Random Forest is highly effective due to its ability to handle large datasets and complex interactions between features. It provides high accuracy, robustness to overfitting, and is relatively easy to interpret compared to some other models. In addition, Boosting methods, such as Gradient Boosting or XGBoost, are also suitable as they excel in capturing complex patterns and improving model performance by focusing on errors from previous models. They can deliver high predictive accuracy but require careful tuning and can be sensitive to noisy data.

Result/Finding

Based on the provided metrics, here's a summary to help determine the best and most accurate model:

Classification Model	F1 Score	Confusion Matrixes	Test Accuracy Score
Logistic Regression	0.83	71:15	0.8255813953488372

KNN	0.91	78:8	0.9069767441860465
SVM	0.85	73:13	0.8488372093023255
Decision Trees	0.98	84:2	0.9767441860465116
Bagging	0.99	85:1	0.9883720930232558
Random Forest	0.97	83:3	0.9651162790697675
Boosting	0.99	85:1	0.9883720930232558

Best Model Analysis:

1. F1 Score:
 - Measures the balance between precision and recall. Higher values indicate better performance.
 - Best Models: Bagging and Boosting (0.99)
2. Test Accuracy Score:
 - Represents the proportion of correctly classified instances out of the total instances.
 - Best Models: Bagging and Boosting (0.988)
3. Confusion Matrix:
 - Shows the number of accurate vs. inaccurate predictions.
 - Best Models: Bagging and Boosting (85:1)

Bagging and Boosting are the best models across all metrics. They both achieve the highest F1 Score (0.99) and Accuracy Score (0.988), and they have the most accurate predictions according to the confusion matrix (85:1). Given these results, Bagging and Boosting are the top choices, with Boosting slightly more favorable in terms of F1 Score, but both perform exceptionally well. In summary, high scores in these metrics generally mean the model is better at distinguishing between classes, making fewer errors, and providing reliable predictions. This is why Bagging and Boosting, with their high scores across these metrics, are considered the best performing models for the given task.

Limitation

1. Dependence on Historical Data

Limitation: The model depends heavily on historical financial data to predict future stock success. This reliance means that any sudden market changes, new economic conditions, or unforeseen events not captured in the historical data may lead to inaccurate predictions. For instance, a model trained on past economic conditions may fail to predict market downturns triggered by new global events or innovations.

Impact: This limitation can lead to biased or outdated predictions, as the model may not adapt well to new trends or disruptions. In your research paper, this could be a significant limitation, as it affects the robustness and reliability of the model's predictions.

Further works: To mitigate this limitation, incorporating real-time data and continuously updating the model with recent information can help. Additionally, using techniques like transfer learning could allow the model to adapt to new data more effectively. Implementing these changes would enhance the model's ability to respond to current market conditions and improve the overall accuracy of predictions.

2. Sample Size Limitations

Limitation: The accuracy of predictions can be limited by the size of the sample used to train the model. A smaller dataset may not capture the full range of market conditions, leading to less reliable predictions. For example, if the model is trained on a limited number of stocks or a narrow time frame, it may not generalize well to unseen stocks or future market scenarios.

Impact: This limitation impacts the model's performance and generalizability. In your research paper, it could be noted that a small sample size may lead to overfitting or underfitting, which could skew the model's effectiveness.

Further works: Increasing the sample size by including more stocks and extending the time frame of the data would provide a more comprehensive training dataset. This would improve the model's ability to generalize across different market conditions and enhance prediction accuracy. This approach would demonstrate a commitment to robust and reliable model development.

3. Need for Model Optimization

Limitation: The current model may not be fully optimized, which can impact its performance. Without fine-tuning parameters and employing advanced techniques, the model might not achieve its best possible performance. For instance, using default hyperparameters may result in suboptimal predictions.

Impact: This limitation could lead to less accurate predictions and inefficient model performance. In the research paper, this may be discussed as a potential area for improvement, indicating that the model's full potential has not yet been realized.

Further works: Systematic tuning of model parameters, such as through grid search or randomized search, and applying techniques like cross-validation can help optimize performance. Enhancing the model in this way would likely lead to improved accuracy and reliability, which would be a positive addition to the research findings.

4. Complexity of Models

Limitation: The current model may be based on relatively simple algorithms, which might not capture complex relationships within the data. More sophisticated models, such as neural networks, could potentially offer better performance by learning non-linear patterns and interactions between features.

Impact: Using simpler models might limit the model's ability to capture intricate patterns in financial metrics, which can affect the accuracy of predictions. In the research paper, this limitation could be addressed as a potential area for future exploration.

Further works: Implementing more complex models, such as deep learning neural networks, could improve the model's ability to capture complex relationships and interactions in the data. This advancement would enhance predictive accuracy and offer deeper insights into stock success, adding significant value to the research.

5. Geographical Limitation

Limitation: The model may currently focus on stocks from a specific region or market, which limits its applicability to global stock markets. Regional biases and differences in financial regulations or market behaviors can affect the model's performance when applied to stocks from other regions.

Impact: This limitation can restrict the model's generalizability and applicability across different markets. In the research paper, this could be noted as a limitation affecting the model's broader relevance and usefulness.

Further works: Expanding the dataset to include stocks from multiple regions and markets would enhance the model's ability to generalize and adapt to different financial environments. This expansion would improve the model's applicability and robustness, making it more relevant to a global audience and enhancing the research's impact.

By addressing these limitations, the model's performance, reliability, and applicability can be significantly improved, leading to more accurate predictions and a stronger research contribution.

Conclusion

After running our Python code and alternative classification models, we can conclude that we created a reliable code that successfully analyzes the unpredictable and unstable growth of startups.

In the context of predicting stock success based on fundamental strength and financial metrics, analysis and modeling are crucial for transforming raw financial data into actionable insights. The process involves several key steps:

- **Data Collection and Preprocessing:** Gathering historical financial data and fundamental metrics of stocks. This includes cleaning and normalizing data to ensure it is suitable for analysis.

- Feature Selection: Identifying and selecting relevant features that are most indicative of stock success, such as earnings growth, P/E ratio, and debt levels.
- Model Building: Applying various classification algorithms (e.g., Logistic Regression, KNN, SVM, Decision Trees, Bagging, Random Forest, Boosting) to train models on the dataset. Each model learns patterns and relationships from the data to predict stock success.
- Model Evaluation: Using metrics such as F1 Score, Accuracy Score, and confusion matrices to assess model performance. This helps in understanding how well each model predicts stock success and identifies areas for improvement.
- Comparison and Selection: Comparing the performance of different models based on evaluation metrics to select the most effective one. This involves analyzing which model provides the highest accuracy and balance between precision and recall.

Overall Findings

Bagging and Boosting emerged as the top-performing models due to their high F1 Scores (0.99), Accuracy Scores (0.988), and confusion matrix results (85:1). They are effective at capturing complex patterns and providing reliable predictions. Random Forest and Decision Trees also performed well but slightly lower than Bagging and Boosting. Logistic Regression, KNN, and SVM had lower scores in comparison, reflecting less effectiveness in predicting stock success accurately.

Impact on Research

The analysis and modeling provide a structured approach to understanding which factors are most indicative of stock success and how different algorithms perform in predicting this success. By identifying the most effective models and understanding their performance, the research can offer valuable insights into stock prediction and guide investment decisions. Enhancing the model's performance and addressing limitations will contribute to more accurate and reliable stock predictions, advancing the field of financial forecasting.

Our next steps revolve around using our model to begin predicting the potential success of startups as well as possible go back and work on refining our model through implementing solutions to the limitations we identified previously. I further plan on sharing my model with Venture Capitalist firms that I have gained internship experience from and gather their thoughts as well as see how and if we could alter and adjust our program to make it more effective, efficient, or useful in several ways.

Works Cited

- Aprio Publisher. "Investing in AI: Opportunities and Risks for Startups and Investors - Aprio." *Aprio*, 24 June 2024, www.aprio.com/investing-in-ai-opportunities-and-risks-for-startups-and-investors-ins-article-tech/#:~:text=Understanding%20the%20Risks%3A%20AI%20investments. Accessed 4 Aug. 2024.
- Arjuna Aqa. "Top 5 Risks for AI Startups." *Arjuna Aqa*, Arjuna Aqa, 25 Oct. 2023, www.arjunaqa.com/post/top-5-risks-for-ai-startups. Accessed 4 Aug. 2024.
- Bhandari, Aniruddha. "Feature Scaling | Standardization vs Normalization." *Analytics Vidhya*, 3 Apr. 2020, www.analyticsvidhya.com/blog/2020/04/feature-scaling-machine-learning-normalization-standardization/.
- "Choosing the Best Machine Learning Classification Model and Avoiding Overfitting." *Www.mathworks.com*, www.mathworks.com/campaigns/offers/next/choosing-the-best-machine-learning-classification-model-and-avoiding-overfitting.html#:~:text=Classification%20models%20are%20Used%20to.
- Communications, KeAi. "Scientists Develop AI to Predict the Success of Startup Companies." *Phys.org*, phys.org/news/2021-09-scientists-ai-success-startup-companies.html.
- Erickson, Jeffrey. "11 Common Challenges of AI Startups & How to Address Them." *Oracle.com*, Oracle, 12 Jan. 2024, www.oracle.com/artificial-intelligence/ai-startup-challenges/. Accessed 4 Aug. 2024.
- "Financials and Accounting Blog." *Financials and Accounting Blog*, 22 Sept. 2023, www.123financials.com/insights/startup-financial-statements-what-investors-and-stakeholders-look-for/.
- Griffith, Erin. "Investors Pour \$27.1 Billion into A.I. Start-Ups, Defying a Downturn." *The New York Times*, 3 July 2024, www.nytimes.com/2024/07/03/technology/ai-startups-funding.html.
- Hayes, Adam. "The Risk and Rewards of Investing in Startups (GOOG)." *Investopedia*, 2019, www.investopedia.com/articles/personal-finance/041315/risk-and-rewards-investing-startups.asp.
- Marciano, Jonathan. "AI: An Extension of the FP&a Brain: Didi Gurfinkel Co-Founder and CEO of Datarails - Datarails." *Datarails*, 30 Jan. 2024, www.datarails.com/ai-an-extension-of-the-fpa-brain-didi-gurfinkel-co-founder-and-ceo-of-datarails/. Accessed 4 Aug. 2024.
- NYT. "Investors Pour \$27.1 Billion into AI Startups, Defying a Downturn." *@Bsendia*, Business Standard, 3 July 2024, www.business-standard.com/world-news/investors-pour-27-1-billion-into-ai-startups-defying-a-downturn-124070301277_1.html. Accessed 4 Aug. 2024.

- Porzucek, Natalia. "3 Problems Startup Founders Do Not Know They Have When Using AI/ML." *Punktum*, 15 Nov. 2023, punktum.net/insights/3-problems-startup-founders-do-not-know-they-have-when-using-ai-ml/. Accessed 4 Aug. 2024.
- "Scientists Develop AI to Predict the Success of Startup Companies." *EurekaAlert!*, 7 Sept. 2021, www.eurekaalert.org/news-releases/927669. Accessed 4 Aug. 2024.
- Stobierski, Tim. "13 Financial Performance Measures Managers Should Monitor | HBS Online." *Harvard Business School Online*, 5 May 2020, online.hbs.edu/blog/post/financial-performance-measures.
- team, AIContentfy. "Investing in Startups: Risks and Rewards." *AIContentfy*, AIContentfy Oy, 8 Feb. 2023, aicontentfy.com/en/blog/investing-in-startups-risks-and-rewards. Accessed 4 Aug. 2024.
- "The Impact of AI and Machine Learning on Startups." *AIContentfy*, aicontentfy.com/en/blog/impact-of-ai-and-machine-learning-on-startups.
- "The Potential of Working Hypotheses for Deductive Exploratory Research." ResearchGate, https://www.researchgate.net/publication/346773952_The_potential_of_working_hypotheses_for_deductive_exploratory_research.
- "The Two Sides of the Coin - The Success and Failure of the Business." ResearchGate, https://www.researchgate.net/publication/364044899_The_Two_Sides_of_the_Coin_-_The_Success_and_Failure_of_the_Business.
- "Yahoo Finance." *Yahoo.com*, 2024, finance.yahoo.com/?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAAAGAtlIZ19kroyZoaFzi8X7P6uId1PkItZJJnwS5114r-15uQzROpKtzbIZ6evGm-L0zsSAm8k5E7pRk4MnWNBesDVcTaArrBEFGETb8-GEDGYZ3hI9oiZUZRMndYoHxVIEChg2GDZWQ1N8FQLwNtRspDdVilY7cE-8K60aBofxO. Accessed 4 Aug. 2024.

Systems of Power: Analyzing Apartheid's Legacy of White Supremacy in South Africa

By Saawan Duvvuri

On June 4, 1948, the Afrikaner-led National Party came to power in South Africa, winning an election that would significantly impact the role of race in South Africa for decades to come.¹ Although racial segregation had existed in South Africa for almost a century, it was not until the rise of the National Party that race would cement itself into almost every aspect of South African society.² “Apartheid,” meaning “Separateness” in the Afrikaner language, was the ideology introduced by the National Party, with the ultimate goal of separating Blacks from the South African state and engraving White supremacy into South Africa’s political, economic, and social structures.³ From 1948 to Apartheid’s end in 1994, the South African government passed legislation enforcing the subjugation of South Africa’s Black population while also oppressing South Africa’s Coloureds and Indians to ensure the maintenance of Whites at the top of South Africa’s racial hierarchy.⁴ The system of Apartheid implemented by South Africa’s Afrikaner National Party led to racial disparities that severely disadvantaged Black, Coloured, and Indian South Africans; furthermore, the ramifications of Apartheid have significant impacts on race relations in contemporary South Africa.

The Population Registry Act of 1950 initiated the National Party's efforts to instill Apartheid's race-based hierarchy in South Africa, identifying and separating people based on arbitrary racial identifications to embed formalized racism into South Africa’s legal systems.⁵ The process was one of South Africa's first systems to classify the country's diverse ethnic population into four “Verwoerdian” categories: White, Indian, Coloured, and Black.⁶ The National Party's racial classification system worked exceptionally for White South Africans, whose race was typically verified through census data and photo records if available.⁷ However, the South African state imposed immense barriers on Indian and Coloured South Africans. Due to the Afrikaaner-led government's mistrust of Indians and Coloureds, the National Party required Coloured and Indian South Africans to present themselves in person for identification

¹“Afrikaner” refers to the White descendants of predominantly Dutch and other European settlers in South Africa & <https://overcomingapartheid.msu.edu/unit.php?kid=163-571-5>.

²Michael Pearson and Tom Cohen, *Life under apartheid: demeaning, often brutal*, (CNN News 2013), <https://www.cnn.com/2013/12/06/world/africa/mandela-life-under-apartheid/index.html>

³Emma Norton, *How South African apartheid was ended*, (Redflag 2021), <https://redflag.org.au/article/how-south-african-apartheid-was-ended>

⁴Saul Dubow, *Apartheid, 1948-1994*, (Oxford University Press 2014),

https://books.google.com/books?id=Nf6BAwAAQBAJ&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false, pg. 11

⁵Akil Kokayi Khalfani and Tukufu Zuberi, *Racial classification and the modern census in South Africa, 1911–1996, Race & Society*, <https://www.sciencedirect.com/science/article/abs/pii/S109095240300007X>

⁶“Verwoerdian” refers to the racial categories created by Hendrick Verwoerd, Former South African Prime Minister, National Party Politician, and White Supremacist. “Coloured” refers to South Africans of mixed-race origin, typically of mixed Black and White descent. & Deborah Posel, *Race as Common Sense: Racial Classification in Twentieth-Century South Africa*, (African Studies Review, 2001), <https://www.jstor.org/stable/525576?seq=1>

⁷Keith Breckenridge, *The Book of Life: The South African Population Register and the Invention of Racial Descent, 1950–1980*, (Kronos 2014), <https://www.jstor.org/stable/24341941?seq=6>

cards.⁸ In response to the lack of voluntary racial identification applications from non-White South Africans, the National Party implemented several policies that caused negative implications for Coloureds and Indians who did not wish to cement their inferior racial identity in the Apartheid hierarchy through the identification card system. By making identification cards a requirement for social benefits, including job training and pensions, the National Party enforced their Verwoerdian racial classifications onto 95% of White, Indian, and Coloured South Africans by 1958.⁹ The National Party's imposition of identification cards was crucial in engraving racial hierarchies into South African society, and the National Party's enforcement policies based on these racial categories led to many of the atrocities committed by the Afrikaaner-led government during Apartheid.¹⁰

Specifically, the National Party's policies invoked a social system in South Africa that was heavily intertwined with race, leading to legally enforced systems of white supremacy that dominated South African culture.¹¹ In 1950, South Africa's government passed the Immorality Amendment Act, prohibiting adultery and sexual intercourse between Black and White South Africans, in an attempt to legally prohibit the promulgation of mixed-race relationships and citizens in South Africa.¹² The 1950 Immorality Amendment guaranteed the rigidness of race as a direct form of social class in South Africa, strictly guarding white supremacy within South Africa's Afrikaner population while limiting the spread of mixed-race South Africans who could assimilate into White South African populations and benefit from Afrikaner systems of power.

Relationship status was one of multiple facets of South African life restricted by skin color; the National Party initiated several laws causing race to significantly impact one's education, income, and political rights. In 1953, The National Party passed the Bantu Education Act, which led to separate public education systems based on race that severely disadvantaged Black South Africans and limited Coloured and Indian education access.¹³ The Bantu Education Act gave the National Party absolute control over the curriculum taught to Black South Africans on reserves and native tribe lands.¹⁴ South Africa's government used its control of Bantu education to brainwash Black South Africans into believing that they were not citizens of their own country. Bantu education mandated by the Apartheid State propagated ideals of black inferiority to Europeans socially and professionally and reinforced principles of Black

⁸Keith Breckenridge, *The Book of Life*

⁹Keith Breckenridge, *The Book of Life*

¹⁰A. Levine and J.J. Stremlau, *Apartheid*, (International Encyclopedia of the Social and Behavioral Sciences 2001), <https://www.sciencedirect.com/science/article/abs/pii/B0080430767010950>

¹¹(U.S. Department of State Archives 2001), <https://2001-2009.state.gov/r/pa/ho/time/pcw/98678.htm>. & Miquel Pellicer and Vimal Ranchhod, *Understanding the effects of racial classification in Apartheid South Africa*, (Journal of Development Economics 2023), <https://www.sciencedirect.com/science/article/abs/pii/S0304387822001407>

¹²*Apartheid Legislation 1850s-1970s*, (South African History Online), <https://www.sahistory.org.za/article/apartheid-legislation-1850s-1970s>

¹³Matthew Gallo, *Bantu Education, and Its Living Educational and Socioeconomic Legacy in Apartheid and Post-Apartheid South Africa*, (Fordham University Digital Research 2020), https://research.library.fordham.edu/cgi/viewcontent.cgi?article=1041&context=international_senior

¹⁴Matthew Gallo, *Bantu Education, and Its Living Educational and Socioeconomic Legacy in Apartheid and Post-Apartheid South Africa*

submission to White leaders in South Africa.¹⁵ Further, the National Party mandated the teaching of unskilled labor to Black children, training them to become unskilled, low wage workers in the Apartheid economy.¹⁶ In the words of South Africa's Prime Minister from 1958 to 1966, Hendrick Verwoerd:

There is no place for [the Bantu] in the European community above the level of certain forms of labour...What is the use of teaching the Bantu child mathematics when it cannot use it in practice? That is quite absurd.¹⁷

By providing significantly inferior education to Native Black South Africans, the National Party aimed to perpetuate segregation and racial inequalities in job attainment and economic mobility for Black South Africans. While Black students were the most significantly impacted by education inequality in Apartheid South Africa, education disparities disproportionately impacted Coloured and Indian communities as well.¹⁸ As a result of Apartheid funding inequalities, White schools had a student-teacher ratio of 1:18, compared to 1:24 in Indian/Asian schools, 1:27 in Coloured schools, and 1:39 in Black schools.¹⁹ Further, 96% of educators in White schools had teaching certificates, compared to just 15% in Black schools, highlighting differences in educational quality for Black students in comparison to White children.²⁰

Along with educational disparities perpetuated by Apartheid-Era legislation, South Africa's government enacted various legislation regulating Black income and wealth, effectively impoverishing South Africa's Black population. In 1958, South Africa's government passed the Natives Taxation and Development Act, requiring Black men to pay higher taxes at a younger age than men of any other racial group, despite the severity of the existing hardships Black South Africans faced.²¹ Due to the National Party's policies on education and job attainment, White South Africans held almost 90% of high-income and middle-class jobs during Apartheid's peak in 1970.²² To further cripple Black South Africans' place in Apartheid society, the National Party passed the Natives (Prohibition of Interdicts) Act to cease the right of Blacks to receive court protection for any oppressive laws inflicted upon them by the South African Government.²³ Similarly, South Africa's government passed the Industrial Conciliation Amendment Act of 1956

¹⁵African Studies Center, *Overcoming Apartheid Building Democracy*, <https://overcomingapartheid.msu.edu/sidebar.php?kid=163-581-2>

¹⁶Matthew Gallo, *Bantu Education, and Its Living Educational and Socioeconomic Legacy in Apartheid and Post-Apartheid South Africa*

¹⁷Matthew Gallo, *Bantu Education, and Its Living Educational and Socioeconomic Legacy in Apartheid and Post-Apartheid South Africa*

¹⁸Maria Ocampo, *A Brief History of Educational Inequality from Apartheid to the Present*, (Global Perspectives on Human Language: The South African Context, Stanford University 2004), https://web.stanford.edu/~jbaugh/saw/Lizet_Education_Inequity.html#:~:text=For%20instance%2C%20Apartheid%20funding%20resulted,affected%20the%20quality%20of%20teachers.

¹⁹Maria Ocampo, *A Brief History of Educational Inequality from Apartheid to the Present*

²⁰Maria Ocampo, *A Brief History of Educational Inequality from Apartheid to the Present*

²¹*Apartheid Legislation 1850s-1970s*, (South African History Online)

²²*Race and class inequality in South Africa*, (Roape 2022)

²³*Apartheid Legislation 1850s-1970s*

to end mixed-race trade unions.²⁴ The Act mandated that unions serving Coloured and Indian workers had to become race exclusive and led by a White executive.²⁵ Black South Africans were strictly prohibited from forming trade unions at all.²⁶ By severely restricting trading rights for non-Whites, the South African government was able to halt social mobility and remove threats from Coloureds and Indians to White power in South Africa, while forcing Blacks into exploitative working conditions. The string of acts passed in quick succession by the National Party led to sharp income and wealth disparities between South Africans of different racial groups, specifically undermining Black South Africans and elevating the lifestyle of Whites.

An analysis of census data from South Africa in the 1980's and early 1990's shows the disparate effects the National Party's policies had on income and education gaps between racial groups.²⁷ In the sample studies, census data showed that during Apartheid's later years, White people obtained an average of 5 more years of education than Coloured people, and White men earned almost 4.5 times the amount of Coloured men.²⁸ In 1960, the average White South African earned almost 12 times the amount of money as the average Black South African.²⁹ While Black South Africans were the most heavily impacted by Apartheid-era education and income legislation, Coloured and Indian communities were also adversely affected by these policies. Through their laws, the National Party developed a rigid system of racial oppression which enveloped South Africa's government, economy, and education systems, placing Blacks at the bottom of society and strategically holding Coloureds and Indians in intermediary positions of inferiority to Whites.³⁰ Most importantly, The National Party's laws ensured that White South Africans could effortlessly dominate all means of South African life.

While the White dominance of South African politics and economy was vital in maintaining systems of White supremacy, the use of land as a tool for segregation was instrumental in perpetuating racial inequality in South Africa.³¹ South African laws on land regulation imposed severe restrictions on residence and property ownership for non-White South Africans, especially limiting how Black South Africans could utilize the ownership of land as a form of building wealth. In 1936, South Africa's government passed the Natives Trust Act, restricting black land ownership and primary residence to just 13% of South Africa's available

²⁴*Apartheid Legislation 1850s-1970s*

²⁵*Apartheid Legislation 1850s-1970s*

²⁶*Apartheid Legislation 1850s-1970s*

²⁷Miquel Pellicer and Vimal Ranchhod, *Understanding the effects of racial classification in Apartheid South Africa*

²⁸Miquel Pellicer and Vimal Ranchhod, *Understanding the effects of racial classification in Apartheid South Africa*

²⁹See Appendix A, *Chart of the Week: How South Africa changed, and didn't, over Mandela's lifetime*, (Pew Research Center 2013),

<https://www.pewresearch.org/short-reads/2013/12/06/chart-of-the-week-how-south-africa-changed-and-didnt-over-mandela-lifetime/>

³⁰Whitney Pirtle, *White People Still Come Out on Top": The Persistence of White Supremacy in Shaping Coloured South Africans' Perceptions of Racial Hierarchy and Experiences of Racism in Post-Apartheid South Africa* (MDPI Social Sciences 2022), <https://www.mdpi.com/2076-0760/11/2/70>

³¹Although the National Party formally established Apartheid in the mid-1950s, racial segregation had existed in South Africa since the late 19th Century, and South Africa's government had passed many laws regarding land distribution and race before the National Party came into power (Including those discussed in this paper).

land.³² As a result of racial segregation, systemic racism, and colonization, 87% of South Africa's land became controlled by White South Africans, who comprised just 15% of South Africa's population.³³ Further, South Africa's indigenous black population was confined to environmentally depleted communal reserves that were wholly incapable of supporting South Africa's entire black population, which comprised almost 70% of South Africa's population at the time.³⁴ Specifically in rural regions, Blacks were significantly disadvantaged in terms of education, healthcare, and lifestyle compared to their White counterparts in developed, urban areas of South Africa.³⁵ Furthermore, White politicians strategically enacted residency laws to restrict Blacks into living in areas near resource mines with brutal working conditions and scarce pay.³⁶ By depriving Black South Africans of vital resources on native reserves, and by forcing Blacks into the dependence on horrific working environments for basic sustenance, South Africa's government was able to construct systems of White supremacy that exploited, abused, and discriminated against Black South Africans.

Along with the atrocities committed against South Africa's black population, Indians and Coloureds had been increasingly segregated in South African townships and suburbs. While South Africa's government had effectively reduced all means of competition and means of opposition from the Black population, White South Africans perceived major threats from Asiatic and Coloured peoples in South African cities.³⁷ Following their antagonization and oppression of Black communities, South Africa's government passed a series of legislation that sought to segregate Indian and Coloured South Africans to reduce political, geographic, and economic threats to the White population.³⁸ In 1922, the Natal Provincial Council passed the Durban Land Alienation Act, which gave the Durban City Council the power to restrict property ownership and land occupation to people by race, effectively segregating Indians and Coloureds from Whites in major South African townships.³⁹ While South Africa's government used disparate land distribution to subject black workers to exploitative working conditions in natural resource mines, much of South Africa's anti-Asiatic legislation sought to cripple Indian trade and business.⁴⁰ In 1923, the Minister of Interior introduced the Class Areas Bill, legalizing the compulsory residential and trading segregation of Coloureds and Indians throughout South Africa.⁴¹ Riding a wave of anti-Indian sentiment, South Africa's government passed the Rural Dealers Ordinance of 1924, debilitating Indian trade in South African cities by prohibiting Indian

³²Saul Dubow, *Apartheid, 1948-1994*, pg. 11

³³Saul Dubow, *Apartheid, 1948-1994*, pg. 11

³⁴*South Africa: Revolution at the Ballot Box*, (TeachDemocracy)

³⁵Saul Dubow, *Apartheid, 1948-1994*, pg. 12

³⁶Erin Blakemore, *The Harsh Reality of Life Under Apartheid in South Africa*, (History.com 2023), <https://www.history.com/news/apartheid-policies-photos-nelson-mandela>

³⁷Kalpana Hiraral, *The Myth of the 'Asiatic Menace' in Post War Natal (1918-1924)*

³⁸*Segregation and Apartheid Laws as Applied to Indians (1859-1994)*

³⁹The Natal Provincial Council was the provincial council for the current province, KwaZulu-Natal, a region on South Africa's east coast that includes the city of Durban, one of South Africa's largest cities. & Kalpana Hiraral, *The Myth of the 'Asiatic Menace' in Post War Natal (1918-1924)*

⁴⁰Kalpana Hiraral, *The Myth of the 'Asiatic Menace' in Post War Natal (1918-1924)*

⁴¹*Apartheid Legislation 1850s-1970s*

land and business ownership in White areas.⁴² Discriminatory land policies during Apartheid were critical in perpetuating White Supremacy and destroying competition from Indians and Coloureds, while capitalizing on Black labor and resource disparities.

The ramifications of inequitable land policies during the 20th Century coupled with White Supremacy in political and economic spheres continues to impact non-Whites in contemporary South Africa. In 2011, 18 years after Apartheid ended, the average annual income of white households was 365,000 rand, followed by Indians at 251,541 rand, Coloureds at 112,172 rand, and trailed by Black households that earned an average of just 60,600 rand.⁴³ More shockingly, these income disparities between Black and White families are in spite of a 170% increase in Black income since 2001.⁴⁴ Correlations between race and income in Modern South Africa are remarkably consistent with Apartheid-Era racial hierarchies, with Blacks at the bottom of the rung, Indians and Coloureds in the middle, and Whites reigning paramount on society.

The wealth gap between Black and White South Africans is even greater than their income disparities, as systemic barriers resulting from Apartheid diminished the ability of Black families to build generational wealth.⁴⁵ Despite the average Black family earning nearly 20% of the average White family's income, the average household wealth of a Black South African was just five percent of the average White family in 2015.⁴⁶ Wealth distribution inequality within South Africa is also very high, with recent research reporting that ten percent of all South Africans, the majority White, own more than 90% of South African wealth, while the considerably black bottom 80% of South Africa's population owns nearly nothing.⁴⁷ In 2018, the World Bank named South Africa the most unequal country in the world.⁴⁸

Land disparities in South Africa are similarly disproportional as a result of Apartheid-Era legislation.⁴⁹ In 2015, Whites comprised eight percent of South Africa's population, yet they owned 72% of agricultural land in the country, while the Black population consisting of 80.5% of the population owned four percent of South Africa's agricultural land.⁵⁰ These statistics are remarkably similar to land ownership rates in South Africa nearing the end of Apartheid in 1987,

⁴²Kalpana Hirral, *The Myth of the 'Asiatic Menace' in Post War Natal (1918-1924)*

⁴³See Appendix B, Statistical release (Revised), pg. 42

⁴⁴*South Africa's census: Racial divide continuing* (BBC World News)

⁴⁵Antony Sguazzin, *South Africa Wealth Gap Unchanged Since Apartheid, Says World Inequality Lab*, (Time Magazine 2021)

⁴⁶See Appendix C, *Cross-Sectional Features of Wealth Inequality in South Africa: Evidence from the National Income Dynamics Study*

⁴⁷Rebecca Fogel, *Informal housing, poverty, and legacies of apartheid in South Africa*, (Urban at the University of Washington 2019), <https://urban.uw.edu/news/informal-housing-poverty-and-legacies-of-apartheid-in-south-africa/>

⁴⁸*In South Africa, Leveling the Playing Field at Birth Critical to Reducing Inequality, Intergenerational Poverty*, (The World Bank 2022),

<https://www.worldbank.org/en/region/afr/publication/in-southern-africa-leveling-the-playing-field-at-birth-critical-to-reducing-inequality-intergenerational-poverty#:~:text=%E2%80%9CLeveling%20the%20playing%20field%20at,World%20Bank%20Country%20Director%20for>

⁴⁹Rudolph Daniels, *The Nature of the Agrarian Land Question in the Republic of South Africa*, (The American Journal of Economics and Sociology 1987), <https://www.jstor.org/stable/3486700>

⁵⁰*Mid-Year Population Estimates*, (Statistics South Africa 2015), <https://www.statssa.gov.za/publications/P0302/P03022015.pdf>

when 81.2% of land was owned by Whites, 3.6% was occupied by Blacks, and Indians and Coloureds owned 2.4% of land.⁵¹

Interestingly, Whites and Blacks alike are discontented by their racial positioning in modern South African society.⁵² While Black South Africans perceive the lasting impacts of Apartheid-Era inequalities on opportunities for social mobility, White South Africans remain resentful of current affirmative action policies enacted to correct systemic inequalities that impact people of color.⁵³ However, the role of Coloureds in modern South Africa is more complex. While research finds that historical white supremacy impacts the experiences of Coloureds in South Africa, their relative “whiteness” holds them in a position of power compared to Black South Africans.⁵⁴ Further research asserts the ambivalence of the position of Coloureds within South African social hierarchies, highlighting their privilege over Black South Africans but subordination to Whites. Scholars argue that their social ambivalence causes South Africa’s Coloured population to be in a constant power struggle to maintain their racial intermediacy in South African society.⁵⁵ As a result of Apartheid’s racial hierarchies, South Africa’s relationship with the Coloured identity differs from other nations such as the United States, where mixed-race Black and White people are generally perceived as Black by the public.⁵⁶ Notably, it seems that despite the vast differences in discrimination between South African racial groups during Apartheid, all racial groups negatively perceive their status within contemporary South Africa. As South Africa continues to transition from its Apartheid-Era systems of oppression to an all-encompassing democratic society, the forms of racial discrimination present during Apartheid continue to have lasting impacts on race in modern South African society.

=

⁵¹Rudolph Daniels, *The Nature of the Agrarian Land Question in the Republic of South Africa*

⁵²Whitney Pirtle, *White People Still Come Out on Top*

⁵³Whitney Pirtle, *White People Still Come Out on Top*

⁵⁴Whitney Pirtle, *White People Still Come Out on Top*

⁵⁵Mohamed Adhikari, *Not White Enough, Not Black Enough, Racial Identity in the South African Coloured Community*, (Ohio University Press 2005), pg. 8, <https://www.jstor.org/stable/j.ctt1j7x9gx>

⁵⁶Arnold Ho, Nour Kteily, & J.M. Chen, *How Are Black–White Biracial People Perceived in Terms of Race?* (Kellogg Insight 2017),

<https://insight.kellogg.northwestern.edu/article/how-are-black-white-biracial-people-are-perceived-in-terms-of-race>

Works Cited

Census 2011. Accessed February 27, 2024.

<https://www.statssa.gov.za/publications/P03014/P030142011.pdf>.

“Economic Research Committee : ‘What Is the Use of Teaching the Bantu Child Mathematics When He Cannot Use It in Practice. : There Is No Place for Him in the European Community above the Level of Certain Forms of Labour’. -Verwoerd- 1953.” SAHA. Accessed February 26, 2024.

https://www.saha.org.za/workers/economic_research_committee_what_is_the_use_of_teaching_the_bantu_child_mathematics_when_he_cannot_use_it_in_practice_there_is_no_place_for_him_in_the_european_community_above_the_level_of_certain_forms_of_labour_verwoerd_1953.htm.

Mbewe, Samson, and Ingrid Woolard. “Cross-Sectional Features of Wealth Inequality in South Africa: Evidence from the National Income Dynamics Study.” openSALDRU, August 1, 2016.

<https://www.opensaldru.uct.ac.za/handle/11090/843#:~:text=Further%2C%20the%20results%20show%20that,by%20a%20typical%20White%20household>.

Nour Kteily Professor of Management & Organizations. “How Are Black–White Biracial People Perceived in Terms of Race?” Kellogg Insight, April 19, 2022.

<https://insight.kellogg.northwestern.edu/article/how-are-black-white-biracial-people-are-perceived-in-terms-of-race>.

Akil Kokayi Khalfani, AbstractThe South African government has previously used racial classification as a tool in its official state policies to control its population. In this essay, and A. McDaniel. “Racial Classification and the Modern Census in South Africa, 1911–1996.” Race and Society, June 14, 2003.

<https://www.sciencedirect.com/science/article/abs/pii/S109095240300007X>.

A. Levine a, a, b, and An Afrikaans word meaning ‘separateness. “Apartheid.” International Encyclopedia of the Social & Behavioral Sciences, November 2, 2002.

<https://www.sciencedirect.com/science/article/abs/pii/B0080430767010950>.

“Apartheid Legislation 1850s-1970s.” South African History Online. Accessed February 26, 2024. <https://www.sahistory.org.za/article/apartheid-legislation-1850s-1970s>.

“Apartheid, 1948-1994.” Google Books. Accessed February 26, 2024.

<https://books.google.com/books?hl=en&lr=&id=Nf6BAwAAQBAJ&oi=fnd&pg=PP1&dq=History%2Bof%2BApartheid&ots=SAAz5HSJ86&sig=zWEznmzm9e0-mXvbbYGuGZQ1tVI#v=onepage&q=History%20of%20Apartheid&f=false>.

Bantu education, and its living educational and ... Accessed February 27, 2024.

https://research.library.fordham.edu/cgi/viewcontent.cgi?article=1041&context=international_senior.

Bloomberg, Antony Sguazzin /. “South Africa Wealth Gap Unchanged since Apartheid - Report.” Time, August 5, 2021.

<https://time.com/6087699/south-africa-wealth-gap-unchanged-since-apartheid/>.

- The book of life: The south african population register and the ... Accessed February 27, 2024.
http://www.scielo.org.za/scielo.php?script=sci_arttext&pid=S0259-01902014000100010.
- Daniels, Rudolph. "The Nature of the Agrarian Land Question in the Republic of South Africa." *The American Journal of Economics and Sociology* 46, no. 1 (January 1987): 1–16.
<https://doi.org/10.1111/j.1536-7150.1987.tb01752.x>.
- DeSilver, Drew. "Chart of the Week: How South Africa Changed, and Didn't, over Mandela's Lifetime." Pew Research Center, December 6, 2013.
<https://www.pewresearch.org/short-reads/2013/12/06/chart-of-the-week-how-south-africa-changed-and-didnt-over-mandelas-lifetime/>.
- "The Harsh Reality of Life under Apartheid in South Africa." History.com. Accessed February 26, 2024. <https://www.history.com/news/apartheid-policies-photos-nelson-mandela>.
- Hiralal Kalpana. "The Myth of the 'Asiatic Menace' in Post War Natal (1918-1924)" *Alternation*, 2002. Accessed February 27, 2024.
https://journals.co.za/doi/pdf/10.10520/AJA10231757_649
- "How South African Apartheid Was Ended." Red Flag. Accessed February 26, 2024.
<https://redflag.org.au/article/how-south-african-apartheid-was-ended>.
- "Informal Housing, Poverty, and Legacies of Apartheid in South Africa." UrbanUW. Accessed February 26, 2024.
<https://urban.uw.edu/news/informal-housing-poverty-and-legacies-of-apartheid-in-south-africa/>.
- Miquel Pellicer a b 1, a, b, 1, c, Highlights•We estimate the causal effect of racial classification on labour market outcomes in Apartheid South Africa. •We use a change in the criteria by which people were racially classified for identification. •Our estimates for education are extremely l, AbstractThe apartheid era began in 1948 in South Africa, et al.
 "Understanding the Effects of Racial Classification in Apartheid South Africa." *Journal of Development Economics*, October 29, 2022.
<https://www.sciencedirect.com/science/article/abs/pii/S0304387822001407>.
- Not white enough, not black enough: Racial identity in the South ... Accessed February 27, 2024.
<https://www.jstor.org/stable/j.ctt1j7x9gx>.
- Pearson, Michael, and Tom Cohen. "Life under Apartheid: Demeaning, Often Brutal." CNN, December 6, 2013.
<https://www.cnn.com/2013/12/06/world/africa/mandela-life-under-apartheid/index.html>.
- Pirtle, Whitney N. Laster. "'white People Still Come out on Top': The Persistence of White Supremacy in Shaping Coloured South Africans' Perceptions of Racial Hierarchy and Experiences of Racism in Post-Apartheid South Africa." MDPI, February 14, 2022.
<https://www.mdpi.com/2076-0760/11/2/70>.
- Race as common sense: Racial classification in twentieth-century ... Accessed February 27, 2024.
<https://www.cambridge.org/core/services/aop-cambridge-core/content/view/EB235D7E4>

5B7D501FC4B2D53E2DD974F/S0002020600029826a.pdf/race_as_common_sense_racial_classification_in_twentiethcentury_south_africa.pdf.

“Resources.” South Africa: Overcoming Apartheid. Accessed February 26, 2024.

<https://overcomingapartheid.msu.edu/unit.php?kid=163-571-5>.

“South Africa’s Census: Racial Divide Continuing.” BBC News, October 30, 2012.

<https://www.bbc.com/news/world-africa-20138322>.

World Bank Group. “In Southern Africa, Leveling the Playing Field at Birth Critical to Reducing Inequality, Intergenerational Poverty.” World Bank, March 14, 2022.

<https://www.worldbank.org/en/region/afr/publication/in-southern-africa-leveling-the-playing-field-at-birth-critical-to-reducing-inequality-intergenerational-poverty>.

Programming Prejudice: Violence Against Women in Media and its Effect on Teenage Attitudes Towards Misogyny By Anita Marie Júlca

Abstract

Within this study, I aimed to highlight the impacts of growing up in a digital landscape oversaturated with depictions of violence against women. Whether it is visual, lyrical, or insinuated, teenagers are face to face with extreme-misogyny across the internet. This study was conducted in 2 phases, the first of which was based upon self-reported data from a group of 1000 American teenagers, 500 males and 500 females. Teens responded to surveys aimed at targeting the effects 3 categories of media: social media, mainstream music, and digital pornography, all of which grapple with violent and disturbing depictions of the female. Within phase 2, teens were divided into 2 groups, high consumption and low consumption. These groups were based on high media consumption in contrast with low media consumption and how the media one consumed impacted their likelihood of promoting statements of misogyny. Both groups were posed with 10 statements, each based upon strong ideologies of misogyny and sexism. Teens were then asked to rate their agreement or disagreement with each statement in order to investigate how teens who consumed high amounts of misogynistic media reacted to misogyny in the ‘real-world.’ This study was targeted at depicting how messages in media are internalized, especially within developing brains, and the impacts of these harmful messages.

Introduction

Internet media is becoming more and more integrated into everyday life, especially among teens, who are becoming increasingly reliant on apps and websites not only for social purposes but also for sharing information. Unfortunately, the anonymity and ubiquity of the internet have further made it easier than ever for misogynistic content and attitudes to spread, which can come with severe developmental consequences for perceptions and behaviors in constructs associated with adolescents. (Cookingham & Ryan)(Khalaf et al.) Several studies have highlighted how social media is over-saturated with misogynistic rhetoric and abuse targeted at women and young girls. (Habib et al.)(TECHNOLOGY-FACILITATED GEND.) This study was an attempt to direct attention to one of the most overlooked facets of such exposure to online misogyny and harmful beliefs and behaviors that consequently followed.

With the evolution of media, including social media, music, and pornography, there is an eventual effect on the attitudes and behaviors of young people. Prior research has established links of exposure to media violence with heightened aggression, and connected sexualized media exposure with the endorsement of harmful beliefs about gender and sexual violence.

(Cookingham & Ryan)

The gendered nature of these outcomes and the potentials for differential effects on male and female youth is an area that is not elaborated upon to a great extent. The purpose of the current research is to examine the relationship between the exposure and consumption of

misogynistic and violent media content material on the endorsement of problematic attitudes towards women and violence among adolescents.

The incidence of violence against women in society today is an alarming fact. Although the roots of gender-based violence are quite complicated and nuanced, growing concern is being expressed over the aspect of the media that normalizes, excuses, and provokes violence against women. Recent studies have shown that exposure to misogynist material on social media greatly affects teenagers' attitudes and beliefs towards gender, power, and acceptability of violence against women. (Kahlor & Eastin) (Cosma & Gurevich)

Social media, in particular, has become a breeding ground for sharing sexist and even violent content. A study that finds Facebook may be drawing young women to objectifying images concludes: "Exposure to objectifying images directly links with body shame and self-objectification, and tolerance of sexual violence" (Manago et al.). Equally, it has been proven through studies that even small exposure to media violence can stimulate desensitization to real-life violence, increased aggressive thoughts and behaviors, and decreased empathy.

The effects of popular music, featuring its misogynistic lyrics and concepts, should not be dismissed. Such music often presents women as submissive, objectified, and deserving of abuse, which can affect the views and perceptions of society. Therefore, the effects that such music may have on children and youth remains an issue of concern. Many depictions of women in mainstream music define the female as being submissive, objectified, and deserving of abuse, which influences societal attitudes and beliefs. The long-term psychological consequences of normalizing violence against women have not yet been studied, however, there is probable reason to believe negative effects such as internalizing harmful gender norms and decreasing sexual self-efficacy among young women are just the beginning.

The proliferation of pornography, often depicting women as passive victims of male aggression, may even contribute to the normalization of women's violence. Regular exposure to this type of content has been found to tell tales of increased acceptance of myths about rape and, in fact, increased sexual aggression, with a higher likelihood of committing sexual assault; proving the psychological ramifications on account of these media portrayals are deep and pervasive. (Menzie) (Bridges)

This commonplace of violent depictions in media is indicative of and can lead to the development of widespread psychological effects among both males and females. There's a disturbing trend among the "alpha-male" archetype in the depths of social media; glorification, tolerance, and, at times, encouragement of unbearable violence against women. Violence against women is used as a twisted symbol of masculinity. In this dreaded e-nirvana, one can find a proper glorification, normalization, and even instigation of violence against women. Specific corners exist on social media designated for an "alpha male"; a breeding ground for hypermasculinity evolved into another digital ecosystem wherein the classic definitions of masculinity are warped into a toxic cocktail of aggression, control, and misogyny. (Kupers)(Russell) This online space pervaded by memes of either subservient or treacherous women lionizes figures like Andrew and Tristan Tate - self-proclaimed lifestyle gurus with a

carefully cultivated image of maximal wealth and masculinity. But beneath the bravado of sports cars and VIP lifestyles is a much darker truth. (Solea & Sugiura)(Shattering Myths and Lives: A Report on Image-Based Sexual Abuse)(Menzie) What started as a small project among the Tate brothers, themselves now notorious for making hyper-misogynistic pronouncements before the cameras, has evolved into serious accusations of human trafficking and even rape in Romania on their parts - charges that they strenuously deny. Their arrests and the courtroom dramas that have followed have only bolstered their standing among many of their young, male acolytes, who portray them as martyrs to a philosophy that vilifies women and maintains the need for male supremacy. The women that have come out against the Tate brothers carry accusations including, but not limited to: rape, violence and sexual exploitation. (Text correspondence with women as described in court document) Fanning the flames of the movement are people like Nico Kenn De Balinthazy, better known by his internet followers as 'Sneako.' Though far less legally entangled, he and others like him echo the Tates' positions in incendiary podcasts, slickly produced videos, and a steady clip of social media statements. Their rhetoric, reframed as self-improvement advice, makes a subtle shift to reframe violence against women as not only permissible but as a necessity to prove dominance. Control is then normalized as "protectiveness," a siege mentality is fostered in which men are pitted against women in an epic struggle for power, and this worldview that a woman has her compulsory value manifest only through obedience and sexual availability to men is continuously reinstated. (THE SEVEN P's OF MEN'S VIOLENCE)(Habib et al.) One study found that men who were exposed to media that objectified women were more likely to endorse rape myths and accept sexual aggression as normative. (Seabrook et al.) This insidious messaging seeps into the minds of impressionable young men, blurring the lines between assertiveness and aggression, leading to a culture where violence against women is tacitly condoned. (Cosma & Gurevich)

The psychological toll of this hyper-present misogyny extends beyond the digital spaces. The exposure of the media to images of feminicides has been linked to greater levels of acceptance and desensitization of such acts and the perpetuation of harmful stereotypes about gender roles. This dangerous rhetoric is one that young boys, seeking validation after having sailed the troubled waters of adolescence, very easily fall into. The algorithm, being geared towards maximum engagement, quickly funnels them through an upward spiral of ever more extreme content, making misogyny and violence not only normalized but framed as both aspirational and acceptable. Such constant exposure can distort one's expectations about what a healthy relationship and masculinity entail, thus perfectly laying down the groundwork for harmful behaviors toward women later in life. The deeply disturbing truth regarding mainstream pornography is the graphic showing of women as a vessel for extreme sexual abuse, normalizing violence and degradation against women as if it is a form of entertainment. (Bridges)(Shaw) Over and over again, women are portrayed as the most extreme and degrading sexual objects, stripped of all agency and dignity, existing for the sole perverted productive supply of men. (Seabrook et al.)(Boyle) The titles such as the following : "Desperate slut gets destroyed by daddy's huge cock," "Pathetic whore stepsister begs to be used," or "No mercy for screaming,

tiny teen," make this point perfectly clear, introducing highly problematic power aggravities and themes of incest, pedophilia and forceful domination, through seemingly innocent captions. In addition to this constant depictions of women being choked, slapped, hit, and penetrated while bleeding and in pain, and the naturalization of lack of consent – that is, sexual coercion - desensitizes viewers from feeling the pithos of violence against women.(Kahlor & Eastin)(Seabrook et al.) Shockingly, there is a well-documented correlation of pornography viewing, particularly for men at a young age, and an increased propensity to commit sexual violence in real life. (Shaw) Coupled with the documented rise in pornography induced sexual dysfunctions, the most common being erectile dysfunction, as well as an inability to participate in a healthy, non forcefully rendered sexual or intimate relationship, quite clearly, pornography has done more harm than good to the world and its users. (An empirical investigation of the role of pornography in the verbal and physical abuse of women - PubMed)

The present research presents strong evidence for the association of exposure to misogynistic and violent media content with the advocacy of deviant views regarding women and violence in teenagers, more specifically, male teenagers. Our findings suggest that a reduction in exposure to this sort of unhealthy media content alone could do much toward engendering more equal and non-violent attitudes among youth. It points to the importance of critically analyzing possible media effects, consumption of which influences the genesis of gender stereotypes and beliefs about sexual violence.

Methods

This review consists of a survey-based research study to examine the impact of misogyny in media on teenagers' attitudes regarding violence toward women. As such, while phase 1 of the study primarily consisted of self-reported media consumption and attitudes towards violence, phase 2 of the survey measured the extent and impact of such exposure among adolescents of the said misogynistic content.

Our sample included 1000 teens, ages 15-18, with an average age of 16.2, including an equal number of males and females, recruited online through forums and social media. (N = 1000; 50% females; 50% males). All participants were United States residents, although the response was paramount from the states of California, New Hampshire, Vermont, Massachusetts, New York, North Carolina, and Florida. Participants were asked to fill out a questionnaire for the first half of the assessment, addressing variables such as social media consumption, music habits, and pornography consumption. They were also asked about exposure to and reactions toward misogynistic content, attitudes and reactions toward women, and self-reported tendencies toward violence or aggression.

This primary phase was aimed at establishing a general database of the consumption habits of media and their potential correlation with individual perceptions and experiences.

In the comparative study within the second phase, we focused on specific media use habits we deemed likely to enhance endorsement of misogynist statements. In this, two groups were distinguished:

Moderate to Heavy Consumption Group: This consisted of a group of those participants who reported using the internet daily for 8+ hours, frequently listened to music containing misogynistic and violence-advocating lyrics, and consumed a significant amount of digital pornography on a weekly basis.

Low Consumption Control Group: This provided a comparison baseline of participants who reported little or no pornography consumption, with less than 2 hours per week of usage overall across social media, and minimal overall exposure to music with misogynistic content.

The responses of these two groups were then taken for statistical analysis to ascertain significant differences in their attitudes towards women and acceptance of violence.

In this regard, we sought to contrast their reactions by pitting both parties against an array of misogynistic statements and aiming to establish a correlation in terms of levels of agreement with or disapproval of such statements, and then to link the consumption of media at a higher rate with a greater acceptance of dangerous ideologies. This two-pronged approach helped us understand better the interplay of media consumption with personal beliefs and probable behavioral outcomes.

The participation in this study was purely voluntary, superseded by the free will to withdraw when desired. Efforts were made to make the participants express themselves as honestly as possible. They were also given information regarding pornography use and addiction, the addiction to social media, and support services that exist for victims/survivors of abuse. All researchers completed the "Protecting Human Research Participants Training Module" developed by the NIH Office of Extramural Research to ensure methodological rigor and ethics in this study. It aims to bring out the relationship between media consumption and the formation of attitudes likely to escalate gender-based violence among youth.

We used a combination of statistical techniques that included regression analysis, ANOVA, and thematic coding for qualitative responses for data analyses.

Exclusion Criteria

The exclusion criteria in the study applied to participants aged less than 15 years and more than 18 years, those who did not respond to the question about their age in the survey, those respondents who were not permanent residents in the United States, and those who did not complete the consent question. Therefore, the total number of participants used in this study was 1000, hence representing a response rate of 100 percent.

Survey Structure

The survey was systematically structured to gather detailed information from participants about demographic data, media use, and attitudes towards violence against women and misogyny. The hypothesis that was tested in this survey is whether teens' attitudes toward violence against women were affected by media portrayals of misogyny.

Demographic Information: This ranges from details such as age and sex to state of residence and school status.

Media Consumption Habits: This section asked about media consumption behaviors: the kinds of media they find themselves accessing most often, such as television, movies, social media, and video games; the kind of content viewed most often; and how much time was spent daily on each kind of media. Further questions were on their exposure to contents detailing violence against women or bearing a tone that is misogynistic in nature.

Attitudes Towards Violence Against Women and Misogyny: The attitudes of the respondents towards violence against women and misogyny were measured with a set of Likert-scale questions. The scale items tapped into statements on gender roles, acceptability of violence towards women, and beliefs about misogynistic activities. Participants rated their agreement level on a scale from "Strongly Disagree" to "Strongly Agree."

Phase 1 had 3 specific sections: the impact of social media, the impact of popular music, and the impact of mainstream pornography. Each of these consisted of multiple-choice questions. They ranged over a variety of more specific elements of the way that media may manipulate attitudes toward misogyny and violence against women. An "Other" option with each question offered space for the participant to fill in a response not offered by the choices.

Ethical Considerations: The survey was ethically conducted using anonymous responses via Google Forms. Full informed consent was appropriately sought from all participants and guardians where participants were less than 18 years, clearly detailing purpose, procedures, and the voluntary nature of participation in this study. They were informed that they were free to withdraw at any point without penalty. They were encouraged to respond as they had experienced, and all relevant services, including those addressing pornography and social media addictions, and victim and survivor supports of abuse, were made available.

Statistical Analysis

Analyses were conducted on complete cases only, $n = 1000$. Chi-squared tests or Fisher's exact tests were computed as appropriate in R version 3.6.2 to investigate differences in proportions of a range of factors pertaining to media consumption habits and attitudes towards violence against women and misogyny, by state, and by sex. In addition, independent samples t-tests and two-way independent measures ANOVAs were carried out using SigmaPlot version

14.0 to determine whether there were differences in attitudes towards violence arising from media consumption between sexes. Where appropriate, Tukey post hoc was applied. An exploratory factor analysis was conducted to explore the relationships between several of the variables to determine specific possible factors relating to attitudes towards misogyny and violence towards women rather than test a hypothesis per se. The factorability of the attitude items was first assessed using Bartlett's test of sphericity and the Kaiser–Meyer–Olkin measure of sampling adequacy via Jamovi 1.1.7. On the basis of the outcome, EFA was conducted in order to establish response patterns within the attitude items. A backward stepwise elimination regression determined how demographics, combined with media consumption habits, correlated with attitudes towards violence against women. In this study, two models were estimated: one in which the predictor variables were indicators of media use, and another in which they were attitudes toward violence against women. In both models, a range of demographic variables was added as predictors, while attitudes toward misogyny and violence against women remained as the dependent variable.

Phase 1 Results

Social Media

Below are the findings from the social media component of our broader research study investigating the effect of portrayals of, and exposure to, misogynistic content across diverse media on the sexist attitudes and responses of teenagers in the 'real' world, towards violence against women. This data provides information on how often individuals encounter misogynistic-type content on social media, the type of reactions that might follow, and the intentions on real attitudes toward women. In doing this, the research will be seeking to emphasize the general and wide exposure attributed to social media misogyny, in addition to emphasizing the potential influence attributed to shaping societal norms and behaviors tied to gender and violence.

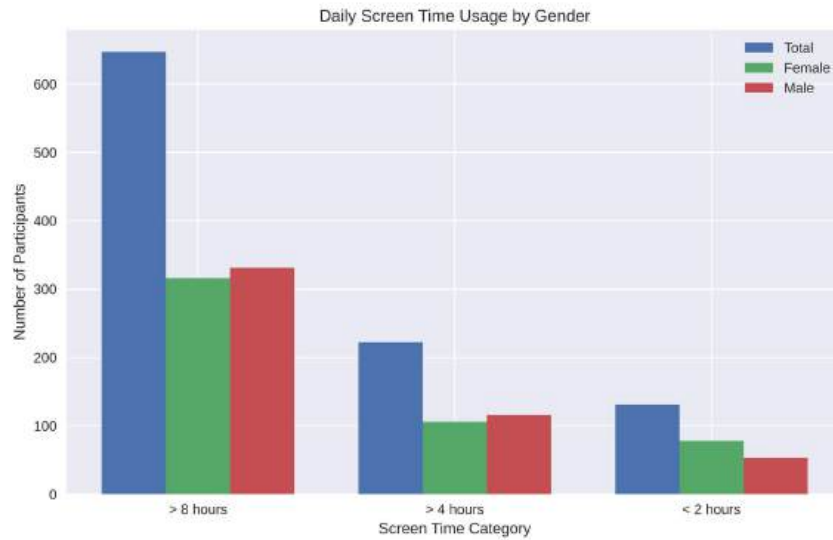


Figure 1: Daily Screen Time Usage by Gender

Figure 1 shows that out of the participants involved in the research, 64.7% said that they devote more than 8 hours a day to using their screens. This group was separated among females and males almost evenly: 316 females and 331 males. On the contrary, the smaller segment, which was 22.2% of the total sample, reported over 4 hours of screen time daily, with almost equal numbers for both males and females: 106 were female and 116 were male. In fact, only 13.1% of the participants reported less than 2 hours of screen time per day, and out of the very few, there were disproportionately more females, 78 vs. 53.

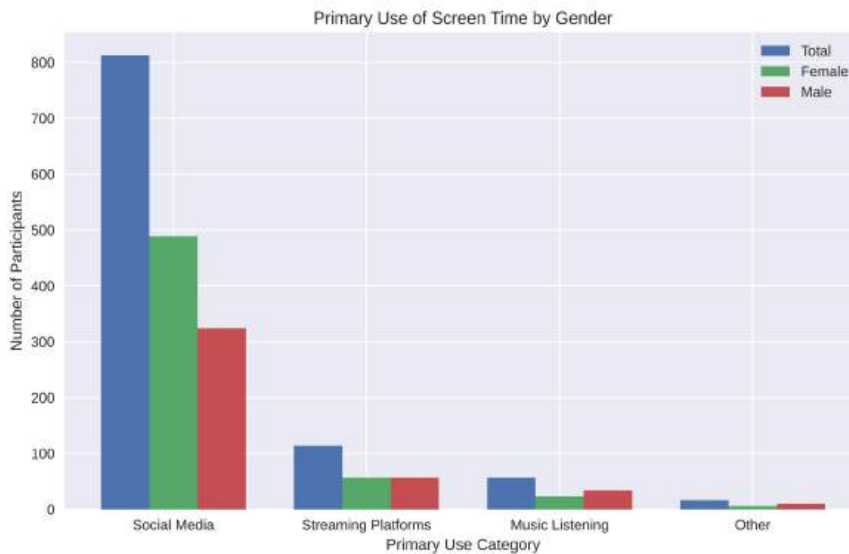


Figure 2: Primary Use of Screen Time by Gender

Figure 2 shows that, overall, in all respondents, the most common use of screen time identified was social media applications, and 81.3% of the respondents devoted their phone usage primarily to these websites and applications, such as TikTok, Instagram, Snapchat, X, and Facebook. The demographic split requires a higher tendency for the female teens of 489 compared to male teens with 324 individuals. Media streaming sites like Netflix, Hulu, and Amazon Prime followed with 11.4% users from the participants. Music listening comes under the third most used, having a proportion of 5.7%. Only a small fraction, 1.6%, belonged to the 'other' category, which included such a variety of activities as online shopping, mobile games, gambling, and educational support. The main findings of these may be the focus on digital media playing in teens' daily lives and a possible increasing shift in their way of being entertained and acquiring information, moving towards more visual and interactive means.

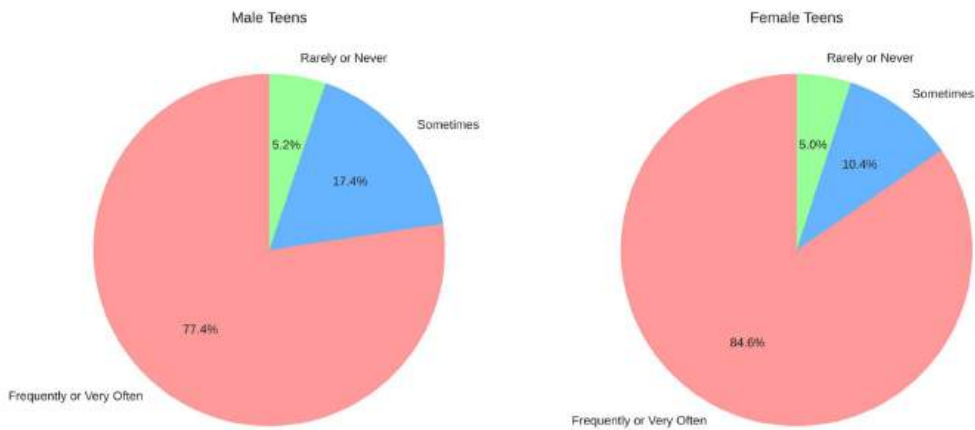


Figure 3: Frequency of Exposure to Misogynistic Content by Gender

In a question asking about the frequency of exposure to misogynistic content on social media among teens, a large majority of both genders reported high exposures. According to Figure 3, 77.4% of male teens and 84.6% of female teens indicated that they often came across this type of content. This is high; more so among the females. Such findings raise a red flag on the effect of social media on youths, more so in propagating gender biases and possible psychological effects on young users.

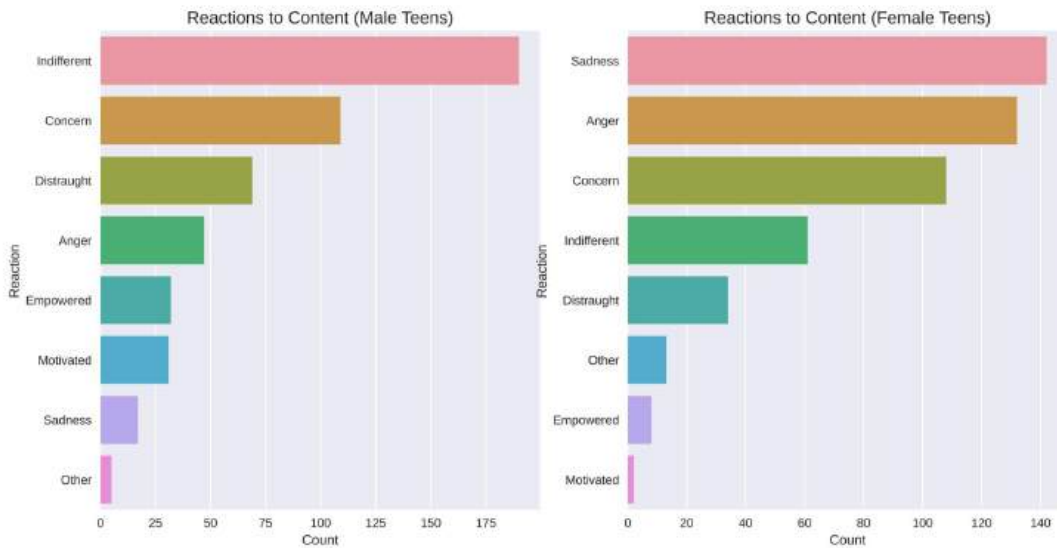


Figure 4: Reactions to Misogynistic Content by Gender

Figure 4—reactions to misogynistic content by gender on social media showed that there were varied emotional responses for male and female teens. The most common reaction for the male teens was indifference at 38.0%. Concern was next at 21.8%, followed by feelings of distraught at 13.8%. Anger and empowerment were equally present but less common, noted at 9.4% and 6.4% respectively.

The female adolescents comparatively showed relatively greater negative responses. The major response noted was sadness, which covered 28.4% of the study participants, followed closely by anger at 26.4%, concern at 21.6%, while indifference in females was relatively very low at 12.2, the feeling of being empowered was minimal at only 1.6%.

These results underline the fact that harmful content is hugely emotion-provocative, resulting in very high rates of feelings of negativity among female adolescents compared with male adolescents. This explains the need for targeted interventions against this backdrop of increased harmful content in support of affected people.

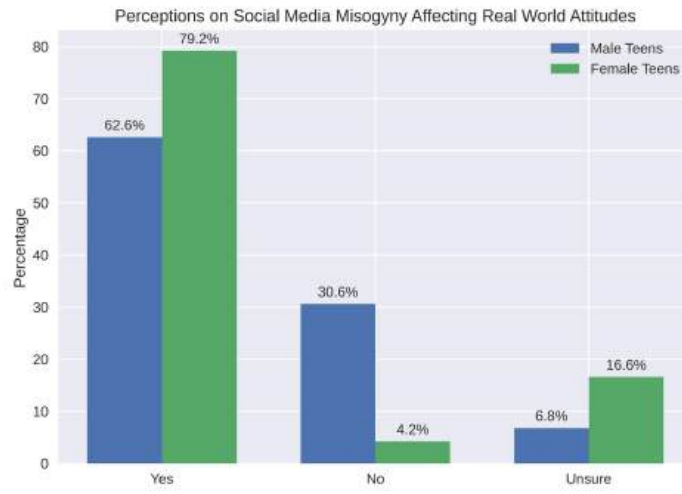


Figure 5: Perceptions on Social Media Misogyny Affecting Real World Attitudes

When asked how they thought social media misogyny might be impacting attitudes to real women, there is a significant gender divide in responses: Based on Figure 5, 62.6% of male teens thought that it does have an impact; 30.6% did not, and 6.8% were unsure. In contrast, a higher percentage of female teens responded with 79.2% acknowledging the impact, with only 4.2% disagreeing and 16.6% remaining unsure.

These results show a gender disparity in attitudes towards the effect of the misogyny on the social websites; essentially, cyber hatred towards women is affirmed by female-teen offenders only. This furthers the necessity to defuse all such misogynistic contents posted in social media sites to exercise a more cognitive and sensitive social place.

Music

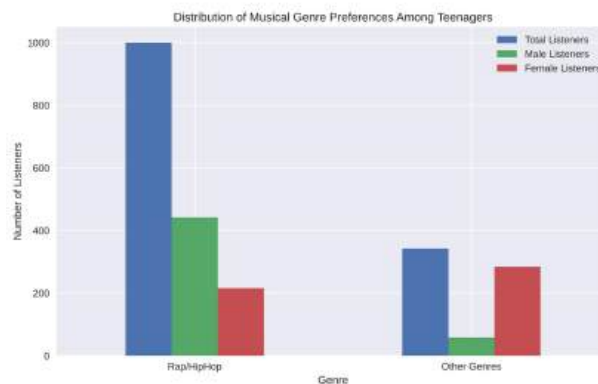


Figure 6: Distribution of Musical Genre Preferences Among Teenagers

In Figure 6, the branch of phase 1 research into the instances of misogyny within music and their impacts upon the teenage listener, it was found that the overwhelming majority, 65.8% (658 out

of 1000), listened primarily to Rap/HipHop. As such, this is the most lyrical of the options offered in the gamut of choices which included Pop, Country, Folk, Classical, Jazz, Indie, R&B, Rock, Metal, and Gospel. The data also indicate some form of gender imbalance: 442 of the listeners to Rap/HipHop are male, with 216 being female. In contrast, for the other genres preferred by 34.2% of the total teenager respondents, 58 are male while the remaining 284 are female. This distribution underlines the strong preference for Rap/HipHop among teenagers, raising significant practices in musical tastes from some different genders.

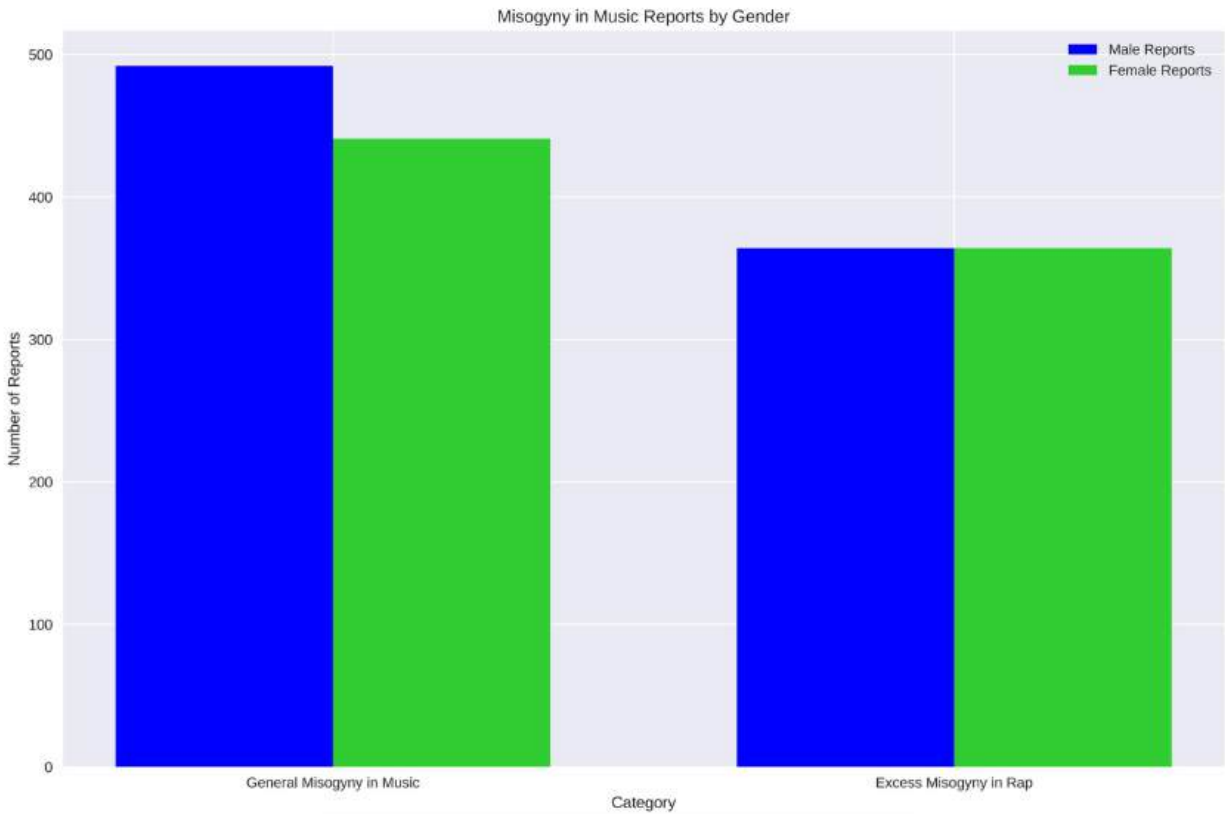


Figure 7: Reports of Misogyny in General Music and Rap Music

When exhibiting the prevalence of sexism in music, and rap specifically in Figure 7, it is indicated that a strong majority of teenagers have been exposed to misogynistic material in music, especially rap. Precisely, as tabulated, while 93.3 % of teens agreed or strongly agreed that there was general misogyny in music, males agreed a bit more than females, 98.4% compared to 88.2%. Secondly, 72.8% of the respondents agreed or strongly agreed that rap music contained excessive misogynist lyrical content, with equal, balanced reporting between both genders: 364 agreements each.

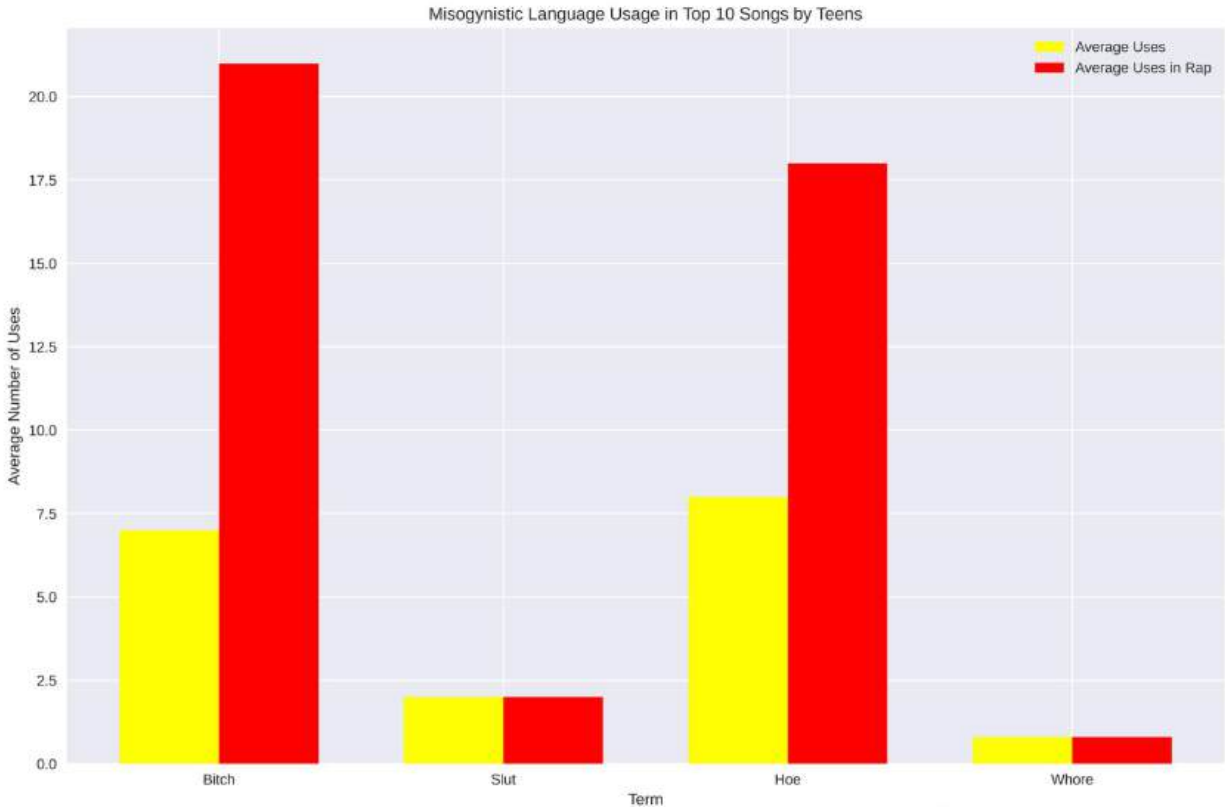


Figure 8: Misogynistic Language Usage in Teen’s Top 10 Songs of the Past Year

The above graph of Figure 8 visualizes the average uses of the following terms: "bitch," "slut," "hoe," and "whore," comparing the general usage of the terms in other genres of popular music to usage within rap music, highlighting stark differences in the context of rap.

Our study found that within the top ten most streamed songs of the last year by each teen listener, there was a high volume of derogatory terms. This involved the examination of the presence of some misogynistic terms within the lyrics by the researchers, using the software 'Receiptify', a service that amalgamates users' top ten most listened to songs annually. More specifically, the terms being sought after in this research were "bitch," "slut," "hoe," and "whore," and their usage in a degrading context towards women.

The results of Figure 8 show that "bitch" was used on average 7 times throughout all the top ten lists; for those who listened to only rap music, this rose to 21 times. "Hoe," on the other hand, was used 8 times on average, with the count increasing in the top ten songs among the rap music listeners to 18 times. In contrast, the word "slut" had average use a little less at 2 times, and again, with no significant increase among the rap listeners. The term that was used the least was "whore" at an average of 0.8 times across all lists.

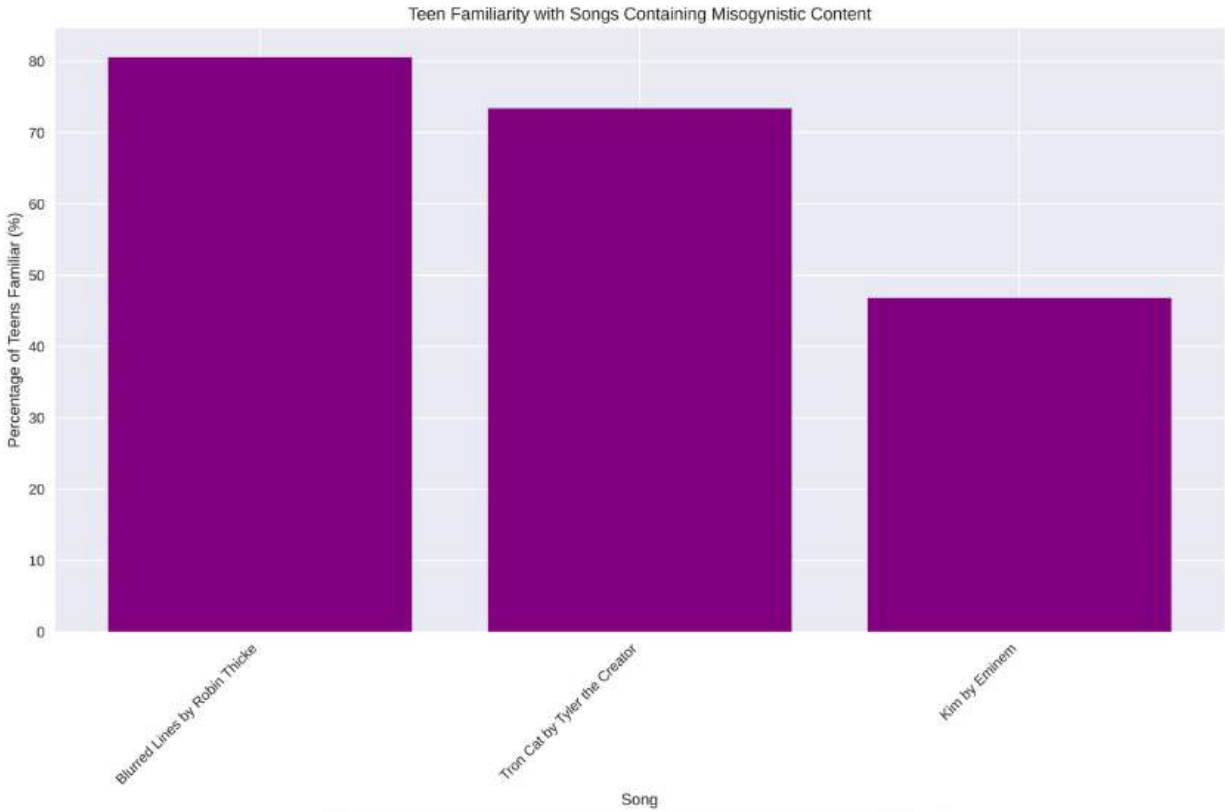


Figure 9: Teen Familiarity with Songs Containing High Misogynistic Content

Indeed, from the data of Figure 9, the familiarity that emerged for the teenage study population in songs possessing misogynistic content is immense. The data reveal that a high percentage of teenagers are familiar with several songs: "Blurred Lines" by Robin Thicke at 80.6 percent, "Tron Cat" by Tyler the Creator at 73.4 percent, and "Kim" by Eminem at 46.8 percent. The figures emanating from the data show high proportions being familiar with songs degrading and objectifying women and even bringing up a scenario of violence against them.

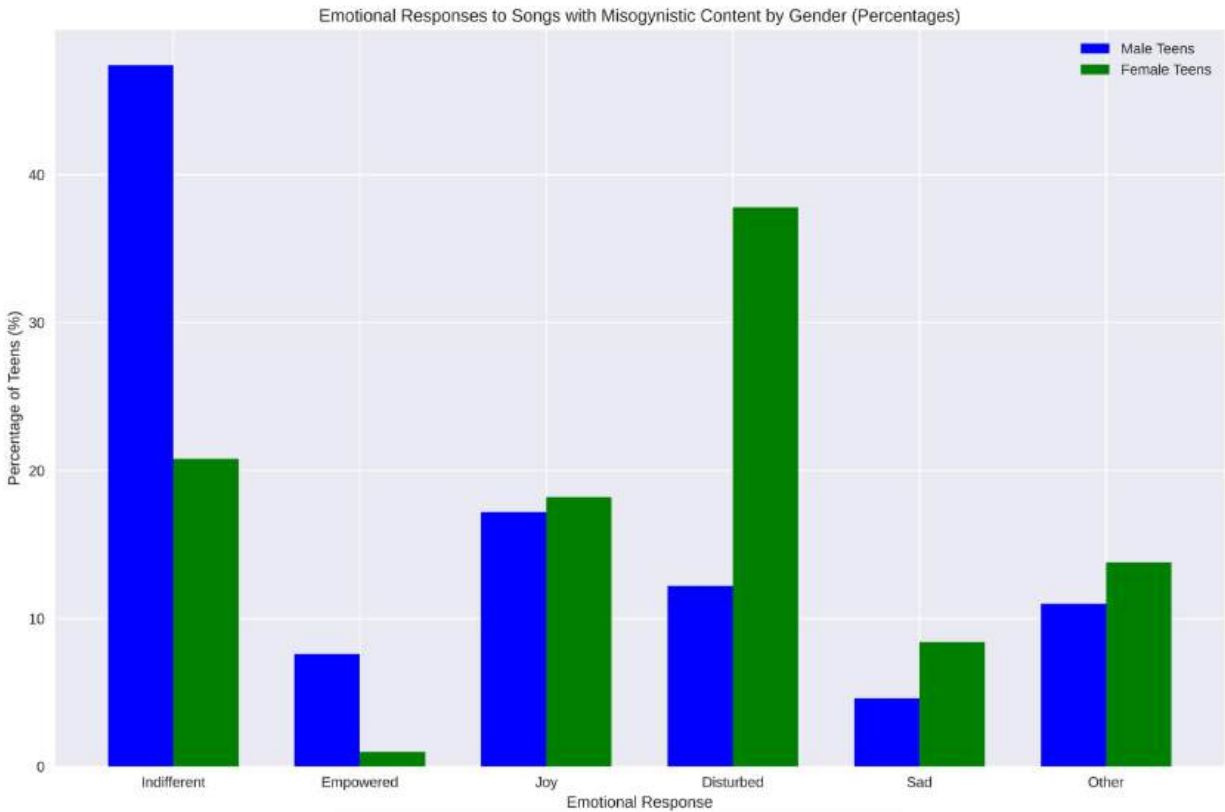


Figure 10: Percentage of Emotional Responses to Songs with Misogynistic Content by Gender

Above, in Figure 10, is a graph of emotional responses to the misogynistic content of the songs for teenagers. In this case, it can be noticed for certain that there were very distinct discrepancies in terms of gender responses: for example, 47.4% of male teens and 20.8% of female teens reported that while listening to the following songs, they felt indifferent. This points to a greater extent of emotional detachment among males. Among females, the emotional impact is much stronger, with 37.8% stating that they feel disturbed—a statistically significant percentage higher than that of males at 12.2%.

What is equally worrying is that nearly 8% of male teenagers experience feelings of empowerment through the songs. This indicates how generalized misogynistic discourse may shape certain dangerous gender role expectations or attitudes in young males. Additionally, almost equally from the genders, teens reported joy: 17.2% for boys and 18.2% for girls, which points out the controversy in meanings connected with music enjoyment loosely hiding the dimensions of lyrics.

These findings emphasize the need for educationally directed programs and interventions that significantly target the reduction of the impact that misogynistic content expressed in music lyrics may have.

Pornography

In venues of pornography viewing, a marked model of difference was identified between male and female teenagers. Similar conditions have been witnessed in nearly all the studies conducted on this topic. Figure 11 shows that 25.2% of male teens, or 126, viewed pornography daily, while 62%, or 310, viewed weekly. Among females, 12.4%—62—viewed pornography daily, while 50.4%, or 252, viewed weekly. Meaning that already from these results, a gender difference arises about frequency in the way teenagers look at pornography.

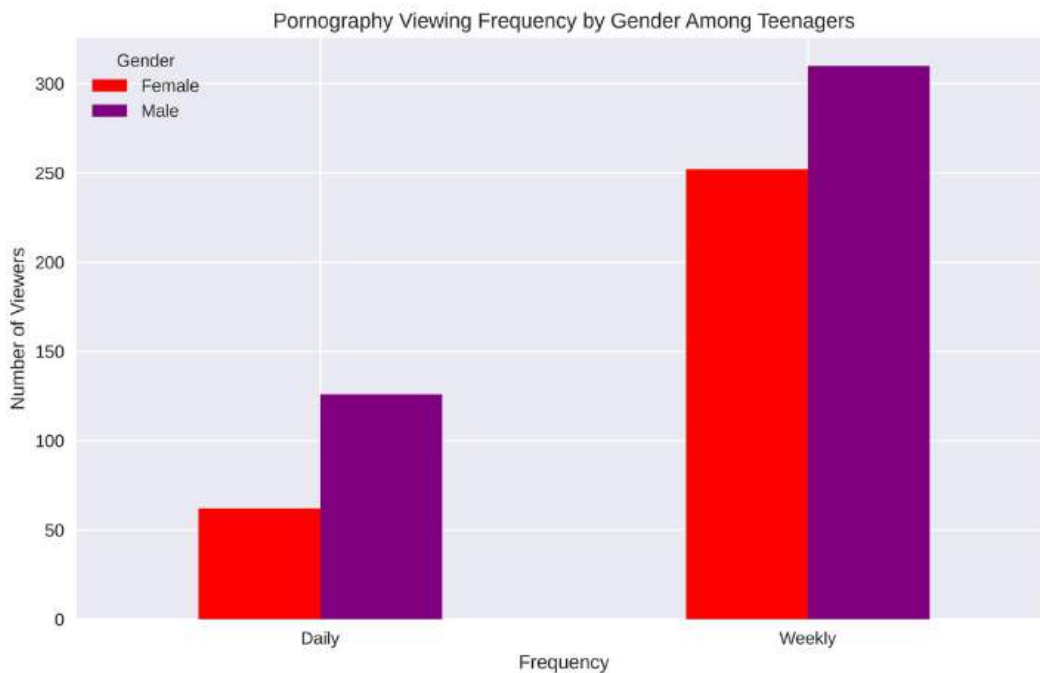


Figure 11: Pornography Viewing Frequency by Gender Among Teenagers

Observed gender differences in pornography consumption in teenagers may reflect broader societal and cultural patterns. For example, male exposures to and encouragement of sexual content could explain the difference in exposure rates. These directions may hold implications for psychological development and social relationships later in life. Further research is required to probe the underlying causes of these disparities and their long-term effects at both individual and societal levels.

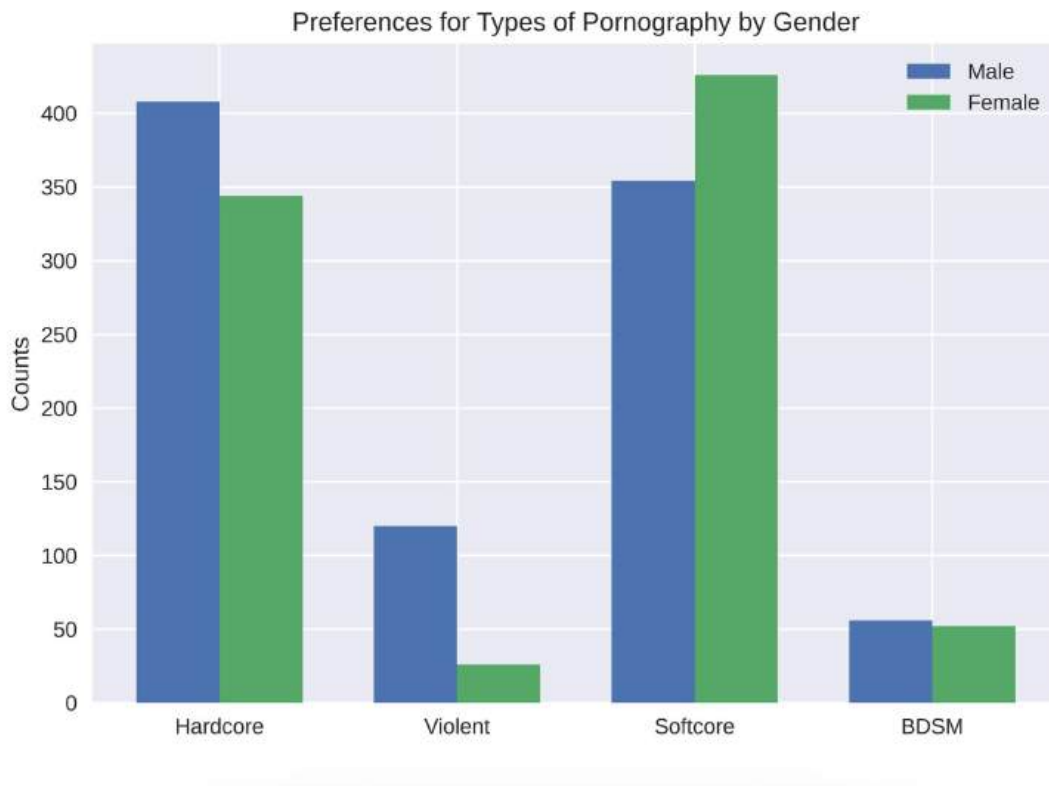


Figure 12: Preferences for Types of Pornography by Gender

Based on Figure 12, teens were asked to select all that applied in their pornography genre preferences. The porn that they liked the most came out strongly in favor of hardcore porn in both male and female teens, with 81.6% of male teenagers selecting this option equating to 408 and 68.8% of female teenagers also selecting this option equating to 344. Some 70.8% of male teenagers selected softcore, equating to 354, while 85.2% of female teenagers favored this type of porn, equating to 426 respondents. Noticeably, fewer teenagers reported viewing violent content alone—the 24% of males, 120, and 5.2% of females, 26,—and BDSM content—11.2% of males, 56, vs. 10.4% of females, 52. These results indicate that male and female teenage youths have different patterns of pornography use.

Gender differences in pornography types might be related to the expectations put into young people by their societal environment and mass media portrayal. It may be that a preference for hardcore and softcore pornography speaks to its higher availability and normalization across media. There is also, in general, a disinterest in violent and BDSM content, showing aversion to hard forms of pornography in the female group. The preferences and their implication hereby invite further investigation to understand their consequence on sexual development and gender dynamics.

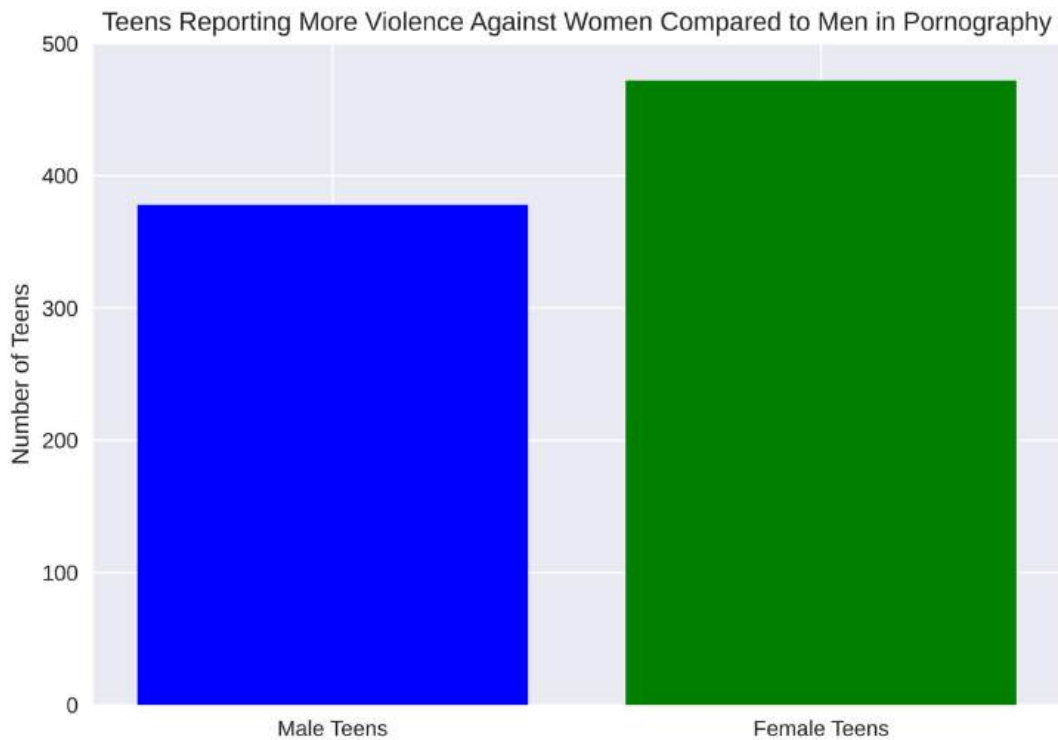


Figure 13: Amount of Teens Reporting More Violence Against Women Compared to Men in Pornography, By Gender

As shown in Figure 13, a strong majority claimed to have seen more violence against women than against men. Precisely, 75.6% of male participants (378) and 94.4% of female participants (472) affirmed so. This shows the reality of modern pornography and the gender patterns of violence in it.

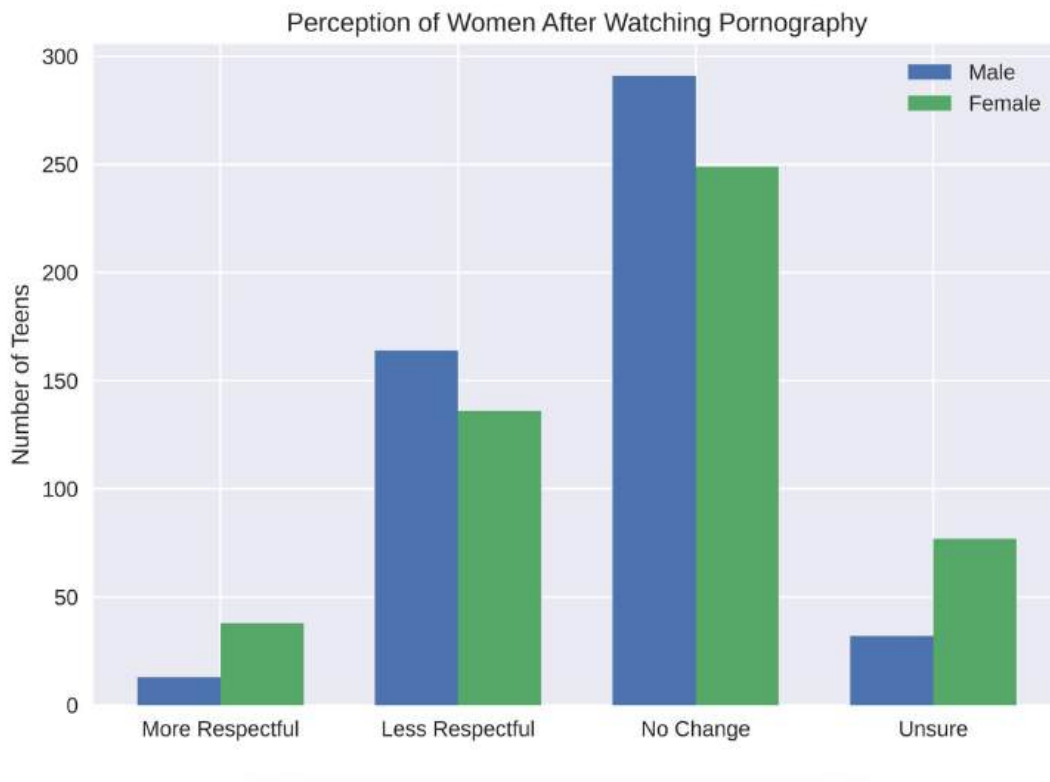


Figure 14: Teens’ Perceptions of Women After Viewing Pornography, by Gender

The chart above, in Figure 14, shows a contrast in terms of respect perceived toward women after watching pornography, as expressed by male and female teens. While majorities of both sexes have answered "No Change" to the change in their perception after watching pornography, the percentage of male teens who feel "Less Respectful" toward women, at 32.8 percent, is somewhat larger than that of female teens at 27.2 percent.

Some interesting additional comparisons came from the feelings toward women after watching pornography: more female teens reported themselves to be "More Respectful" towards women than their male counterparts. This could indicate that some content might be portraying women in an empowering or positive light, to which females are more receptive.

The fact that there is a greater percentage of female teens, 15.4%, who are "Unsure" compared to male teens at 6.4% of the impact, suggests that the former might be more reflective or even unsure about the impact of pornography on their perceptions, hence betraying a more complex engagement with the content.

These findings give an insight into how pornography subtly shapes young people's views of gender, suggesting that more education or discussion is needed around media literacy and gender in pornography.

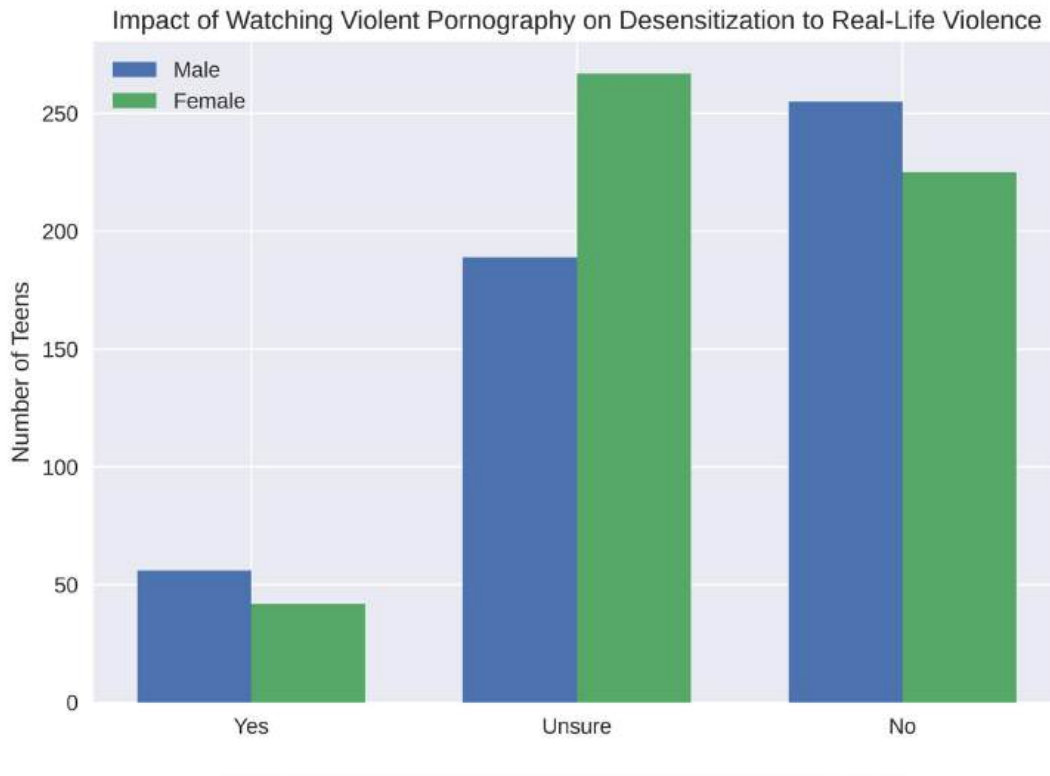


Figure 15: Teens’ Perceptions on the Impact of Violent Pornography to Desensitization to Real-World Violence

Figure 15 above indicates that most of the male and female teens break off from the view of being desensitized towards violent acts against women in real life after viewing pornographic violence, as 51% of males and 45% of females responded with "No." However, a significantly high percentage of both remained "Unsure," specifically among females with 53.4%, as compared to males having only 37.8%. This should also explain some ambiguity or confusion over the effects of such content, maybe due to varying levels of exposure or personal sensitivity.

A lower percentage of teens answered in the affirmative "Yes": 11.2% of males and 8.4% of females, which may indicate that though some feel desensitized by the media, it is not felt as the predominant response. These insights underline the complexity in media influence on youth perceptions of violence and call for more research and dialogue to further understand the possible desensitization effects coming from violent media content.

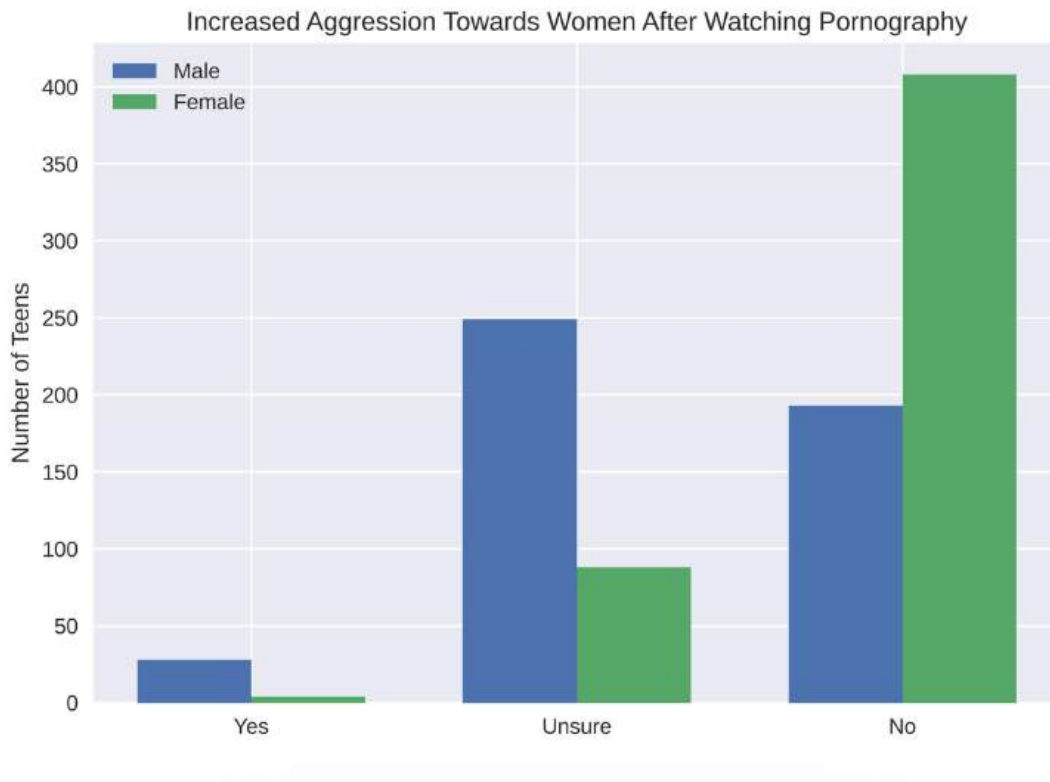


Figure 16: Teens’ Perceptions on Their Aggression towards Women After Viewing Pornography

Figure 16 indicates that a small number of male teens reported that viewing pornography increased their aggressive thoughts toward, or behaviors of aggression against, women. Although this may seem like a low total, the number nevertheless reflects 28 individuals, which is quite a problematic number when considering the possible harm related to such acts. This shows the need for awareness and interventions informed by how media, such as pornography, can affect young people's attitudes and behaviors towards violence.

The vast majority—81.6 percent—of young women reported no increase in aggression, a considerably higher amount when compared to young men. This result, at the group level, very strongly suggests that the effects of pornography views on aggression will certainly vary across gender, perhaps because the consumption patterns may differ or because there is another sort of psychological impact.

This is further supported by the high percentage of teens who were "Unsure" in both males, 52.98%, and females, 17.6%, of the impact, indicating a need to create more dialogue and education on the impact of one's consumption of violent or aggressive content.

Phase 2 Results

Statement	Consumption	Male Teens	Female Teens	Statement Number
I judge a woman's sexual behavior differently than I do a man's sexual behavior	HIGH CONSUMPTION	72.0	4.8	1
I judge a woman's sexual behavior differently than I do a man's sexual behavior	LOW CONSUMPTION	38.4	3.2	2
A woman's main life purpose is to be a mother, above all else	HIGH CONSUMPTION	30.8	23.2	3
A woman's main life purpose is to be a mother, above all else	LOW CONSUMPTION	27.2	10.4	4
Women are somewhat responsible for protecting themselves from rape and sexual assault	HIGH CONSUMPTION	66.0	10.8	5
Women are somewhat responsible for protecting themselves from rape and sexual assault	LOW CONSUMPTION	22.4	7.2	6
Men are overall better at sports than women are	HIGH CONSUMPTION	62.4	2.4	7
Men are overall better at sports than women are	LOW CONSUMPTION	32.8	1.6	8
There are circumstances in which a man hitting a woman, NOT IN SELF DEFENSE, is acceptable	HIGH CONSUMPTION	16.4	0.8	9
There are circumstances in which a man hitting a woman, NOT IN SELF DEFENSE, is acceptable	LOW CONSUMPTION	4.4	1.2	10
A husband has the right to punish his wife if he feels she is not living up to her duties	HIGH CONSUMPTION	14.4	5.2	11
A husband has the right to punish his wife if he feels she is not living up to her duties	LOW CONSUMPTION	2.8	0.4	12
If a man assaults a woman when he is intoxicated, he should not be held fully accountable	HIGH CONSUMPTION	53.6	7.2	13
If a man assaults a woman when he is intoxicated, he should not be held fully accountable	LOW CONSUMPTION	12.8	2.4	14
Rape is a naturally occurring, unavoidable, biological reaction by men towards women	HIGH CONSUMPTION	3.6	0.8	15
Rape is a naturally occurring, unavoidable, biological reaction by men towards women	LOW CONSUMPTION	0.4	0.0	16
Women are overdramatic when they are assaulted	HIGH CONSUMPTION	4.8	0.4	17
Women are overdramatic when they are assaulted	LOW CONSUMPTION	0.0	0.0	18
It is possible for a woman to encourage being assaulted by behaving in certain ways	HIGH CONSUMPTION	28.8	1.2	19
It is possible for a woman to encourage being assaulted by behaving in certain ways	LOW CONSUMPTION	7.6	0.4	20

Figure 17: Teen's Rates of Agreement with Statements Promoting Misogyny and Violence Against Women, By Gender and Consumption Group

Male Teens

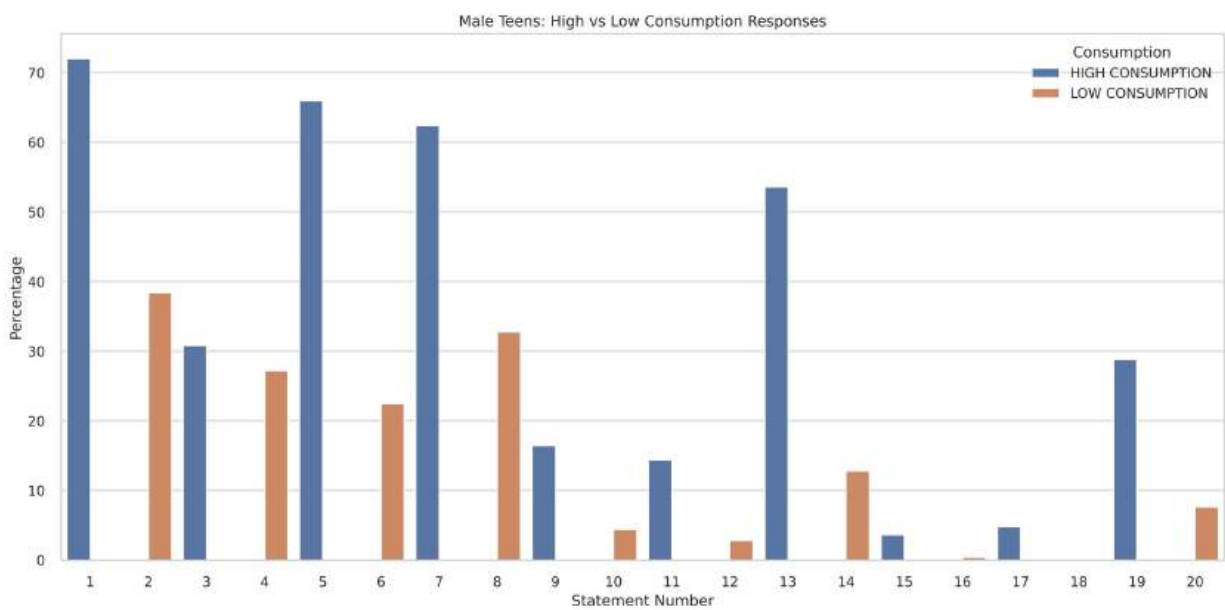


Figure 18: Male Teens' Rates of Agreement with Statements Promoting Misogyny and Violence Against Women, By Media Consumption Group

Figure 18, a graph of the male teenagers' responses to statements that promote violence against women and misogyny (Figure 17) immediately shows huge differences between high and low consumers of misogyny in media. One can easily notice how big of a difference in attitude and beliefs exists across these two groups just from this data.

Among male teens with high consumption of misogynistic media, 72.0% were far more likely to judge a woman's sexual behavior more harshly than a man's compared to the proportion among those with less consumption of this kind of media, which was 38.4%, an additional 33.6 percentage points. (Statements 1-2, Figures 17 & 18) Similarly, 30.8% of high-consumption male teens believe that women's most important role in life is the maternal one, whereas 27.2% of their low-consumption peers hold this view, which means a difference of 3.6 percentage points. (Statements 3-4, Figures 17 & 18)

An even greater difference is noted in the belief that women are somewhat responsible for protecting themselves from rape and sexual assault, with 66.0 percent of high consumption male teens agreeing with this statement compared to only 22.4 percent of low consumption male teens, a full 43.6 percentage points apart. (Statements 5-6, Figures 17 & 18) However, while 62.4 percent of high consumption male teens agree with the statement that men are generally better at sports than women, only 32.8 percent of low consumption male teens do, a difference of 29.6 percentage points. (Statements 7-8, Figures 17 & 18)

High-consumption male teens are also more likely to accept circumstances in which a man hitting a woman, not in self-defense, is deemed acceptable; 16.4 percent of them have agreed with this statement, compared to 4.4 percent for low consumption male teens—an absolute 12.0 percentage-point difference. (Statements 9-10, Figures 17 & 18) Additionally, whereas only 2.8 percent of low-consuming male teens aged 15-18 years reported that a husband has the right to punish his wife if he feels she is not living up to her duties, as much as 14.4 percent of their high-consuming counterparts did, a gap of 11.6 percentage points. (Statements 11-12, Figures 17 & 18)

Whereas 53.6% of high consumption male teens say that a man who assaults a woman when he is drunk should not be held fully responsible, only 12.8% among low consumption male teens hold this view—that is, a gap of 40.8 percentage points. (Statements 13-14, Figures 17 & 18) Furthermore, 3.6% of high-consuming male teens believe that rape is a naturally occurring, unavoidable biological reaction of men toward women, as compared to 0.4% of low-consuming male teens, which presents a difference of 3.2 percentage points. (Statements 15-16, Figures 17 & 18)

Furthermore, 4.8% of high consumption male teens, compared with 0.0% of low consumption male teens, agree that women over dramatize situations of assault. (Statements 17-18, Figures 17 & 18) Finally, 28.8% of high consumption male teens agree that a woman can be responsible for an assault by the way she behaves. The respective figure among low consumption male teens is 7.6%, representing a difference of 21.2 percentage points. (Statements 19-20, Figures 17 & 18)

The data is very clearly indicative that male teens who have highly consumed misogynistic media are much more likely to support attitudes related to violence against women and misogyny than those who have low consumption. Percentage-wise, the differences are dramatic for many of these, which goes on to show an impact on a young person's inference by media consumption regarding attitude towards gender and violence. The high consumption group always maintains higher percentages in accepting statements justifying and downplaying violence against women and reinforcing gender stereotypes that further misogynistic views. On the other hand, the low consumption group has very low percentages for agreeing with the same statements, indicating a more egalitarian and less violent attitude towards women.

The results demonstrate a disturbing influence of media consumption on the shaping of attitudes and beliefs among male teens: higher consumption goes with higher acceptance of violence against women, especially more misogynistic points of view.

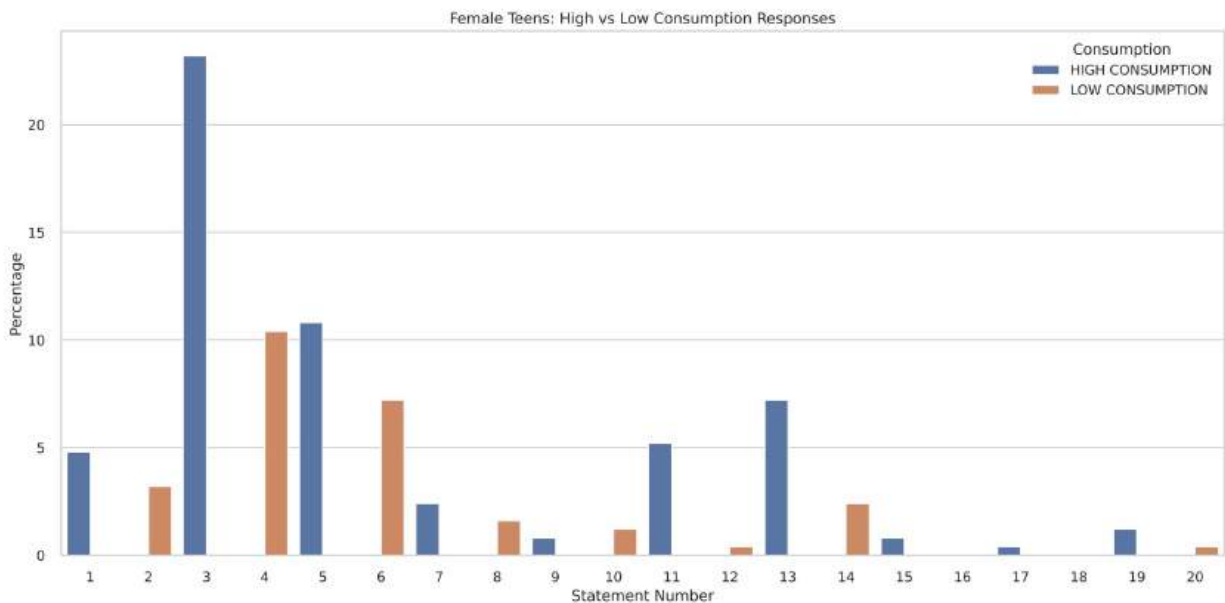


Figure 19: Female Teens’ Rates of Agreement with Statements Promoting Misogyny and Violence Against Women, By Media Consumption Group

Female Teens

The bar graph above, in Figure 19, of female teenager's responses to statements that promote violence against women and misogyny shows large differences between high and low consumption individuals. The two different graphs drawn by these data show huge differences in attitude and belief between these two groups.

Female teenagers who consume this type of media frequently are 1.6 percentage points more likely to agree that women, sexually, should be judged differently than men. (Statements 1-2, Figures 17 & 19) Specifically, the belief that being a mother represents a greater life purpose for women above all else, 23.2% of the high-consumption female teenagers agree with this

statement, as opposed to only 10.4% of low-consumption female teens—an absolute difference of 12.8 percentage points. (Statements 3-4, Figures 17 & 19)

Also important is the divergence IDV noted in the belief that women are somewhat responsible for protecting themselves from rape or sexual assault, whereby 10.8% of high consumption female teens agree with this statement against 7.2% of low consumption female teens, which is a difference of 3.6 percentage points. (Statements 5-6, Figures 17 & 19) Furthermore, 2.4% of high consumption female teens agreed that men are generally better at sports than women, compared with 1.6% of low consumption female teens, a difference of 0.8 percentage points. (Statements 7-8, Figures 17 & 19)

Interestingly, the belief in circumstances in which a man hitting a woman, not in self-defense, being acceptable is held by only 0.8% of high consumption female teens, compared with 1.2% of low consumption female teens, showing a difference of -0.4 percentage points. (Statements 9-10, Figures 17 & 19) Additionally, 5.2 percent of high consumption female teens agree a husband has the right to punish his wife if he feels she is not living up to her duties, compared with only 0.4 percent of low consumption female teens, a gap of 4.8 percentage points. (Statements 11-12, Figures 17 & 19)

On the accountability of a man who assaults a woman when he is intoxicated, 7.2% of high-consuming female teens consider him not fully accountable, against 2.4% of low-consuming female teens—a gap of 4.8 percentage points. (Statements 13-14, Figures 17 & 19) Regarding sexual assault, 0.8% of the female high-consuming teenagers agreed that rape is a naturally occurring and unavoidable biological reaction of men toward women, compared to 0% of the female low-consuming teenagers—an absolute difference of 0.8 percentage points. (Statements 15-16, Figures 17 & 19)

0.4% of high consumption female teens, compared with 0.0% of low consumption female teens, believe women can be overdramatic when assaulted, a difference of 0.4 percentage points. (Statements 17-18, Figures 17 & 19) Lastly, 1.2% of high consumption female teens believe a woman can invite being assaulted by her behavior, compared with 0.4% of low consumption female teens, a difference of 0.8 percentage points. (Statements 19-20, Figures 17 & 19)

Data conclusively shows that female teenagers who consume large amounts of misogynistic media are more likely to harbor attitudes that perpetuate violence against women/misogyny, compared to those rated as low consumers. Though percentage differences were generally smaller than the male teenager results, they still show effects of media consumption on attitudes toward gender and violence. The high consumption group response is significantly higher for statements supporting or minimizing violence against women, justifying stereotypes about rigid gender roles, and fostering misogynistic beliefs. In contrast, in the low consumption group, lower percentages agreed with statements on perception of women, which indicates attitudes of a more egalitarian and less violent nature toward women.

This means that, after all, this research will point out that media consumption is related to attitude and belief formation among female teens. Higher rates of misogynistic media

consumption show a more permissive attitude towards violence against women and a more misogynistic set of views.

Discussion — Phase 1

Social Media

The research findings that looked to determine how misogyny propagated through the variety of media modes affect the real-world attitudes of teen males and females in light of violence against women, focused on three points: Social Media, Popular Music, and Mainstream Pornography. This section is devoted to findings at the level of social media.

64.7% of the respondents used screens for more than 8 hours a day, and this was almost equally distributed between both genders. When asked about their primary use for their phones, most respondents, 81.3%, indicated social media applications, with the trend more toward female teenagers at 489 responses, as opposed to male teenagers at 324 responses.

Many teens reported frequently seeing misogynistic content. As many as 77.4% of male teens and 84.6% of female teens said they would often come across such content. More specifically, this prevalence turned out to be much higher in females, which shows that social media is going to take a toll on the young generation and further feed more biases based on gender.

For male teens, a far more ambivalent reaction to misogynistic content was reported: 38.0 percent of the time it was met with indifference; the next most common was concern, at 21.8 percent; then distress at 13.8 percent. The young women responded that they had stronger feelings of emotion: the top one being feelings of sadness, at 28.4 percent; feelings of anger ranked as the next highest emotion, at 26.4 percent; then concern ranked at 21.6 percent. The findings underline that male teens are often unaware of the dangers of sexism in media and ignorant to the promotion of hatred, while female teens may be all too aware.

On perceptions of impact on real-world attitudes, 62.6% of male teens agreed that social media misogyny impacts real-world attitudes toward women; 30.6% did not agree, and 6.8% were not sure. A greater proportion of the female teen sample acknowledged the impact at 79.2%, with 4.2% disagreeing and 16.6% unsure. It can be seen from these results that there is a clear, sharp gender gap in terms of recognizing real-world implications from social media misogyny: this makes the issue of education and awareness very pressing indeed.

Specifically, the key implications of this study are those oriented toward policies on more stringent content moderation, educational programs, and targeted interventions to help curtail the wide spreading of misogynistic content over social media and create a safer, more respectful digital environment for all users.

Music

The data analysis underlines trends as to the prevalence of misogynistic music, mostly rap/hip-hop, and how it impacts teen listeners.

Data showed that 65.8% of the teenagers surveyed, or 658 out of 1000, listened largely to rap/hip-hop music. That is the most lyrical of the options given, some of which were pop, country, folk, classical, jazz, indie, R&B, rock, metal, and gospel. More interesting was the gender difference in this preference: while 442 listen to rap/hip-hop music are male, 216 of them were females. Oppositely, 34.2 percent of teens, or 342 out of a total of 1000 who prefer other genres, are 58 males and 284 females. This clearly indicates a high liking for rap/hip-hop music problematically preferred by teenagers, more so by male listeners, which identifies key trends in musical taste between different genders.

Following is the prevalence of misogynistic content encoded in the music. A majority of those teens, 93.3 %, agreed to the presence of general misogyny in the music. The males, at 49.2 %, slightly above the females at 44.1 %, showed response rates for general misogyny. In this case, it is inference, based on the evidence, that there is general recognition among teens of themes of misogyny across the music they have consumed. Furthermore, 69.5% of teenagers reported that they had listened to songs with lyrics which explicitly describe violence against women. The statistic in itself is deeply alarming because it shows evidence that half of the teenage population is venturing into content that normalizes and glorifies violence against women.

When focusing on specific songs like "Blinded Lines" by Robin Thicke, "Tron Cat" by Tyler the Creator, and "Kim" by Eminem aids in giving a better understanding of this phenomenon. All of the above songs were selected based on previous surveys and data on popular music among teens. The lyrics of "Blurred Lines" raise some serious concerns about the depiction of consent and the objectification of women. There are a whole host of issues, including referencing drugging and non-consensual sex, that further objectify women and truly perpetuate rape culture. This becomes especially concerning with regard to teenage listeners, who may be powerfully swayed by such lyrics and foster unhealthy attitudes toward sex, consent, and the treatment of women.

The lyrics to both "Tron Cat," by Tyler the Creator, and "Kim," by Eminem, are extremely violent and misogynistic, detailing sexual assault and murder. In "Tron Cat," vivid descriptions are given of the dismemberment and disposal of female victims' bodies. Conversely, "Kim" provides a scenario of horrific domestic violence in which the singer threatens, berates, and eventually murders his partner. The use of profanity, violent threats, and the lack of empathy or remorse create an atmosphere of terror and helplessness.

Reactions of participants to the violent content in these songs are also telling. As indicated, a high percentage of teenagers, 78.4%, responded to the survey saying that they were disturbed by the violent lyrics. A greater number of females, 82.1%, responded as disturbed, compared with males, at 75.6%. This does show some general awareness of the negative character of this type of content. However, 21.6% of teens responded that they were neither upset nor offended, with a greater percent of males 24.4% than females 17.9%. A further level of this study's findings is that even more worrisome is this finding, which may support that there is a lack of critical engagement with the important harmful messages being conveyed. A small, yet

noteworthy proportion of male teens, too, reported feeling empowered by the songs, 7.6%, a suggestion that some misogynistic messages may reinforce certain harmful gender norms and attitudes in young males. This is similar for both males and females who reported feelings of joy, 17.2 percent and 18.2 percent respectively, hence showing that there is some instance when the enjoyment of the music can mask the negative implications of their lyrics.

Further research on the mechanisms of how such lyrics affect young people, especially in formulating ways to curb this trend, needs to be conducted.

Given the potent impact of music on young minds, the prevalence of misogyny in lyrics raises concerns. One will document, from a diversity of studies, that repeated exposure to violent and derogatory content influences behavior and attitudes; therefore, high exposure rates among teens could further the normalization of negative gender attitudes and behaviors.

What is required here is some kind of intervention, which brings together different stakeholders in this perspective. It should, therefore, be sharply critical about the possible impact of such media on the music industry. Impose upon artists, producers, and label companies the self-regulation to create works that genuinely treat all genders with respect and give all of them honor equally. Lastly, it is also one of the prime functions of educators as well as parents to make sure that proper media literacy is learned and inculcated at the level of teenagers so that they could critically evaluate the feed of music they are consuming. Regulatory measures or guidelines concerning distribution manners of explicitly misogynistic content may be taken into consideration with balancing of the goal to combat negative messages and maintenance of artistic freedom.

Therefore, with more reason, it underlines an immediate call for collective effort in attenuating impacts that arise from misogynistic music content toward a culture of greater inclusivity and respect.

Pornography

The following research operationally defines the consumption patterns of pornography usage among teenagers with results offering numerous discrepancies between male and female teenagers in their watching habits and preferences, along with some possible psychological and social consequences that could occur from the use of this type of media.

The results indicate clear differences in pornography-watching behavior among male and female adolescents. While 25.2% of males responded to having watched pornography daily, 82% of the male respondents indicated watching it during the past week. For female respondents, the frequency is 12.4% daily and 50.4% weekly. Such a pattern of results seems to reflect a potentially problematic gender gap regarding the consumption of pornography. It could, therefore, be that higher male consumption rates reflect broader social and cultural influences in which males are likely to have more exposure and encouragement to such sexual content.

The survey has also exposed important preferences for different categories of pornography among both sexes. The overwhelming proportion of 81.6% male teens, or 408 responses, and 68.8% of female teens, or 344 responses, chose hardcore pornography as their

preferred category. Further, 70.8% of male teens, or 354 responses, and 85.2% of female teens, or 426 responses, identified softcore pornography as their preferred category. Results clearly point to a very high level of exposure to such explicit content bound to affect psychological development and social relationships in a number of ways.

The belief that, after being exposed to violent media content, most teenagers felt desensitized to violence was interrogated in this study. Large numbers in both groups remain "Unsure," although this is higher in females at 53.4% compared to males at 37.8%. This therefore creates some level of ambiguity or confusion over the effects of such content, possibly due to varying extents of exposure and personal sensitivity. The lower percentage of teens that answered with a clear "Yes" (11.2% of males and 8.4% of females) suggests that though some do feel desensitized, it is not the answer across the board. This could show that while there is an impact from violent content in media, many teens are either capable of differentiating between what they see in the media and actual reality or have not been hit too hard in terms of their empathetic response toward actual violence.

The data shows that 5.96% of male teens reported increased aggressive thoughts or behaviors toward women as a result of watching pornography. Although this percentage alone may not seem high, it still equates to 28 individuals, a large and concerning number when considering the potential harm of such behaviors. This therefore brings about awareness and the need for interventions targeting how media, such as pornography, can further change young people's attitudes and behaviors toward violence. About 81.6% of female teens showed no increase in aggression, which is substantially higher compared to males. It has been observed in previous studies that the effect of pornographic movies on aggression might vary greatly between genders due to different use patterns or psychological effects. (Bridges et al.) (Huesmann & Taylor)(The effects of sexually arousing and violent films on aggressive behavior) This comparative analysis between male and female teenagers has returned one trend in common: high media consumption is correlated with a choice of more explicit content and possible desensitization to violence. These findings are supported by previous studies that showed that exposure to explicit and violent media material increases aggressive behavior or causes desensitization to violence. (Krahé et al.)(Anderson et al.) (The Role of Violent Media Preference in Cumulative Developmental Risk for Violence and General Aggression - Journal of Youth and Adolescence) Such normalization of deviant and violent themes among high exposure groups is not conceptual; on the contrary, it is a very real problem with critical implications for society.

In summary, this study inspects the critical influence of pornography consumption on teenagers' attitudes and behaviors. These large differences in male and female teenagers' consumption patterns and preferences further call for focused research and interventions to reduce the possible negative effects of explicit media.

Discussion — Phase 2 Results

The current study offers an estimation of how women-hating media consumption is related to the attitudes and beliefs of male and female teenagers, in which it is associated with higher media consumption and high aggression toward women, with reduced disturbance against violence toward women.

These results demonstrate some very large and alarming differences related to high, as opposed to low, misogynistic media consumption male teens. Specifically, among these, 72.0% of male teens who report a high consumption of such media agree that a woman's sexual behavior should be judged differently than a man's, compared to just 38.4% of the male teens who reported a low consumption of this media model. This epitomizes a difference of 33.6 percentage points, which is very significant. This highly significant difference underlines how the normalization of double standards and gender biases becomes common among high consumers of misogynistic content.

Additionally, while only 22.4 percent of low-consumption male teens agreed that women are somewhat responsible for protecting themselves from rape and sexual assault, 66.0 percent did among high-consumption male teens—an incredibly large gap of 43.6 percentage points. As this is a measure of victim-blaming attitude, an attitude that helps cement harmful stereotypes in people's minds and thus leads to poorer chances of adequately dealing with sexual violence, the difference is deeply disturbing.

It is also the case with the belief that men are better in sports, with 62.4% of the high consumption male teens compared to 32.8% of their peers in the low consumption category, which is a gap of 29.6 percentage points. Each of these factors further enhances the traditional gender roles and also doubly acts to undermine the gains made toward equality between the sexes in all walks of life, including this one in sports.

Alarming is the fact that 16.4 percent of high-consumption male teens find it acceptable for a man to hit a woman, not in self-defense, while only 4.4 percent of their low-consumption counterparts do, leaving a further difference of 12 percent. The acceptance of violence against women becomes a clear indicator for the desensitization toward violence that could be the result of media consumption with degrading content against women.

Furthermore, 14.4% of high consumption male teens believe that a husband has the right to punish his wife if he feels she is not living up to her duties, compared with 2.8% of low consumption male teens—a gap of 11.6 percentage points. This belief in patriarchal control and punishment further underlines regressive attitudes fostered by misogynistic media consumption.

Moreover, perceptions of lesser accountability for intoxicated assaulters are more likely to be held by high consumption male teens, compared to low consumption male teens, at 22.4 percent—a gap of 31.2 percentage points. This dangerous attitude belittles the crime of assault and condones violent behavior under the influence of alcohol.

This is also evident with regard to the impact of misogynistic media on female teens. For example, whereas 30.8 percent of high-consuming female teens report that the central function of life for a woman lies in being a mother, above all else, only 27.2 percent of low-consuming

female teens agreed—a disparity of 3.6 percentage points. This reinforces traditional gender roles and limits the perceived potential of women beyond motherhood.

Moreover, following a trend, 5.2% of high-consumption female teens strongly believe that a husband has the right to punish his wife when he feels she is not living up to her duties, compared with 0.4% of low-consumption female teens—a gap of 4.8 percentage points. This acceptance of patriarchal control is worrying and underlines internalization on the part of female teens.

Additionally, the belief in reduced accountability for intoxicated assaulters is held by 7.2 percent of high-consumption female teens, as opposed to only 2.4 percent of the low-consumption female teens, which represents a gap of 4.8 percentage points. This evidences a troubling trend of excusing violent behavior and lessens the serious nature of assault.

Comparative analysis between the high and low consumption groups comes out with the fact that quite consistently, with an increase in media consumption, attitude towards women becomes regressive and injurious. Accordingly, these findings were replicating earlier researches, such as the studies by Ferguson and Kilburn, which prove exposure to violent media contents increases aggression and leads to desensitization towards violence.

The normalization of such misogynist beliefs and behaviors in groups of high consumption is thus not a theoretical but very real problem, bearing important implications for society. Rather, it illustrates quite suitably why interventions aimed at reducing exposure and increasing media literacy are so urgently needed. It is possible to chart the course toward a more just and respectful society for all people only by grasping these issues.

Limitations

These limitations raise broad implications when considering various forms of media, such as social media, music, and pornography. The results from this study were based on individuals located in one geographic area; hence, it cannot be determined with complete certainty whether the results would generalize easily to other parts of the world where individuals may use these many forms of media quite differently. Furthermore, much of the data is self-reported; therefore, recall bias could occur, especially in the case when participants report their intake of various media, whose consumption may not always be remembered.

The study did not distinguish between the different media content and their impact on attitudes in regards to misogyny; for example, whether social media interactions were invidious or supportive, sub-genres of music, and sub-categories of pornography. This is what is critical because different contents may have different impacts on different people. Moreover, though it did look into how often people consume the media, it did not consider the duration of time that a subject would spend in a single session, which is very important in determining whether brief moments of intense media consumption have the same impact as longer periods.

Several other variables of probable influence were not taken into consideration, such as financial requirements needed for media consumption, personal values and beliefs, and individual kinds of behavior singled out by the various modes of media use. There is also a call

to break down in greater detail resilience variables that may reduce probable negative effects from media consumption, examining support variables and educational programs.

There is, therefore, an apparent requirement for future studies that will include not only quantitative metrics but also qualitative insight into how various forms of media change mental health and behavior. These studies should consist of the interplay of personal faith, community support, and mediated consumption to give a full view of the complex interactions between media consumption and individual behavior in the digital age.

Conclusion

The current study finds strong empirical evidence to assert that the exposure to misogynous and violent media content is predictive of problematic attitude endorsement toward women and violence in teenagers and most characteristically in male teens. The results revealed a remarkable difference in terms of percentage of acceptance of misogynistic and violence-approving statements across high and low media-consuming groups. Its findings are in line with earlier research showing links between exposure to media violence and increased levels of aggression, and the function of sexualized media in guiding support for harmful beliefs about gender and sexual violence.(Huesmann & Taylor) (Anderson et al.)(Anderson et al.) This is also a place where it really is necessary to further investigate the possibly differential impact on male and female youth.

Violence against women is very prevalent in contemporary society. While causes of this condition are complex and multifactorial, over the years, growing concern has been raised regarding the role of media in normalizing, excusing, and even encouraging violence toward women. (Sarieddine)(REPRESENTATIONS OF VIOLENCE AGAINST WOMEN IN THE MASS MEDIA) The worrying truths about the effects of media on attitudes to VAW cannot be overemphasized.

Moreover, domestic violence surged to new records again in 2023, exceeding any level over the last 50 years as documented by the National Domestic Violence Hotline in 2023. With this trend, it also needs definitive solutions to stop the causes of domestic violence, especially from the media.

To put it differently, the results of this research acted as a wake-up call to underscore the importance of media literacy programs in making young people aware of the effect of the media on their values and actions. Furthermore, these norms are to be wrestled with altercation and changed by policymakers and educators in environments. The attention received by the media concerning misogyny and sexism will be a result of endeavors directed at taming dangerous effects on attitudes—the only hope toward reducing violence against women in our society.

Works Cited

- Cookingham, L M., & Ryan, G L. (2015, February 1). The Impact of Social Media on the Sexual and Social Wellness of Adolescents. Elsevier BV, 28 (1), 2-5.
<https://doi.org/10.1016/j.jpap.2014.03.001>
- Khalaf, A., Alubied, A A., Khalaf, A., & Rifaey, A A. (2023, August 5). The Impact of Social Media on the Mental Health of Adolescents and Young Adults: A Systematic Review. Cureus, Inc.. <https://doi.org/10.7759/cureus.42990>
- Habib, H., Srinivasan, P., & Nithyanand, R. (2022, November 7). Making a Radical Misogynist. Association for Computing Machinery, 6(CSCW2), 1-28.
<https://doi.org/10.1145/3555551>
- TECHNOLOGY-FACILITATED GENDER-BASED VIOLENCE. (n.d).
https://www.svri.org/sites/default/files/attachments/2018-07-24/ICRW_TFGBVMarketing_Brief_v8-Web.pdf
- Wiedeman, A M., Black, J A., Dolle, A L., Finney, E J., & Coker, K L. (2015, November 1). Factors influencing the impact of aggressive and violent media on children and adolescents. Elsevier BV, 25, 191-198. <https://doi.org/10.1016/j.avb.2015.04.008>
- Schooler, C., & Flora, J A. (1996, January 1). Pervasive Media Violence. Annual Reviews, 17(1), 275-298. <https://doi.org/10.1146/annurev.pu.17.050196.001423>
- Kahlor, L A., & Eastin, M S. (2011, May 25). Television's Role in the Culture of Violence Toward Women: A Study of Television Viewing and the Cultivation of Rape Myth Acceptance in the United States. Taylor & Francis, 55(2), 215-231.
<https://doi.org/10.1080/08838151.2011.566085>
- Papadamou, K., Zannettou, S., Blackburn, J., Cristofaro, E D., Stringhini, G., & Sirivianos, M. (2021, October 13). "How over is it?" Understanding the Incel Community on YouTube. Association for Computing Machinery, 5(CSCW2), 1-25.
<https://doi.org/10.1145/3479556>
- Ging, D., & Σιαπέρα, E. (2018, July 4). Special issue on online misogyny. Taylor & Francis, 18(4), 515-524. <https://doi.org/10.1080/14680777.2018.1447345>
- Manago, A M., Ward, L M., Lemm, K M., Reed, L A., & Seabrook, R C. (2014, December 12). Facebook Involvement, Objectified Body Consciousness, Body Shame, and Sexual Assertiveness in College Women and Men. Springer Science+Business Media, 72(1-2), 1-14. <https://doi.org/10.1007/s11199-014-0441-1>
- Anderson, C A., Bushman, B J., Donnerstein, E., Hummer, T A., & Warburton, W. (2015, November 4). SPSSI Research Summary on Media Violence. Wiley-Blackwell, 15(1), 4-19. <https://doi.org/10.1111/asap.12093>
- Impact of Music, Music Lyrics, and Music Videos on Children and Youth. (2009, November 1). American Academy of Pediatrics, 124(5), 1488-1494.
<https://doi.org/10.1542/peds.2009-2145>

- Primack, B A., Gold, M A., Schwarz, E B., & Dalton, M A. (2008, September 1). Degrading and Non-Degrading Sex in Popular Music: A Content Analysis. *SAGE Publishing*, 123(5), 593-600. <https://doi.org/10.1177/003335490812300509>
- Bridges, A J. (2019, January 1). Pornography and Sexual Assault. *Springer Nature*, 129-149. https://doi.org/10.1007/978-3-030-23645-8_8
- Seabrook, R C., Ward, L M., & Giaccardi, S. (2019, September 1). Less than human? Media use, objectification of women, and men's acceptance of sexual aggression.. *American Psychological Association*, 9(5), 536-545. <https://doi.org/10.1037/vio0000198>
- Manzoor, A., Rehman, D., & Rauf, S. (2016, September 8). Analysis Of Gender Stereotypes In Movies. *University of Karachi*, 4(1), 95-109. <https://doi.org/10.46568/pjass.v4i1.298>
- Ging, D. (2017, May 10). Alphas, Betas, and Incels: Theorizing the Masculinities of the Manosphere. *SAGE Publishing*, 22(4), 638-657. <https://doi.org/10.1177/1097184x17706401>
- Anderson, C A., Berkowitz, L., Donnerstein, E., Huesmann, L R., Johnson, J D., Linz, D., Malamuth, N M., & Wartella, E. (2003, December 1). The Influence of Media Violence on Youth. *SAGE Publishing*, 4(3), 81-110. https://doi.org/10.1111/j.1529-1006.2003.pspi_1433.x
- Russell, E L. (2021, January 1). Alpha Masculinity. *Springer Nature*. <https://doi.org/10.1007/978-3-030-70470-4>
- Kupers, T A. (2005, January 1). Toxic masculinity as a barrier to mental health treatment in prison. *Wiley-Blackwell*, 61(6), 713-724. <https://doi.org/10.1002/jclp.20105>
- Solea, A I., & Sugiura, L. (2023, August 12). Mainstreaming the Blackpill: Understanding the Incel Community on TikTok. *Springer Science+Business Media*, 29(3), 311-336. <https://doi.org/10.1007/s10610-023-09559-5>
- Shattering Myths and Lives: A Report on Image-Based Sexual Abuse. (n.d). <https://claremcglynn.files.wordpress.com/2019/10/shattering-lives-and-myths-revised-aug-2019.pdf>
- Menzie, L. (2020, August 18). Stacys, Beckys, and Chads: the construction of femininity and hegemonic masculinity within incel rhetoric. *Taylor & Francis*, 13(1), 69-85. <https://doi.org/10.1080/19419899.2020.1806915>
- The women who spoke out against the Tate brothers: rape, violence and sexual exploitation. (2023, May 31). <https://www.riseproject.ro/en/projects/the-women-who-spoke-out-against-the-tate-brothers-rape-violence-and-sexual-exploitation/>
- Morris, L. (2023, February 2). Tate rape allegations, text exchanges with women detailed in court document. https://www.washingtonpost.com/world/2023/02/02/andrew-tate-rape-investigation-romania/?utm_source=twitter&utm_medium=social&utm_campaign=wp_main

- Bratich, J Z., & Banet-Weiser, S. (2019, September 22). From Pick-Up Artists to Incels: Con(fidence) Games, Networked Misogyny, and the Failure of Neoliberalism. USC Annenberg School for Communication & Journalism, 13. <http://eprints.lse.ac.uk/102183/>
THE SEVEN P's OF MEN'S VIOLENCE. (n.d).
<http://www.michaelkaufman.com/wp-content/uploads/2009/01/kaufman-7-ps-of-mens-violence.pdf>
- Cosma, S., & Gurevich, M. (2019, August 8). Securing sex: Embattled masculinity and the pressured pursuit of women's bodies in men's online sex advice. SAGE Publishing, 30(1), 42-62. <https://doi.org/10.1177/0959353519857754>
- Gadd, D., Fox, C L., & Hale, B. (2013, June 1). Young people and violence against women. Taylor & Francis, 92(1), 36-37. <https://doi.org/10.1080/09627251.2013.805377>
- Shaw, S M. (1999, January 1). Men's leisure and women's lives: the impact of pornography on women. Routledge, 18(3), 197-212. <https://doi.org/10.1080/026143699374925>
- Boyle, K. (2011, November 1). Producing abuse: Selling the harms of pornography. Elsevier BV, 34(6), 593-602. <https://doi.org/10.1016/j.wsif.2011.09.002>
- Bridges, A J., Wosnitzer, R., Scharrer, E., Sun, C., & Liberman, R. (2010, October 1). Aggression and Sexual Behavior in Best-Selling Pornography Videos: A Content Analysis Update. SAGE Publishing, 16(10), 1065-1085.
<https://doi.org/10.1177/1077801210382866>
- Huesmann, L R., & Taylor, L D. (2006, April 1). The Role of Media Violence in Violent Behavior. Annual Reviews, 27(1), 393-415.
<https://doi.org/10.1146/annurev.publhealth.26.021304.144640>
- Meyer, T P. (2014, October 28). The effects of sexually arousing and violent films on aggressive behavior. <https://www.tandfonline.com/doi/abs/10.1080/00224497209550766>
- Krahé, B., Möller, I., Huesmann, L R., Kirwil, L., Felber, J., & Berger, A. (2011, April 1). Desensitization to media violence: Links with habitual media violence exposure, aggressive cognitions, and aggressive behavior.. American Psychological Association, 100(4), 630-646. <https://doi.org/10.1037/a0021711>
- Anderson, C A., Bushman, B J., Donnerstein, E., Hummer, T A., & Warburton, W. (2015, November 4). SPSSI Research Summary on Media Violence. Wiley-Blackwell, 15(1), 4-19. <https://doi.org/10.1111/asap.12093>
- Dominic, B P R H L B B J O M M. (2008, September 9). The Role of Violent Media Preference in Cumulative Developmental Risk for Violence and General Aggression - Journal of Youth and Adolescence. <https://link.springer.com/article/10.1007/s10964-008-9335-2>
- Sariieddine, M. (2018, January 1). Violence against Women: A Phenomenological Perspective. OMICS Publishing Group, 04(01). <https://doi.org/10.4172/2471-2701.1000180>
- Henricksen, M G J. (2018, January 1). REPRESENTATIONS OF VIOLENCE AGAINST WOMEN IN THE MASS MEDIA.
https://www.academia.edu/36846155/REPRESENTATIONS_OF_VIOLENCE_AGAINST_WOMEN_IN_THE_MASS_MED

An Analysis and Discussion of the Various Signal Acquisition Techniques Used in Speech Decoding Technology By Sai Akhil Madurakavi

Abstract

In the field of speech decoding, there are numerous signal acquisition techniques used to collect signals from the body and translate them into speech. The most prominent techniques are reviewed in this paper: electroencephalography (EEG), non-invasive techniques like electrocorticography (ECoG) and Utah intracortical electrode array (UIEA), electromyography (EMG) and other non-neural techniques, and functional magnetic resonance imaging (fMRI). By reviewing various papers on BCIs using these techniques, this review has identified the advantages and disadvantages of each signal acquisition technique. The paper provides a comprehensive analysis of each signal acquisition technique by discussing studies using each technique with the following factors: portability, non-invasiveness, accuracy, vocabulary size, and decoding speed. The findings show that EEG and EMG are non-invasive, portable, and have good accuracy and decoding speed but have a very limited vocabulary size. ECoG and UIEA are very accurate, have large vocabulary sizes and decoding speed, but are less portable and very invasive. fMRI is non-invasive, has a very accurate spatial resolution, and has a high potential for future development, but is not portable and has slow temporal resolution and direct speech translation speed. This review concludes with an exploration of future research directions and open questions in the field.

1. Introduction

Decoding speech from thoughts or brain signals could be considered extremely difficult due to the task's complexity and ambiguity. However, ever since brain-computer interfaces (BCIs) were developed in the 1970s, researchers have been able to understand and obtain readings of signals from the brain which could be translated into thoughts or speech¹. In the 2000s, major advancements in machine learning enabled scientists and engineers to analyze these signals and turn them into actual speech using advanced machine learning algorithms and models². Neuroscientists, neuroengineers, signal processing specialists, and machine learning experts collaborate to produce and improve speech decoding models and devices. From improving the lives of the disabled to revolutionizing communication on an immense scale, speech decoding has countless applications bound to advance humanity's technology to the next level. Currently, numerous speech decoding technologies using different signal acquisition techniques can decode one's thoughts into speech. Although major progress has been made, more specific development is required since each signal acquisition technique's speech decoding technology faces different challenges. As such, no one optimal speech decoding device exists. This paper will review the advantages and disadvantages of each signal acquisition technique used in speech decoding technology using the factors of portability, non-invasiveness, accuracy, vocabulary size, and decoding speed to identify the specific improvements that have to be made for each technology.

1.1. Biology of Speech

Humans have evolved to be exceptional language-processing communicators. To allow for our advanced communication, various regions of the brain become active when one thinks about saying something or when one starts to formulate a sentence in their mind³. Some of these regions include Broca's and Wernicke's regions which play a crucial role in language production and understanding, respectively⁴. Semantic processing occurs across various brain regions, including the temporal lobes, to associate the thought one tries to express with certain words and concepts with similar meanings⁵. In Broca's region, syntactic processing also occurs while structuring a sentence to process syntax and grammatical rules in a language⁶. Concurrently, regions in the brain, like the prefrontal cortex, that control the speech articulators (lips, tongue, vocal cords, etc), are activated to plan the precise movements of these articulators to convey the planned speech⁷. After speech planning, the brain will send signals through the spinal cord to the speech articulators and vocal cords to command them to move in a certain way to create the correct sounds, words, and sentences^{7,8}. These speech signals start across various regions in the brain and continue along the spinal cord to reach the speech articulators and other muscles and organs involved with speech⁸. The basis of speech decoding utilizes this concept and uses signal acquisition techniques to capture these speech signals in the various regions through which the signal travels to then convert into speech.

2. The Various Signal Acquisition Techniques Used in Speech Decoding

Researchers have done extensive research using technologies like EEG, ECoG, EMG, or fMRI to make efficient and accurate BCIs to successfully decode speech signals from the brain. While there are many factors to consider when studying BCIs, including preprocessing, feature extraction, training data collection, and decoding models, it is crucial to consider the most viable signal acquisition techniques in speech decoding since each technique has its advantages, disadvantages, and situational uses.

2.1. Electroencephalography (EEG)

EEG is a non-invasive, electrophysiological technique used to measure the electrical activity in the brain². More specifically, it measures the changes in the electrical activity of large groups of neurons in the region of the brain where the EEG electrodes are placed². The electrodes are placed on the scalp and can measure the action potentials of the groups of neurons². Due to the nature of the brain and its abundance of rapid electrical signals, EEG can detect the precise changes in this electrical activity in real-time². Consequently, EEG is great for real-time speech decoding due to its exceptional temporal resolution¹⁰. However, EEG lacks spatial resolution and cannot clarify whether the electrical activity is produced in deeper parts of the brain or the surface¹⁰. This means that EEG is unreliable for determining the exact region a relevant signal comes from in the practical use of speech decoding. This is because the electrodes generally show the activity for a relatively large region compared to other signal acquisition

techniques with better spatial resolution, which allows for more specific and a greater amount of brain data.

On the other hand, EEG is a very feasible signal acquisition technique because it is generally easy to use, portable, and lower cost, simpler EEG devices have recently come onto the market. However, because EEG does track electrical activity, it is greatly prone to noise from other electrical sources which can interfere with the data being collected¹¹. For example, electrical activity from technology in the environment, like power lines and lighting, and from personal electronic devices, such as televisions, computers, and phones, all contribute to the noise¹². Noise also arises from the electrical activity of physiological sources such as the heart, muscle contractions, and eye blinks¹². This requires EEG users to rely on noise removal techniques to filter out the irrelevant electrical activity and amplify the relevant electrical activity¹¹.

2.1.1 EEG-based Studies

Covert or imagined speech is a common type of speech decoded with speech decoding systems using EEG. Panachakel and Ganesan test their transfer-learning-based approach for decoding imagined speech using EEG¹³. This study used 15 participants within their 4 experiments with different prompts. The first experiment consisted of the long words “independent” and “cooperate”¹³. The second experiment prompted subjects to imagine saying short words: “in,” “out,” and “up”¹³. The third experiment consisted of vowels: “/a/,” “/i/,” and “/u/”¹³. And finally, the fourth experiment consisted of short-long words: “in” and “cooperate”¹³. In all experiments, subjects were instructed to repeatedly imagine saying the prompts without moving their vocal muscles¹³. The researchers used a 64-channel EEG headset to collect the data¹³. Their decoding algorithm achieved an accuracy between 83% and 92% for all of the prompts¹³.

A study by García-Salinas et al. also uses transfer learning to develop and test an imagined speech decoder¹⁴. In the study, researchers recorded EEG data from 27 participants using a 14-channel EEG headset¹⁴. They instructed the subjects to imagine saying 5 different Spanish words on separate occasions¹⁴. In English, these 5 words translate to “up,” “down,” “left,” “right,” and “select.”¹⁴ On average of all the participants and prompts, the decoder achieved an accuracy of $65.65\% \pm 13.39\%$ ¹⁴.

Lastly, a study by Tøttrup et al. investigated the feasibility of decoding imagined speech from EEG and combining it with motor imagery to allow individuals with motor deficiencies to operate external devices such as wheelchairs and prosthetic robotic arms¹⁵. In their study, the researchers recorded EEG with twenty-five channels in seven healthy subjects during six different actions: speaking three words (both covert and overt speech), two arm movements (both motor imagery and execution), and one idle class¹⁵. For the covert speech decoding, the participants were instructed to imagine saying three words, “go”, “stop”, and the control word “Viborg” (a Danish city), on separate occasions¹⁵. Their decoder achieved an average classification accuracy of $67 \pm 9\%$ for the covert speech¹⁵.

These three experiments use different decoding models, preprocessing techniques, and feature extraction techniques which make them all quite different from each other^{13,14,15}. Except for the first study, the other two studies have very similar, and slightly above-average accuracies^{13,14,15}. On the other hand, the first study has an exceptional accuracy that surpasses the accuracy of the other two studies^{13,14,15}. As shown in these experiments, EEG is quite good for smaller BCIs that do not use EEG speech decoding as their main feature. But in general, EEG is not the best signal acquisition technique for standard speech decoding devices since its technology is currently limited to very small vocabulary sets as shown by the aforementioned studies. Additionally, these studies do not present data on how fast their decoder can decode the imagined speech which likely means that they are not developed enough for real-time speech decoding capabilities^{13,14,15}. Speech decoding BCIs using EEG are not as developed relative to other decoders using other signal acquisition techniques. However, the fact that EEG is a non-invasive signal acquisition and the fact that more advanced and efficient classifying algorithms and decoding models can be used, gives EEG a lot of potential in the future.

2.2. Invasive Signal Acquisition Techniques: ECoG & UIEA

Electrocorticography (ECoG), also known as intracranial electroencephalography (iEEG), is a commonly used invasive technique used for signal acquisition in BCIs. ECoG records electrical activity through a grid of flat electrode contacts placed on a thin silicone sheet on the cerebral cortex¹⁶. Because it is invasive, participants of studies involved in ECoG would have to undergo a craniotomy¹⁷. To overcome this, researchers do their studies on participants who are brain surgery patients and have already undergone a craniotomy¹⁷. Despite the difficulty in finding brain surgery patients to conduct ECoG research, ECoG offers many advantages that other signal acquisition techniques lack.

Unlike EEG, ECoG offers higher spatial resolution and the ability to capture more detailed neural signals¹⁸. Also, ECoG has excellent temporal resolution which is similar to the temporal resolution of EEG¹⁸. Additionally, a lot of the physiological noise from muscle contractions that is characteristic of EEG is eliminated while using ECoG since the electrodes are not attached to the skin¹⁸. ECoG also has a considerable bandwidth of frequencies of up to 200 Hz or higher which can be beneficial in some practical cases¹⁹. Although ECoG offers excellent signal quality, low noise levels, and high temporal and spatial resolutions, it requires researchers to be very flexible with their study. Specifically, researchers are limited to where they can place their ECoG electrode grid depending on where the patient has gotten their craniotomy. This is a major factor, especially in the field of speech decoding.

The UIEA (Utah intracortical electrode array) is another invasive signal acquisition technique that is used in BCIs. Functionally, it is almost identical to ECoG electrodes except that the UIEA has a different electrode design. It is made up of 100 silicon microelectrodes that are 1 mm long and spaced 0.4 mm apart²⁰. These microelectrodes penetrate the cerebral cortex rather than sitting on the surface like ECoG²¹. Due to the very small size of the penetrating

microelectrodes, they can record action potentials from individual neurons whereas ECoG can only record action potentials from a group of neurons^{21,22}. This allows for more precise data which can be useful for researchers. However, the small size of the array is also a disadvantage since it has very little coverage on the brain, which is less practical, causing researchers to use multiple arrays on their participants²³. The main drawback of UIEA is that it can cause acute cell death and chronic degeneration in the brain area where the array has been placed²⁴. This can negatively impact the quality of neural data being collected and can also cause unnecessary harm to the participant²⁵.

2.2.1 Invasive Studies

The latest, most successful speech decoding BCI developed using ECoG has been the high-performance neuroprosthesis for speech decoding and avatar control by Metzger et al²⁶. Their participant has severe limb and vocal paralysis and anarthria and is unable to produce vocal noises, but can marginally move her orofacial muscles when attempting to silently speak²⁶. Because the participant can move her orofacial muscles, Metzger et al. were able to create a multimodal speech neuroprosthesis that collects neural data from the ECoG electrodes placed in the supplementary motor complex speech and superior temporal gyrus which are cortical areas relevant for orofacial movements²⁶. They collected data from these areas since they contain the neural signals traveling to the motor neurons of the face, or signals that can be translated into speech²⁶. In text decoding, the participant was instructed to try and say 249 random sentences that were not used during model training, from a 1024-word-General set²⁶. Their decoding model was able to decode the speech at a rate of 78 WPM (words per minute) and a WER (word error rate) of 25.5% in real-time²⁶.

Another very successful and recent speech decoding BCI is the high-performance speech neuroprosthesis by Willett et al²³. Their participant can no longer speak comprehensibly due to amyotrophic lateral sclerosis but retains some limited orofacial movement and an ability to vocalize with sounds²³. The study used 4 intracortical electrode arrays, placed in the ventral premotor cortex and a part of Broca's area, instead of ECoG electrodes, to collect neural data at single-neuron resolution²³. Their decoder decoded at 62 WPM and achieved a 9.1% WER on a 50-word vocabulary set and a 23.8% WER on a 125,000-word vocabulary set in real-time²³.

Although numerous other factors differ between the two neuroprosthesis studies, such as preprocessing techniques and feature extraction strategies, it can be concluded that the electrode type and the regions the electrodes were placed in are two crucial factors that greatly affect the results of the studies^{23,26}. These major differences between the two studies are likely what caused Metzger et al.'s BCI to decode speech at a much faster rate than Willett et al.'s BCI^{23,26}. On the other hand, Willett et al.'s BCI was able to achieve lower word error rates with both of its vocabulary sets than Metzger et al.'s BCI^{23,26}. Additionally, Willett et al.'s BCI surpasses Metzger et al.'s BCI in terms of vocabulary set size^{23,26}. However, Metzger et al.'s BCI offers avatar control for emotional expressions while Willett et al.'s BCI does not^{23,26}. Both studies' BCIs have

considerable advantages over one another and have proved themselves to be breakthroughs in the current state of speech decoding technology^{23,26}.

Signal acquisition using invasive techniques, such as ECoG and the UIEA, offer greater qualities and quantities of data since they obtain the brain signals directly from the source which can greatly increase the speech decoding BCI's overall performance. This is exactly what happened in the two studies mentioned above. By collecting valuable and precise data to train their advanced decoding models, the studies were able to obtain impressive results relative to modern-day speech decoding technology. Overall, speech decoders relying on invasive signal acquisition techniques offer above-average accuracy, excellent vocabulary sizes, and fast real-time speech decoding, making it a reliable device for mute or disabled individuals.

2.3. EMG and Non-neural Signal Acquisition Techniques

EMG is similar to EEG in that it measures electrical activity. But instead of measuring electrical activity from neurons in the brain, it measures electrical activity in muscles and the motor neurons that control them²⁷. In other words, EMG measures the motor unit action potentials sent to the muscle from the brain when the muscles contract²⁷. EMG electrodes can be invasive and non-invasive. Invasive EMG electrodes include needle electrodes and fine-wire electrodes which are placed within the muscle of interest²⁸. Surface EMG (sEMG) electrodes are usually the preferred electrodes for studies since they are non-invasive and can simply be placed onto the skin of the relevant muscle²⁸. In the context of speech decoding, researchers often place sEMG electrodes on areas of the face and neck to collect electrical signals sent from the brain as the participant silently vocalizes the speech²⁹. sEMG electrodes are also similar to EEG in that they are feasible, inexpensive, and non-invasive. Additionally, they are highly susceptible to noise and artifacts which can lead to misinterpreted data and can hinder a speech decoder's efficiency³⁰. Like EEG, sEMG speech decoders would need to filter out the noise and amplify the relevant EMG signal. EMG also exhibits reduced spatial resolution which can cause separation and the accurate reconstruction of the bioelectric sources to become challenging and infeasible³¹.

2.3.1 EMG and Non-neural Signal Acquisition Technique-Based Studies

AlterEgo is a personalized wearable silent speech interface that allows users to silently converse with a computing device without any voice or any discernible movements of the face²⁹. It enables users to communicate with AI assistants, applications, or other people, providing a seamless form of intelligence augmentation and a form of communication²⁹. The device uses 7 sEMG electrodes placed in specific areas of the laryngeal, hyoid, buccal, mental, oral, and infraorbital regions of the face since these regions are muscle articulators involved in speech production²⁹. Since the goal of the device is to primarily act as an interface between AI assistants and humans, rather than a specialized speech decoder, the vocabulary set of the study of the device included numbers and arithmetic functions, and words correlated with issuing reminders and scheduling events, accessing times around the world, and basic communicative replies²⁹. Because the device is not specialized to be a speech decoder, the study did not record a discrete

number of vocabulary words in their vocabulary set²⁹. After training their decoding model with 3 participants, Kapur et al. tested the accuracy of their device with 10 participants who had no prior experience with the device²⁹. Although the device was said to be trained using the vocabulary sets aforementioned, the participants were only instructed to silently speak numerical digits, from which an overall WER of 7.99% had resulted²⁹. Instead of WPM, latency (the time it takes for the transcription of the silent speech to be produced) was tested and resulted in an average of 0.427 seconds²⁹.

Another notable speech decoding study using sEMG is a study by Meltzner et al³². They developed a speech decoding system capable of recognizing silently mouthed words and phrases entirely from the recorded sEMG signals³². The system used a total of 11 sEMG electrodes which were placed precisely on muscles of the face and neck involved in speech production³². Unlike AlterEgo, this system was specifically made to decode speech³². The researchers had the goal of finding the optimal number of sEMG electrodes to use and the optimal speech decoding algorithm to use for the system³². They learned that 8 sEMG electrodes were the optimal number of electrodes to use for decoding speech with sEMG³². Among the 6 participants for the experiment containing 1,200 phrases from a 2,200 vocabulary set, Meltzner et al. were able to achieve an average WER of 8.9%³². Results regarding the speed of decoding were not included in the study³².

A rather distinct non-neural speech decoding device is the SilentSpeller³³. SilentSpeller uses a unique signal acquisition technique that does not technically collect biosignals³³. Instead, it uses Complete Speech's SmartPalate, a dental retainer-type device with 124 binary capacitive touch sensors, which is placed at the roof of the mouth to capture tongue movements³³. Unlike other speech decoding devices and studies, SilentSpeller requires its users to silently mouth each letter in the context of spelling a word³³. While the user silently spells the word, the device decodes every letter from the positions of the tongue³³. These positions are obtained from when the user's tongue touches the capacitive sensors³³. This is quite different but also simpler than the other speech decoding devices or studies in this paper since it does not obtain or decode neural signals or images³³. The SilentSpeller's main advantage over other speech decoding devices and studies is its portability and ability to decode speech while the user is walking as well as seated³³. The two participants of the study had an average seated WER of 7.5% and a walking WER of 7%³³. The difference in walking WER between the two participants is 0.5% which is a considerable difference³³. This shows that there can be variability in accuracy for different users³³. Out of three sessions, the participants' average maximum session speed was 48 WPM³³. The vocabulary set consisted of 321 words although the decoding model was trained on 2328 words³³.

The results from these 3 studies show that speech decoding systems using sEMG and non-neural signal acquisition techniques are capable of high and impressive accuracies or low WERs^{29,32,33}. All 3 of these studies have WERs below 10% which is exceptionally good^{29,32,33}. Some drawbacks that these studies' systems have is that they vaguely report their speech decoders' speed of decoding in words per minute or latency since they focused more on accuracy

instead^{29,32,33}. An example of this is SilentSpeller since it achieved an average WPM of 48 which is relatively slow compared to other speech decoders in this paper³³. On the other hand, although AlterEgo did not report average WPM since it is not a standard speech decoding device, it did report its latency of 0.427 seconds which is exceptionally fast for its function²⁹. However, a drawback to AlterEgo is its very limited vocabulary since it is only able to decode a few very specific groups of words²⁹. Similarly, SilentSpeller is also reported to have a relatively small vocabulary set which makes it inferior to other speech decoding devices³³. In contrast, Meltzner et al.'s device shows great results due to its comparatively large vocabulary set of 2,200 words and its low WER³². Unfortunately, they did not report latency or WPM for their device, making its real-time, applicable decoding capabilities very unclear³². Overall, speech decoders relying on sEMG and non-neural signal acquisition techniques offer exceptional accuracy but mediocre applicability in real-time decoding for users due to their smaller vocabulary limits and slower decoding speeds.

2.4. Functional Magnetic Resonance Imaging (fMRI)

Functional magnetic resonance imaging (fMRI) is a non-invasive neuroimaging technique used to measure brain activity by detecting changes in blood flow. This technique is based on the principle that cerebral blood flow and neural activity are associated or that blood flow to an area of the brain will increase as that brain area is in use³⁴. Unlike the other signal acquisition techniques which are often electrodes, implants, or devices, the fMRI machine is a large circular tube-like machine in which the patient would have to be resting on a table and inserted into the machine. The machine uses a powerful magnetic field to detect changes in blood flow by utilizing the blood's slight magnetic properties³⁵. Since it is a neuroimaging technique, fMRI reconstructs a high-resolution image of the brain which provides information on the degree of activation of each point of the brain.

fMRI is a very widely used tool in the research community because it associates cognitive functions, tasks, and stimuli with cognitive anatomy and specific regions and points in our brain³⁶. In the context of speech decoding, fMRI can provide crucial information regarding how and where our brain organizes certain ideas, concepts, thoughts, or words³⁷. Previous studies generally aimed to create real-time, accurate, and efficient speech decoders for patients and customers to communicate with daily³⁷. Contrastingly, fMRI speech decoders aim to decode thoughts or speech through semantic reconstruction: a process that uses brain recordings to generate word sequences that recover the meaning of thoughts and imagined or silent speech³⁷. In other words, fMRI speech decoders decode speech through the meanings, intentions, and ideas that the participant is trying to convey by correlating these specific meanings to specific areas of the brain³⁷. Researchers accomplish this with the notion that certain meanings and thoughts will consistently reemerge in possibly multiple, specific regions of the brain which enables them to train machine learning models to decode speech³⁷. Dissimilar to all other speech decoders that use other signal acquisition techniques, fMRI speech decoders strive to achieve the overall meaning of the silently spoken or imagined speech³⁷. Because of this, fMRI's decoded speech

could include synonyms or slightly different words, and varying meanings, structures, and wording rather than outright random errors that can be found in other non-fMRI speech decoders³⁷.

On the contrary, since fMRI machines are extremely large, it is practically impossible for them to become portable, applicable, speech decoding devices for patients and customers, like devices using other signal acquisition techniques. However, researchers continue to research fMRI to understand the brain better, as well as in hopes of further developing fMRI speech decoding technology for applications other than communication. In addition to fMRI being impractical for portability, it offers a very slow temporal resolution of around 3 to 6 seconds, which makes it even more impractical for real-time speech decoding³⁸. fMRI's temporal resolution is very limited because of its dependence on changes in blood flow in regions of the brain, which is a very slow process compared to electrical processes from other signal acquisition techniques like EEG. However, fMRI is known to have an exceptional spatial resolution due to its unique properties³⁸. Additionally, fMRI does not involve radiation and is non-invasive, which makes it safe and accessible for more participants and can lead to more experimentation, data collection, and development³⁹.

2.4.1 fMRI Studies

One of the most successful and recent fMRI speech decoders is the BCI by Tang et al⁴⁰. With their fMRI scanner and the decoder that they developed, Tang et al. conducted various experiments to test their fMRI speech decoder across a range of tasks: decoding while listening to a story, decoding while imagining a spoken story, and language descriptions from brain responses to non-linguistic tasks⁴⁰. For the imagined speech decoding, participants were instructed to imagine telling five 1-minute stories while being recorded with fMRI⁴⁰. These same participants were also separately told the same stories outside of the scanner to provide reference transcripts and compare the decoded reference transcript with the imagined story transcript⁴⁰. For each 1-minute scan, the researchers correctly matched the story (with 100% identification accuracy) that the subject was imagining with the reference transcripts⁴⁰. They did this by using BERTScore, a method that uses machine learning to show whether the decoded speech and the reference transcripts share a meaning since semantic decoding can result in different vocabulary to be used but still convey the same meaning⁴⁰. Although the study did not list a quantitative value that shows the similarity between the imagined decoded speech and reference transcript, the BERTScore graphics showed that the decoder can successfully recover the meaning of imagined stimuli⁴⁰. Also, it is important to note that the decoding model that was used to decode this imagined speech was trained on perceived speech data instead of attempted or imagined speech data because the researchers wanted to focus on having their decoder decode speech that people are listening to⁴⁰. If the researchers had trained their decoder on imagined or attempted speech, they claim that they would have likely achieved even more precise imagined speech decoding⁴⁰.

Another influential fMRI study is by Huth et al. in which they decode various categorized objects and actions from movie scenes³⁷. Although this study doesn't decode imagined or attempted speech, it is still relevant because it decodes a combination of perceived speech and images which is known to produce similar results to stimuli like imagined speech, as shown in the previous study³⁷. In the experiment, participants were shown 2 hours of a movie while under fMRI scans from which 1-second scenes, containing a specific action or object, were labeled, categorized, and decoded³⁷. This study presents AUC values, a common statistic similar to accuracy (with a value of 1.0 representing perfect decoding), to show the results they achieved with their decoder³⁷. For the word "talk", the decoder achieved an AUC of 0.918, demonstrating that it is very accurate at decoding that certain word, scene, or action from the movie that participants were shown³⁷. The AUC for "animal" was 0.911; The AUC for "vehicle" was 0.758; And the AUC for "thing" was 0.694³⁷. This means that categories such as "talk" and "animal" are more distinct in the brain, making it more accurate for the decoder to decode³⁷. But vague categories or ideas such as "thing" can be very indistinct for fMRI to capture which makes it less accurate for the decoder to decode³⁷.

Studies involving fMRI and decoding stimuli have served to unveil fMRI's capabilities in various other applications similar to speech decoding, as well as reveal more information about the semantic systems of the brain. Although real-time, portable fMRI speech decoding BCIs may never exist because fMRI is not portable, it is still a valuable signal acquisition technique since it is non-invasive and presents a high accuracy. In conclusion, fMRI is a unique signal acquisition technique due to its ability to decode by semantics which gives it great potential for futuristic applications and practical studies and experiments.

3. Discussion

The factors that make up the ideal speech decoding device for customers and patients include *portability, non-invasiveness, high accuracy, large vocabulary, and fast real-time speech decoding speed*. All signal acquisition techniques comply with most of these factors to a certain extent, but a signal acquisition technique that exceptionally meets all of these criteria does not yet exist. Instead, each signal acquisition technique has its advantages and disadvantages when using them to develop a speech decoding device. Most of the disadvantages some of these signal acquisition techniques have can be overcome by further development within the speech decoders using these signal acquisition techniques.

3.1 Review of the Signal Acquisition Techniques with Regard to the Factors

The factors of portability and non-invasiveness are generally positively correlated in judging signal acquisition techniques. For example, both EEG and EMG are completely non-invasive signal acquisition techniques, and speech decoding devices that use these techniques are often made to be very portable. Conversely, speech decoders that use invasive signal acquisition techniques are not portable due to an unexpected reason. Although the implants themselves are small and portable, the server stacks and other processing hardware used

in the BCI to decode speech are very large which makes the speech decoder not viably portable²⁶. Since these invasive speech decoders are the best-performing due to their exceptionally large vocabulary set sizes and accuracies, these large parts of the speech decoder should be developed in the future to become much smaller and portable to turn the BCIs into superior speech decoding devices for patients. fMRI on the other hand, is an exception to the positive correlation between portability and non-invasiveness. Although fMRI is a non-invasive technique, it cannot be used for portable speech decoding devices since it is a very large machine and simply can not be viably integrated within a portable BCI or transported easily.

The factors of accuracy and vocabulary set size are generally negatively correlated with these signal acquisition techniques. A majority of the studies using EEG, EMG, and fMRI all had high accuracy in their speech decoding but had a relatively small number of words in their vocabulary sets. This relationship is logical because it is easier for a speech decoding algorithm to correctly identify a word if the vocabulary set is very small compared to identifying a word from a very large vocabulary set. Moreover, the invasive studies had very large vocabulary sets relative to the studies using other signal acquisition techniques which resulted in these invasive studies having lower accuracy than the other studies. Despite the lower accuracies, these invasive studies' speech decoders are the best out of all the studies in this paper since they still achieved exceptional accuracy compared to studies with similar vocabulary sizes.

Regarding the factor of real-time speech decoding speed, the invasive studies are, again, the best out of all the other studies. Invasive speech decoders have exceptional speed in decoding speech in real-time. The only main disadvantages of decoders using invasive signal acquisition techniques are portability and invasiveness. Since it is invasive, these speech decoders are limited for use only to brain surgery patients and individuals who are disabled and willing to get surgery for such a device¹⁷. For people in the general consumer market interested in speech decoding BCIs, sEMG, EEG, and other non-invasive signal acquisition speech decoders would be the best option since they provide high accuracy, decent real-time speech decoding speed, and contain a moderate to low vocabulary set size. fMRI decoders are quite different from the other signal acquisition techniques because they are currently incapable of decoding speech in real-time. Instead, fMRI provides important semantic information and a functional understanding of the brain. It has great potential for nonportable speech decoding and other futuristic applications such as idea or concept sharing

4. Conclusion

The field of speech decoding has made great progress in developing efficient and reliable speech decoding BCIs using various signal acquisition techniques and other technologies. Each group of BCIs using a certain signal acquisition technique has its advantages and disadvantages. In the future, speech decoding BCIs should aim to combine the advantageous features of these BCIs and eliminate the disadvantageous features by using a signal acquisition technique that is portable and non-invasive and has shown to be capable of high accuracy, vocabulary, and fast real-time speech decoding speed.

5. Methods

Since this paper is a literature review, the methods were centered around quantitative and qualitative research of studies relevant to signal acquisition techniques discussed in this paper. Once the prominent signal acquisition techniques to be included in this paper were decided, the databases Google Scholar and PubMed were used to identify the relevant studies to analyze for each signal acquisition technique. Keywords used to identify the relevant studies include the signal acquisition technique names as well as “speech decoding” or “speech decoder.” The most cited papers were displayed at the top of the list within these databases. These papers were preferred to include within this review because they are likely more representative of each of their respective signal acquisition techniques. Research papers that were included in this review were found on the first page of the databases and were made sure to have been cited highly relative to other papers on the page. Additionally, specific quantitative and qualitative factors were looked at within these relevant studies to confirm whether these studies would be a great fit to analyze and include in this paper. For qualitative factors, the studies were analyzed to identify whether the BCIs were tested on participants, whether they decoded speech, and whether they collected brain data using a particular signal acquisition technique. For quantitative factors, the studies were confirmed on whether they reported data on vocabulary size and speech decoding accuracy and speed. After 2 to 3 relevant studies were chosen to represent each signal acquisition technique, the qualitative and quantitative data were compared between the studies to create an overview of the advantages and disadvantages of using each signal acquisition technique for speech decoding.

6. Acknowledgements

I would like to thank the Lumiere team, especially my mentor, Stephanie Brener, for supporting, guiding, and helping me through the research and writing process.

Works Cited

- J.J. Vidal, Toward Direct Brain-Computer Communication. *Annual Review of Biophysics and Bioengineering*. **2**, 157-180 (1973).
- A. Kawala-Sterniuk, N. Browarska, A. Al-Bakri, M. Pelc, J. Zygarlicki, M. Sidikova, R. Martinek, E.J. Gorzelanczyk, Summary of over Fifty Years with Brain-Computer Interfaces—A Review. *Brain Sciences*. **11**, 43 (2021).
- M. Jahanaray, A. Jahanaray, Z. Zohoorian, Brain regions involved in speech production, mechanism and development. *Neuroscience Research Notes*. **5**, 178 (2022).
- C.J. Price, J.T. Crinion, M. MacSweeney, A generative model of speech production in Broca's and Wernicke's areas. *Frontiers in Psychology*. **2**, (2011).
- A. Horne, R. Zahn, O.I. Najera, R.C. Martin, Semantic Working Memory Predicts Sentence Comprehension Performance: A Case Series Approach. *Frontiers in Psychology*. **13**, (2022).
- C. Rogalsky, W. Matchin, G. Hickok, Broca's area, sentence comprehension, and working memory: an fMRI study. *Frontiers in Human Neuroscience*. **2**, (2008).
- Z.K. Agnew, C. McGettigan, B. Banks, S.K. Scott, Articulatory movements modulate auditory responses to speech. *NeuroImage*. **73**, 191-199 (2013).
- B. Gick, I. Wilson, D. Derrick, *Articulatory Phonetics*. 6 (2013).
- P.L. Nunez, R. Srinivasan, *Electric fields of the brain: the neurophysics of EEG*. 3-610 (2006).
- B. Burle, L. Spieser, C. Roger, L. Casini, T. Hasbroucq, F. Vidala, Spatial and temporal resolutions of EEG: Is it really black and white? A scalp current density view. *International Journal of Psychophysiology*. **97**, 210-220 (2015).
- P. Khatwani, A. Tiwari, A survey on different noise removal techniques of EEG signals. *International Journal of Advanced Research in Computer and Communication Engineering*. **2**, (2013).
- P. Fiedler, J. Haueisen, A.M.C. Alvarez, G. Cheron, P. Cuesta, F. Maestú, M. Funke, Noise characteristics in spaceflight multichannel EEG. *PLOS ONE*. **18**, (2023).
- J.T. Panachakel, R.A. Ganesan, Decoding Imagined Speech From EEG Using Transfer Learning. *Institute of Electrical and Electronics Engineers*. **9**, 135371-135383 (2021).
- J.S. García-Salinas, L. Villaseñor-Pineda, C.A. Reyes-García, A.A. Torres-García, Transfer learning in imagined speech EEG-based BCIs. *Biomedical Signal Processing and Control*. **50**, 151-157 (2019).
- L. Tøttrup, K. Leerskov, J. T. Hadsund, E. N. Kamavuako, R. L. Kæseler, M. Jochumsen, Decoding covert speech for intuitive control of brain-computer interfaces based on single-trial EEG: a feasibility study. *Institute of Electrical and Electronics Engineers*. 689-693 (2019).
- F. Fallegger, G. Schiavone, E. Pirondini, F.B. Wagner, N. Vachicouras, L. Serex, G. Zegarek, A. May, P. Constanthin, M. Palma, M. Khoshnevis, D.V. Roost, B. Yvert, G. Courtine, K. Schaller, J. Bloch, S.P. Lacour, MRI-Compatible and Conformal Electro-corticography Grids for Translational Research. *Advanced Science*. **8**, (2021).

- W. Chiong, M.K. Leonard, E.F. Chang, Neurosurgical Patients as Human Research Subjects: Ethical Considerations in Intracranial Electrophysiology Research. *Neurosurgery*. **83**, 29-37 (2017).
- J. Parvizi, S. Kastner, Human Intracranial EEG: Promises and Limitations. *Nature Neuroscience*. **21**, 474-483 (2018).
- N.E. Crone, A. Sinai, A. Korzeniewska, High-frequency gamma oscillations and human brain mapping with electrocorticography. *Progress in Brain Research*. **159**, 275-295 (2006).
- D.D. Zhou, R.J. Greenberg, Microsensors and microbiosensors for retinal implants. *Frontiers in Bioscience-Landmark*. **10**, 166-179 (2005).
- P.J. Rousche, R.A. Normann, Chronic recording capability of the Utah Intracortical Electrode Array in cat sensory cortex. *Journal of Neuroscience Methods*. **82**, 1-15 (1998).
- A. Dubey, S. Ray, Cortical Electrocorticogram (ECoG) Is a Local Signal. *The Journal of Neuroscience*. **39**, 4299-4311 (2019).
- F.R. Willett, E.M. Kunz, C. Fan, D.T. Avansino, G.H. Wilson, E.Y. Choi, F. Kamdar, M.F. Glasser, L.R. Hochberg, S. Druckmann, K.V. Shenoy, J.M. Henderson, A high-performance speech neuroprosthesis. *Nature*. **620**, 1031-1036 (2023).
- P.A. Cody, J.R. Eles, C.F. Lagenaur, T.D.Y. Kozai, X.T. Cuia, Unique electrophysiological and impedance signatures between encapsulation types: An analysis of biological Utah array failure and benefit of a biomimetic coating in a rat model. *Biomaterials*. **161**, 117-128 (2018).
- P.R. Patel, E.J. Welle, J.G. Letner, H. Shen, A.J. Bullard, C.M. Caldwell, A. Vega-Medina, J.M. Richie, H.E. Thayer, P.G. Patil, D. Cai, C.A. Chestek, Utah array characterization and histological analysis of a multi-year implant in non-human primate motor and sensory cortices. *Journal of Neural Engineering*. **20**, 1741-2552 (2023).
- S.L. Metzger, K.T. Littlejohn, A.B. Silva, D.A. Moses, M.P. Seaton, R. Wang, M.E. Dougherty, J.R. Liu, P. Wu, M.A. Berger, I. Zhuravleva, A. Tu-Chan, K. Ganguly, G.K. Anumanchipalli, E.F. Chang, A high-performance neuroprosthesis for speech decoding and avatar control. *Nature*. **620**, 1037-1046 (2023).
- A. Michell, *Understanding EMG*. OUP Oxford. (2013).
- V. Gohel, N. Mehendale, Review on electromyography signal acquisition and processing. *Biophysical Reviews*. **12**, 1361-1367 (2020).
- A. Kapur, S. Kapur, P. Maes, AlterEgo: A Personalized Wearable Silent Speech Interface. *IUI '18: 23rd International Conference on Intelligent User Interfaces*. 43-53 (2018).
- M. Boyer, L. Bouyer, J. Roy, A. Campeau-Lecours, Reducing Noise, Artifacts and Interference in Single-Channel EMG Signals: A Review. *Sensors*. **23**, (2023).
- T. Klotz, L. Gizzi, O. Röhrle, Investigating the spatial resolution of EMG and MMG based on a systemic multi-scale model. *Biomechanics and Modeling in Mechanobiology*. **21**, 983-997 (2022).

- G.S. Meltzner, J.T. Heaton, Y. Deng, G.D. Luca, S.H. Roy, J.C. Kline, Development of sEMG sensors and algorithms for silent speech recognition. *Journal of Neural Engineering*. **15**, 1741-2552 (2018).
- N. Kimura, T. Gemicioglu, J. Womack, R. Li, Y. Zhao, A. Bedri, A. Olwal, J. Rekimoto, T. Starner, Mobile, Hands-free, Silent Speech Texting Using SilentSpeller. *CHI EA '21: Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1-5 (2021).
- W.D. Haselden, R.T. Kedarasetti, P.J. Drew, Spatial and temporal patterns of nitric oxide diffusion and degradation drive emergent cerebrovascular dynamics. *PLOS Computational Biology*. **16**, (2020).
- L. Pauling, C.D. Coryell, The Magnetic Properties and Structure of Hemoglobin, Oxyhemoglobin and Carbonmonoxyhemoglobin. *Proceedings of the National Academy of Sciences of the United States of America*. **22**, 210-216 (1936).
- N.K. Logothetis, What we can do and what we cannot do with fMRI. *Nature*. **453**, 869-878 (2008).
- A.G. Huth, T. Lee, S. Nishimoto, N.Y. Bilenko, A.T. Vu, J.L. Gallant, Decoding the Semantic Content of Natural Movies from Human Brain Activity. *Frontiers in Systems Neuroscience*. **10** (2016).
- G.H. Glover, Overview of Functional Magnetic Resonance Imaging. *Neurosurgery Clinics of North America*. **22**, 133-139 (2011).
- C. Longworth, G. Honey, T. Sharma, Functional magnetic resonance imaging in neuropsychiatry. *British Medical Journal*. **319**, 1551-1554 (1999).
- J. Tang, A. LeBel, S. Jain, A.G. Huth, Semantic reconstruction of continuous language from non-invasive brain recordings. *Nature Neuroscience*. **26**, 858-866 (2023).

The Application of CRISPR-cas9 Gene Editing Technology for the Synthesization of Hypoallergenic Food Items By Rahul Jagadish Nadig

Abstract

The effects of allergic reactions aren't just detrimental anatomically and physiologically, but also have various effects on one's social life and interactions. Hence, there is a great need or urgency to resolve this matter. Immunotherapy, which has been the current/on-going source of treatment, isn't a standardized form of treatment due to the variability of outcomes regarding an individual's reaction when exposed to the particular allergen. Hence, alternate sources of treatment are required, whereby this article explores CRISPR-cas9 gene-editing technology. This particular technology is specifically used to alter the genetic-code/genome of the crop-producing the allergen, whereby a specific gene coding for a specific protein is targeting, inducing various mutations. This renders the gene "faulty", causing limited production of the protein, thus an inhibition in the severity/scale of the allergic reaction. This technology has been applied for common food allergies including peanut, egg (specifically from *Gallus domesticus*), and gluten allergy.

Key words CRISPR-cas9; food allergies; hypoallergenic; allergen;

Introduction

Food allergies—an example of an inappropriate immune reaction—is a condition that is caused when a certain food item is consumed, containing a particular antigen that the body develops a hypersensitive reaction to. This condition has become exponentially prevalent over the last few decades, with statistics revealing increased food allergy prevalence by 50% amongst children between 1997 to 2011 and another 50% between 2007 to 2021 [1]. In addition to the increased prevalence, the impacts of food allergies, both anatomically and socially, are devastating. In fact, according to a research study where 87 families were subjected, 49% or more have indicated that food allergies affected family social activities, and 34% have stated the impact of food allergies on their child's attendance in school, with a shocking 10% choosing to homeschool their kids due to the food allergies [2]. Some of the symptoms of food allergies include rashes, hives, stomach pain/cramp, and in some cases—where the magnitude of the allergic reaction is at the highest—individuals may go into anaphylactic shock, where the blood pressure decreases, and when there is narrowing of the throat/airways, making it difficult to breathe.

In order to treat an individual who's in an anaphylactic shock, one must administer an "EPIPEN" or an epinephrine injection. Individuals whose weight is above 30 kilograms shall be administered with the regular EPIPEN with a dosage of 0.3 mg epinephrine, whereas individuals within the weight range of 15 to 30 kilograms shall be administered with the EPIPEN JR with a dosage of 0.15 mg epinephrine [3]. In order to alleviate mild allergic reactions, individuals are administered with antihistamines. During any allergic reaction, Histamine (a chemical

messenger), is produced. This chemical messenger increases vascular permeability, meaning there is an increased flow of fluid from capillaries into the surrounding tissue [4]. This leads to increased swelling and dilation of the blood vessels, thus lowering the blood pressure [4]. Antihistamines, specifically H1 antihistamines which are used to specifically treat allergies and allergic rhinitis will act as antagonists and bind to the H1 receptors, thereby preventing swelling and dilation of blood vessels [4].

The EPIPEN and antihistamine, while extremely beneficial products can only provide short-term treatment. Individuals will still continue to have allergies and will develop allergic reactions if they consume a particular food-allergen/s. Research investigating long-term treatment for food allergies has mainly looked into immunotherapy. Food-allergen immunotherapy exposes an individual to a particular allergen, gradually increasing the dosage of the allergen to eventually prevent the hypersensitive reaction, resulting in desensitization and tolerance [5]. There are three main types of immunotherapy that are being studied, specifically pertaining to allergies: oral, sublingual, and epicutaneous immunotherapy. The efficacy of immunotherapy is debatable. According to De Silva et al, while immunotherapy has shown promise, it does present its own risks.

Outcome	% active	% control	Absolute difference	Relative risk (95% CI)	NNT	Certainty of evidence	Trials and participants
Desensitization	68	15	53%, $p < .05$	5.7 (1.9–16.7)	2	Moderate	7 ($n = 249$)
Adverse reactions	85	17	67%, $p < .05$	3.9 (2.1–7.5)	2	Low	6 ($n = 210$)
Severe adverse reactions	3	0	3%, $p > .05$	7.0 (0.4–124.8)	NA	Very low	4 ($n = 204$)

Figure 1. (De Silva et al, 2022) Depicts a significant increase in desensitization towards the allergen, specifically cow’s milk, with a relatively decent certainty of evidence (moderate) as a result of Oral immunotherapy.

Outcome	% active	% control	Absolute difference	Relative risk (95% CI)	NNT	Certainty of evidence	Trials and participants
Desensitization	84	5	79%, $p < .05$	8.9 (4.4–18.0)	2	Moderate	6 ($n = 259$)
Sustained unresponsiveness	35	4	30%, $p < .05$	7.1 (1.7–29.4)	4	Low	2 ($n = 91$)
Adverse reactions	79	8	71%, $p < .05$	7.0 (2.4–19.8)	2	Moderate	6 ($n = 291$)
Severe adverse reactions	6	0	6%, $p > .05$	3.4 (0.6–19.6)	NA	Very low	4 ($n = 211$)

Figure 2. (De Silva et al, 2022) Results from the Hen’s egg desensitization experiment

However, Figure 2. while also depicts an increase in desensitization towards the allergen (Hen’s egg) as a result of Oral immunotherapy, it shows a greater increase in adverse reactions with the certainty of evidence being relatively high (moderate), supported by great statistical significance with the p-value being less than 0.05. Greater evidence to support the outcome of adverse reactions implies potentially dangerous outcomes for the patient, lowering the efficacy of the treatment. This highlights the lack of standardization of immunotherapy and high variability of patients’ reactions, making it ineffective as long-treatment for food-allergies as a whole.

Mechanism of CRISPR-Cas9 Gene-editing tool

CRISPR (clustered regularly interspaced short palindromic repeats) cas systems are acquired immune systems that are present in different microorganisms like archaea and bacteria and have been investigated as potential long-term solutions for genetic diseases/disorders. The first hint of their existence was in 1987, where an unusual repetition of sequences was observed in an Escherichia coli genome [7]. This was subsequently termed as “CRISPR” [7]. A crucial step to discovering the applications of CRISPR-cas technology was discovering the link between CRISPRs and certain CAS-proteins, which was initially hypothesized to be responsible for DNA repair in hyperthermophilic archaea [7].

The mechanism in which CRISPR-cas9 gene editing takes place is through 3 main steps: recognition, cleavage and repair. The designated sgRNA (single guide RNA) will direct the cas9 protein to the specific gene of interest that needs to be modified [8]. This is done through the 5’ crRNA (crispr RNA) complementary base pair component [8]. The cas-9 protein remains inactive in the absence of sgRNA [8]. The cas9 nuclease produces DSBs (double-stranded breaks) at a site located near to the PAM sequence [8]. PAM sequence is a very short DNA sequence, consisting of 2-5 base-pairs approximately with its size varying based on the species

of bacteria [8]. Following which, the cas-9 recognized the PAM sequence at 5'-NCG-3', with "N" representing any nucleotide [8]. Once cas-9 has found the target site with the appropriate PAM sequence, it triggers local DNA melting, leading to the formation of the DNA-RNA hybrid [8]. Then, the cas-9 protein is then used for DNA cleavage whereby the HMN domain cleaves the complementary strand, and the RuvC domain cleaves the non-complementary strand of the target DNA to form "blunt-ended" DSBS [8]. This is then repaired by the host cellular machinery [8]. NHEJ (Non-homologous end joining) and HDR (Homology-directed repair) are two mechanisms to repair the DSBs synthesized by the CRISPR-cas9 mechanism. NHEJ is responsible for the repair of DSBs by joining DNA fragments and occurs in all phases of the cell-cycle, but is prone to errors that may result in frameshift mutations via small random insertions or deletions [8]. HDR, while only active in the late S and G2 phases of the cell-cycle, executes the gene insertion or replacement with very high precision [8].

However, to subject this technology to the entire human genome is very risky and could potentially trigger unwanted and harmful changes towards various genes. Hence, this particular technology is applied to the plant, whereby the gene that codes for the protein/allergen, is rendered faulty, thus inhibiting the production of that specific protein. In other words, this technology can be applied to synthesize hypo-allergenic food products, designed to mitigate and prevent the emergence and the effects of an allergic reaction. Various scientific experiments had applied this technology for some of the most frequent allergies including allergies to gluten, peanuts and eggs, specifically from the organism domestic chicken or "*Gallus domesticus*".

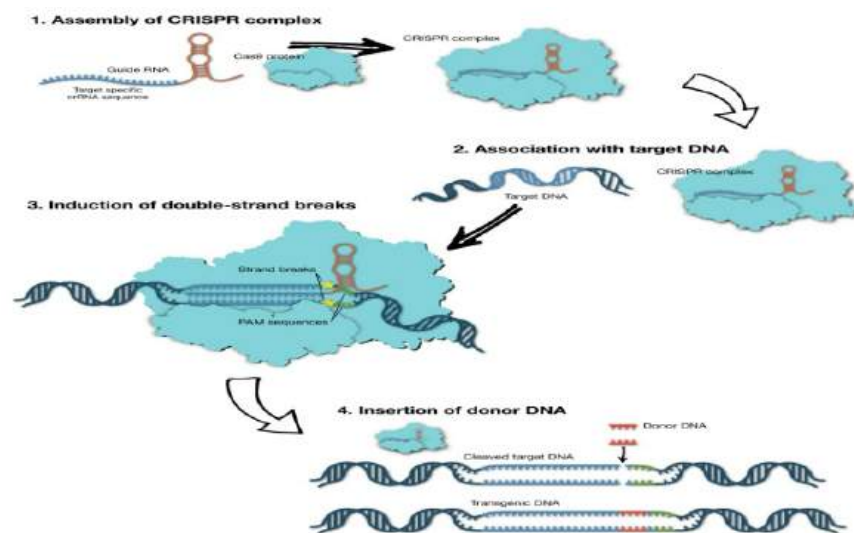


Figure 3. (National Center for Biotechnology Information, 2021) Depiction of the mechanism of CRISPR-cas9 gene editing technology

Exploring the wide range of applications, scientists are definitely hoping to use this technology for a variety of purposes, with the primary focus to permanently treat genetic diseases and disorder and also be used for other areas of interests, like the extermination of pathogenic/infectious diseases and maybe develop and mass-produce more nutritious crops and

vegetation. CRISPR-cas9 gene editing has already had some promise for curing certain genetic diseases like sickle cell anemia, cystic fibrosis and muscular dystrophy [8]. Generally speaking, the heritability of food allergies is estimated to be as high as 81% [10]. With a strong causal link between genetics and food allergies, it's definitely worth looking into the potential efficacy of CRISPR-cas9 gene editing as a long-term/permanent source of treatment for food allergies.

However, to subject this technology to the entire human genome is very risky and could potentially trigger unwanted and harmful changes towards various genes. Hence, this particular technology is applied to the plant/crop producing the antigen, whereby the gene that codes for the protein/allergen is rendered faulty, thus inhibiting the production of that specific protein. In other words, this technology can be applied to synthesize hypo-allergenic food items, enabling the individual to safely consume the food-item/allergen.

CRISPR-Cas9 Gene-editing tool for Wheat

Wheat is a food-crop and is a common ingredient, present in large quantities in various, frequently consumed food-products like bread. Gluten is the protein-component that is present in wheat and is the common allergen in which individuals develop an allergic reaction to. Many studies/experiments have confirmed that the gluten's main antigen "Gliadin", is responsible for triggering allergic reactions [11]. This has been attributed due to its toxicity, having agglutinating activity, inducing apoptosis(programmed cell-death), inhibiting cell-growth etc [11]. Sanchez-leon et al had performed an experiment, where they had used CRISPR-cas9 gene-editing tool, specifically constructed two sgRNAs to target the main immunodominant peptide (33-mer), responsible for triggering drastic effects for individuals with coeliac disease. Within this experiment, twenty-one mutated lines were generated, displaying a significant decrease in the production of alpha-gliadins [12]. Up to 35 genes were mutated amongst the 45 genes present within the wild-type in one of the mutated lines [12]. In addition to that, gluten content/immunoreactivity towards gluten had reduced by 85% [12]. This experiment was able to synthesize transgenic wheat-crops, containing low gluten-content, successfully mitigating the magnitude/scale of the allergic reaction, and potentially could prevent the emergence of an allergic reaction.

CRISPR-Cas9 Gene-editing tool for Peanuts

Peanut allergies are one of the most common and prevalent food allergies, with nearly 2.5% of children within the US may have an allergy to peanuts [13]. Peanuts contain multiple proteins, and therefore multiple allergens including the glycoprotein Ara-h2, one of the six main allergens responsible for triggering an allergic reaction towards peanuts [14]. An experiment conducted by Dodo et al, attempted to eliminate and "knockout" the immunodominant glycoprotein Ara-h2 by using a technique called RNA interference. A total of 59 "kanamycin-resistant" plants were regenerated with the phenotype. PCR(polymerase chain reaction) and "Southern Analysis" revealed that 44% of the plants contained the transgene [15]. In addition to that, there was a significant reduction in Ara-h2 content as evident with the p-value

being less than 0.05 [15]. Most importantly, the data revealed a significant decrease in the IgE binding capacity of the transgenic seeds, successfully demonstrating a mitigation and alleviation of peanut allergy and its effects.

CRISPR-Cas9 Gene-editing tool for Egg-white protein in chickens (*Gallus domesticus*)

Another very common allergy is egg allergy. The allergenic proteins in eggs from the organism domestic chickens or rather “*Gallus domesticus*” are majorly formed in egg-whites, with common examples including ovalbumin and ovomucoid [16]. In an experiment conducted by Oishi et al, Both of these genes were targeted using CRISPR gene-editing. This was done through the usage of circular plasmids encoding cas9 protein and an sgRNA. Both ovalbumin and ovomucoid had undergone mutations to a significant extent, specifically mutagenized by more than 90% within cultured chicken primordial germ cells (PCGs) [17]. By applying the concepts of Mendelian genetics, they were then able to produce ovomucoid homozygous mutant offspring in the 2nd generation by crossing heterozygous mutant chickens from the first generation, thus successfully creating and producing hypoallergenic eggs [17]. This experiment had greatly considered the ethical implications in using animals as test-subjects for their experiments by strictly adhering to the protocols approved by the institutional animal care and used committees of both *AIST* (National institute of Advanced industrial science and technology) and *NILGS* (NARO institute of livestock and greenland sciences) [17].

Conclusion

In conclusion, by using the results from the experiments conducted by Sanchez-leon et al, Dodo et al, and Oishi et al, we’re able to find and utilize an alternate method for the long-treatment of allergies through the application of the revolutionary technology of CRISPR-cas9 gene-editing tool, whereby the production of the protein/allergen is inhibited or minimized, thus successfully mitigating the emergence and the effects of an allergic reaction. This enables us to deviate from the accepted long treatment of immunotherapy whose benefits aren’t standardized or applicable to the entire general public due to the variability of an individual’s reaction when exposed to the particular allergen.

Works Cited

- [1] Food Allergy Research & Education. "Facts and Statistics." Food Allergy Research & Education, <https://www.foodallergy.org/resources/facts-and-statistics>. Accessed 25 Sept. 2024.
- [2] Broome, Susan Brantlee, et al. "Becoming the Parent of a Child With Life-Threatening Food Allergies." *Journal of Pediatric Nursing*, vol. 30, no. 4, Oct. 2014, pp. 532–42. <https://doi.org/10.1016/j.pedn.2014.10.012>. Accessed 25 Sept. 2024.
- [3] EPIPEN® (Epinephrine Injection, USP) Auto-Injectors| Dosage and Administration. www.epipen.com/hcp/about-epipen-and-generic/dosage-and-administration. Accessed 25 Sept. 2024.
- [4] Farzam, Khashayar, et al. "Antihistamines." StatPearls - NCBI Bookshelf, 10 July 2023, www.ncbi.nlm.nih.gov/books/NBK538188/#:~:text=Mechanism%20of%20Action,-Histamine%20. Accessed 25 Sept. 2024.
- [5] "New Approaches to Food Allergy Immunotherapy." *The Journal of Allergy and Clinical Immunology in Practice*, vol. 12, no. 3, Oct. 2023, pp. 546–52. <https://doi.org/10.1016/j.jaip.2023.10.018>. Accessed 26 Sept. 2024.
- [6] de Silva D, Rodríguez del Río P, de Jong NW, et al; the GA2LEN Food Allergy Guidelines Group. Allergen immunotherapy and/or biologicals for IgE-mediated food allergy: A systematic review and meta-analysis. *Allergy*. 2022; 77: 1852–1862. doi:10.1111/all.15211. Accessed 26 Sept. 2024.
- [7] Ishino, Yoshizumi, et al. "History of CRISPR-Cas From Encounter With a Mysterious Repeated Sequence to Genome Editing Technology." *Journal of Bacteriology*, vol. 200, no. 7, Jan. 2018, <https://doi.org/10.1128/jb.00580-17>. Accessed 26 Sept. 2024.
- [8] Mengstie, Misganaw Asmamaw, and Belay Zawdie Wondimu. "Mechanism and Applications of CRISPR/Cas-9-Mediated Genome Editing." *Biologics*, vol. Volume 15, Aug. 2021, pp. 353–61. <https://doi.org/10.2147/btt.s326422>. Accessed 27 Sept. 2024.
- [9] Costa, Joana R., et al. Figure 2: [Mechanism of CRISPR/Cas9 Genome Editing...]. - Assay Guidance Manual - NCBI Bookshelf. www.ncbi.nlm.nih.gov/books/NBK464635/figure/gen_edit.F2/#_ncbi_dlg_citbx_NBK464635. Accessed 27 Sept. 2024.
- [10] Johansson, Elisabet, and Tesfaye B. Mersha. "Genetics of Food Allergy." *Immunology and Allergy Clinics of North America*, vol. 41, no. 2, Mar. 2021, pp. 301–19. <https://doi.org/10.1016/j.iac.2021.01.010>. Accessed 28 Sept. 2024.
- [11] ---. "Divergence of Gut Permeability and Mucosal Immune Gene Expression in Two Gluten-associated Conditions: Celiac Disease and Gluten Sensitivity." *BMC Medicine*, vol. 9, no. 1, Mar. 2011, <https://doi.org/10.1186/1741-7015-9-23>. Accessed 28 Sept. 2024.
- [12] Sánchez-León, Susana, et al. "Low-gluten, Nontransgenic Wheat Engineered With CRISPR/Cas9." *Plant Biotechnology Journal*, vol. 16, no. 4, Sept. 2017, pp. 902–10. <https://doi.org/10.1111/pbi.12837>. Accessed 28 Sept. 2024.

- [13] “Peanut Allergy | Causes, Symptoms and Treatment | ACAAI Public Website.” ACAAI Patient, 1 June 2022, acaai.org/allergies/allergic-conditions/food/peanut. Accessed 28 Sept. 2024.
- [14] De Jong, None, et al. “Identification and Partial Characterization of Multiple Major Allergens in Peanut Proteins.” *Clinical & Experimental Allergy*, vol. 28, no. 6, June 1998, pp. 743–51. <https://doi.org/10.1046/j.1365-2222.1998.00301.x>. Accessed 30 Sept. 2024.
- [15] Dodo, Hortense W., et al. “Alleviating Peanut Allergy Using Genetic Engineering: The Silencing of the Immunodominant Allergen Ara H 2 Leads to Its Significant Reduction and a Decrease in Peanut Allergenicity.” *Plant Biotechnology Journal*, vol. 6, no. 2, Sept. 2007, pp. 135–45. <https://doi.org/10.1111/j.1467-7652.2007.00292.x>. Accessed 30 Sept. 2024.
- [16] Caubet, Jean-Christoph, and Julie Wang. “Current Understanding of Egg Allergy.” *Pediatric Clinics of North America*, vol. 58, no. 2, Mar. 2011, pp. 427–43. <https://doi.org/10.1016/j.pcl.2011.02.014>. Accessed 1 Oct. 2024.
- [17] Oishi, I., Yoshii, K., Miyahara, D. et al. Targeted mutagenesis in chicken using CRISPR/Cas9 system. *Sci Rep* 6, 23980 (2016). <https://doi.org/10.1038/srep23980>. Accessed 1 Oct. 2024.

Oral Immunotherapy (OIT): Evaluating An Effective But Unknown Treatment

By Anya Jolly

Abstract

This paper evaluates the effectiveness of Oral Immunotherapy (OIT) as a treatment for food allergies and the impact the treatment has on the quality of life of patients under the age of 18. OIT is a relatively new treatment for food allergies, and it is not a well known allergy treatment for food allergens. This paper evaluates the responses of a survey that was completed by 202 patients or the patients' parents. To have the survey reach a wide range of patients, the form was dispersed through social media such as Facebook groups. Overall, the respondents felt that OIT is a very effective treatment option and significantly improved the patient's quality of life. Specifically, 88.9% of participants indicated that the patient's quality of life improved after completing OIT, and 78.7% of the respondents indicated that the severity of their allergic reactions has decreased since beginning OIT. In addition, 97.5% of respondents were satisfied with their experience with OIT. A definitive 96.0% of patients state that they would strongly recommend OIT to other children, yet 81.7% of the respondents believe that OIT is not a widely known treatment option. This report concludes that OIT can transform the lives of people struggling with food allergies.

Introduction

This report helps to explain the numerous aspects of Oral Immunotherapy (OIT) as a treatment for food allergies and discusses the effectiveness of this treatment. As of June 2024, thirty-three million people in the United States struggle with food allergies (Food, 2024). In addition, there are food-allergy related cases in emergency rooms every ten seconds in the United States, classifying food allergies as a national "epidemic" (Food, 2024).

According to John Hopkins Medicine, allergic disease is one of the most prominent chronic health conditions worldwide (Allergies, 2024). The Center for Disease Control (CDC) reports food allergies affect about 6% of adults and 8% of children in the United States, which is about 1 in 13 children. The CDC also reported a 50% increase in food allergies in children between 1997 and 2011.

Atopic individuals are individuals that have the genetic tendency to develop allergic diseases (What, 2024). When atopic individuals are first exposed to allergens, they get immune reactions that lead to "allergic inflammation" (What, 2024). An allergy can derive from a wide variety of allergens, including pollen, food, dust, medicines, animal dander, insects, and mold. Furthermore, an allergy can develop at any time in a person's life regardless of their age. A food allergic reaction can have a wide spectrum of severity; some are categorized as minor and on the other end of the spectrum a reaction can be anaphylactic. *Anaphylaxis* is the most severe state of an allergic reaction and is life-threatening. The symptoms of an anaphylactic reaction are hives, low blood pressure, wheezing (blocking airways), nausea, vomiting, diarrhea, dizziness, and

fainting (Mayo, 2021). The most common food allergens that cause serious reactions in the United States include milk, eggs, fish, wheat, soy, peanuts, tree nuts, and crustacean shellfish.

Allergens can enter an individual's body through many modes of transmission. For example, a food allergy related reaction can be caused not only by ingestion, but also by inhaling, smelling, or directly touching the allergen. A common misconception is that food allergies can only cause severe reactions through direct ingestion. There are other ways for a person to go into anaphylactic shock. Severe allergic reactions can occur when the allergic person directly touches contaminants on a surface, holds the food item, or is in close proximity to the allergen (Mayo, 2022). Adolescence, familial history, race, sex, and asthma contribute to a higher risk of developing a food allergy (Mayo, 2022). Allergies are unpredictable and hard to prevent, but by determining the cause, allergies can be better treated and managed.

Treatment Options

Although there are no direct cures for food allergies, there are a multitude of treatments available. After a minor food allergic reaction, antihistamines are a short-term remedy to help reduce the immediate symptoms. For an anaphylaxis or life-threatening allergic reaction, epinephrine (“EPI-pens”) can be injected in emergency scenarios (Epinephrine, n.d.).

One of the management strategies for certain food allergies is epicutaneous immunotherapy (EPIT) (Food, 2024). This treatment involves wearing a patch with the allergens on the skin to help desensitize the immune system. According to the National Institute of Allergy and Infectious Diseases (NIAID), the Consortium for Food Allergy Research conducted an experiment and found that after one year of treatment with this patch, the individual's tolerance to the allergen had increased dramatically (Food, 2024). The limitations of the patch are that the patient should take precaution when working out or showering to limit the amount of water that touches the patch (Food, 2024). Additionally, sweating with the patch can lead to discomfort and skin irritation. Currently, EPIT is mainly used for peanut allergies, but some patients have used it for milk and egg allergies (Food, 2024).

Sublingual immunotherapy is another option for managing food allergies, which involves putting the allergen underneath the tongue (Food, 2024). The NIAID concluded that researchers are still working on adapting this therapy for food allergies, rather than other types of allergies like environmental, seasonal, etc. (Food, 2024).

Lastly, oral immunotherapy (OIT) is the practice of having the patient eat the food allergen in subsequently increasing amounts to eventually reach a period of tolerance. OIT is a new approach to combating severe food allergies through a deliberate process of allergen dosing (Philadelphia, 2021). The purpose of OIT is to help patients have a less severe or even no reaction when accidental ingestion or cross contamination with the allergen occurs. OIT is not a cure for allergies, however the treatment helps prevent severe reactions. OIT helps increase the quality of life of patients because these individuals do not have to worry as much about a severe reaction. The patients receive initial dosing and then go through “updosing” under medical supervision. Throughout the therapy, the body’s resistance to the food is decreasing and

increasing tolerance of the allergen at higher quantities. The ultimate goal is for the patient to eat the food allergen without an anaphylaxis or life-threatening reaction.

OIT Process

The process of OIT begins with allergy testing at an allergist or immunologist's office. By doing this, the doctor can determine the severity of the allergy or allergies and how many different foods the patient should undergo OIT for (FARE, n.d.). After the initial visit, the patient is given a miniscule portion (e.g. a few milligrams) of the allergen at the doctor's office (FARE, n.d.). The allergen is mixed with a creamy substance such as pudding or Jell-O to disguise the taste and to help the patient digest the allergen better. The patient remains under observation for an hour in the doctor's office to make sure that no reaction has occurred (FARE, n.d.). The patient is then sent home and directed to eat a constant amount of the allergen daily for one to several weeks, and then returns back to the office once the time period has ended for an increased dose, called an *updose*.

The primary up dosing period lasts approximately one year, and by the end of the year, the dose has been increased to a much larger quantity compared to the initial dose (FARE, n.d.). For example, if a patient completes OIT for walnuts, at the first appointment, the patient will eat specks of walnuts no larger than a grain of salt. After two weeks of eating that quantity, the patient will return to the allergist and be instructed to eat a larger amount. This up dosing process continues until the patient eats approximately the equivalent of nine full-size walnuts. The final quantity is dependent on the allergen and the patient's tolerance level. The gradual increase of the allergen allows for the atopic person to develop a higher tolerance (FARE, n.d.). The up dosing phase concludes after a year, at which point the patient has reached a *maintenance* level.

OIT Benefits

The purpose of undergoing OIT is to desensitize the immune system to the allergen and reduce the risk of reactions and the side effects associated with severe reactions. Hence, 80% of OIT patients have been successful in reaching the final maintenance stage (Philadelphia, 2021). Consequently, the probability of anaphylactic allergic reactions have been proven to decrease to very small percentages (Philadelphia, 2021). Successful immunotherapy can result in *desensitization*, in which the patient can eat a defined amount of food allergen without a reaction.

OIT Limitations

OIT is a large time commitment and requires lifelong maintenance. Even after a patient has reached the maintenance phase, a patient has to dose for OIT on a daily basis. After some time, once the patient has built up a high enough tolerance, some doctors have approved patients to ingest maintenance doses only a few times a week rather than daily. A difficult aspect of the treatment is that the patient cannot exercise or take a warm shower for two hours after dosing. This preventative measure is to avoid elevation of the body's general temperature which can

increase blood flow, thereby potentially increasing the allergens spread throughout the body (AAAAI, 2024). Doctors recommend this protocol to reduce the chances of adverse reactions (AAAAI, 2024). Another limitation of OIT is that a small demographic is not eligible to participate in the treatment. This includes people with eosinophilic esophagitis (EoE), a severe allergic condition occurring in the esophagus, leading to severe allergic reactions. It is simply too dangerous to start the OIT process for people with EoE because the allergic reactions can be anaphylactic and extremely severe even with such miniscule food quantities (Wilson, 2023).

Methods

The research for this study was conducted via an online survey, distributed to current and former OIT patients in the United States. The survey was created using Google Forms for ease of use and data organization. The survey was sent to patients through allergy groups on different social media platforms such as Facebook. The survey was completely voluntary and anonymous, and no compensation was given to recipients. Consent was received from the respondent prior to completing the survey. Participants were asked to answer a set of 12 questions and were provided with information about the study objectives. The survey received 202 responses from patients, or from the patients' caregivers, who have undergone or are currently undergoing OIT. The data was compiled and analyzed by putting all the graphs side by side and drawing conclusions on the impact of OIT on patients' lives.

Results

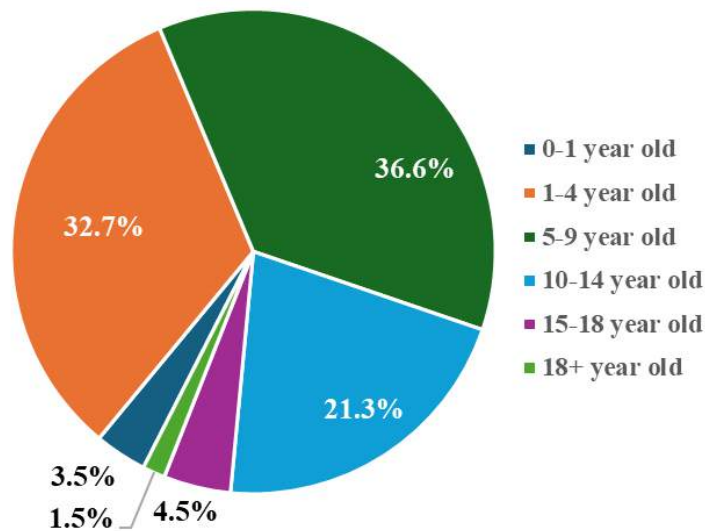


Figure 1: The age distribution of the individuals who underwent the OIT process.

As OIT is a relatively new treatment option, 94.1% of the patients were under the age of 14 when they started OIT (Fig. 1). Additionally, the highest demographic of children undergoing OIT fall into the 5-9 year old range, according to the graph above (Fig. 1).

The data also evaluated where individuals fell in the two main phases of OIT, the dosing phase and the maintenance phase (Burbank et. al, 2016). The *dosing phase* is when the patient has a daily dose of the allergen at the prescribed quantity, and the patient is increasing the dose approximately every two weeks for about a year (Burbank et. al, 2016). After the dosing phase is the maintenance phase. The *maintenance phase*, however, can be a life-long process, with a patient ingesting the allergen usually on a daily basis. To further clarify, the maintenance phase is the final stage of OIT in which the patient eats a constant amount of the allergen for the rest of their life, rather than continually up dosing. In the case when a patient is completing OIT for multiple allergies, the *mixed phase* can occur when one set of allergens are in the maintenance phase, while the other set of allergens are still in the dosing phase (Burbank et. al, 2016).

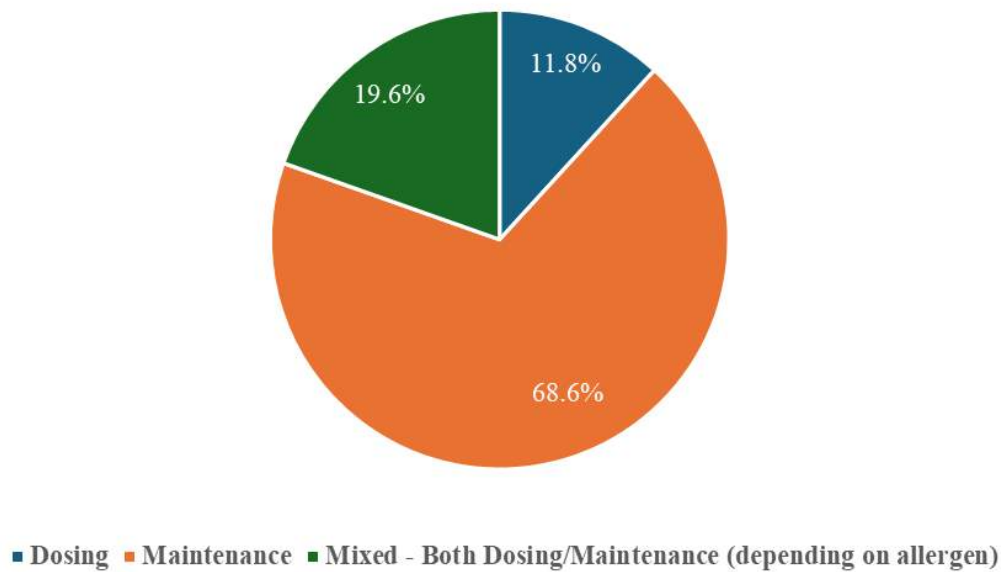


Figure 2: The phase of the patient's OIT treatment.

Another question the survey addressed was how long the patient has participated in OIT at the time of completion of the survey. Of the participants, 42.6% have undergone OIT for over three years, 34.2% have undergone OIT for one to three years, and the remaining 23.3% of patients had been in OIT for less than one year.

The majority of the responses, 69.3% of all participants, state that they underwent OIT for peanuts. This data makes sense as peanuts are the only allergen that the Food and Drug Administration has formally approved for OIT (AAAAI, 2024). After peanuts, the next largest category was tree nuts, which 37.1% of respondents have completed. Other common allergens

include (in subsequent order) eggs, milk, sesame, and wheat. Fewer than 5.0% of those surveyed underwent OIT for mustard, shrimp, pine nut, sunflower seed, and flax seeds.

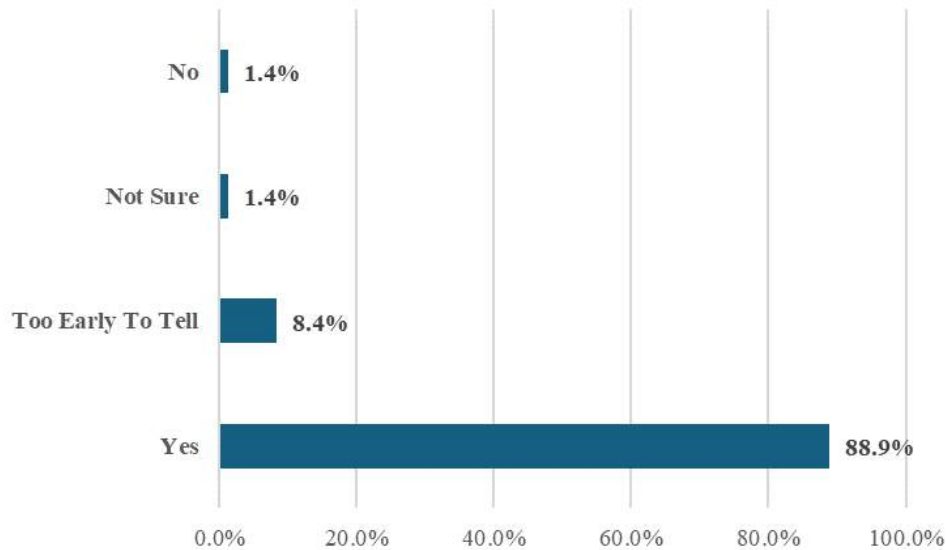


Figure 3: Respondents were asked whether the patient’s quality of life had improved.

The results suggest that the patients are overwhelmingly content with their decision to pursue and follow through with OIT (Fig. 3). A vast majority – 88.9% of participants – indicated that their child’s quality of life improved after completing OIT, and 8.4% stated that it was too early to tell (Fig. 3). Only 1.4% of respondents stated that their child’s quality of life had decreased resulting from OIT (Fig. 3).

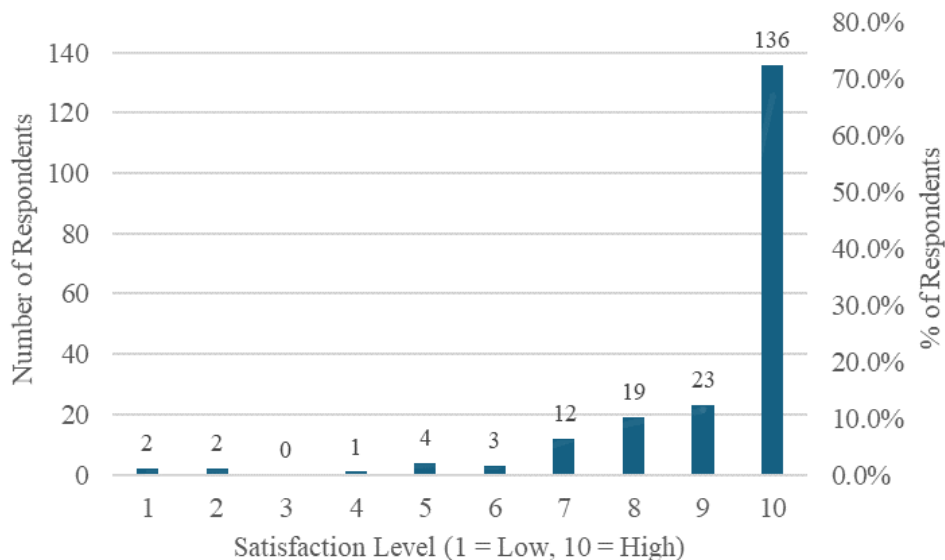


Figure 4: The patient’s level of satisfaction with OIT from one (not at all happy) to ten (extremely happy).

Furthermore, people were generally satisfied with their child’s experience with OIT. 67.3% of people stated that they were “extremely happy” (rating of 10) with their child’s experience with OIT (Fig. 4). Additionally, 94.0% of respondents rated the level of satisfaction at 7 or higher (Fig. 4). As seen in the graph above, only 4.5% of respondents indicate a satisfaction rate of 5 or below (Fig. 4).

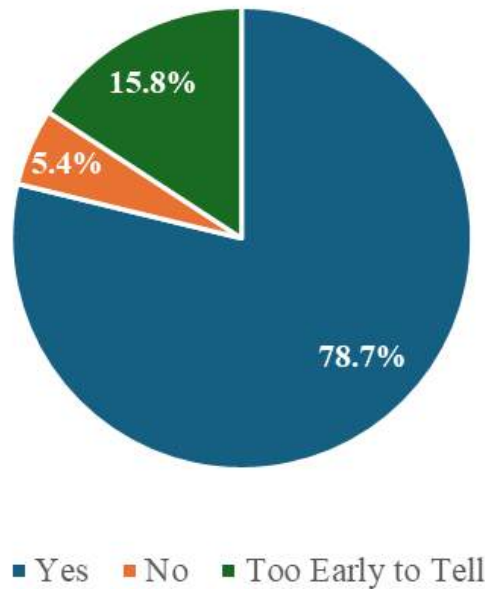


Figure 5: The percentage of individuals whose severity of allergic reactions has decreased since the OIT treatment.

The reason why people are satisfied appears to be the decrease in severity of allergic reactions since the OIT treatment. This graph shows that 78.7% of the sampled demographic indicate that their allergic reaction has decreased (Fig. 5).

Additionally, based on data collected from another question in the survey, 68.8% of responders have been able to safely eat the allergen since beginning OIT. This represents a life-altering shift because the majority of these children had life-threatening allergies and are now safely able to eat their allergens in moderation.

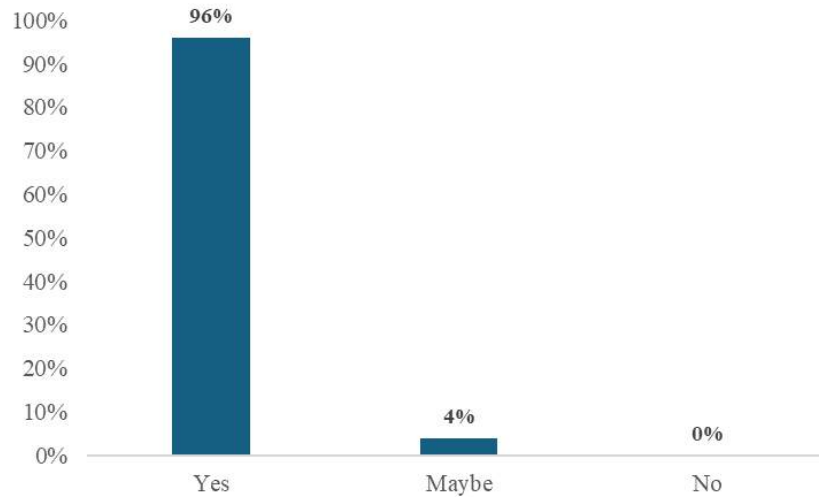


Figure 6: The percentage of participants who would recommend OIT to other children dealing with food allergies.

As further evidence of individual satisfaction, a definitive 96.0% of respondents state that they would strongly recommend OIT to other children dealing with a food allergy of any severity level (Fig. 6). Moreover, there is not one participant out of 202 participants who stated that they would not recommend OIT (Fig. 6). It can be concluded that participants were satisfied with the overall OIT process.

One of main issues with expanding the use of OIT is that even among those with food allergies, the patients are not aware that OIT exists as a possible treatment. This is because a majority of allergist and immunologist offices do not provide OIT treatments. Furthermore, as the FDA has not approved OIT for food allergens besides peanuts, most atopic people are not aware of OIT as a possible treatment option.

Only 34.2% of respondents to my survey found out about OIT from their allergist, immunologist, or medical specialist, suggesting that in order to increase awareness, allergists and immunologists need to promote and inform their patients of the possible treatment. On the other hand, 40.1% of participants found out about OIT from the internet, and 17.3% of responders claim that they found out about OIT from family and or friends.

Lastly, 81.4% of participants believed OIT was not a well-known treatment for food allergies. This indicates a problem in allergy treatment because the majority of people who took this survey believe that OIT increased their quality of life, but many families are unaware of this treatment. One responder exclaimed in response to the survey about OIT, “Love this! I have been saying for years to anyone who will listen that OIT is NOT well known enough.”

Conclusion

The results of the survey indicated that people who underwent or are undergoing the OIT process are overwhelmingly satisfied with OIT, the severity of their allergens have significantly decreased, and these patients would definitely recommend this treatment option to others. The

vast majority of respondents (over 88%) indicated that the patient's quality of life improved after completing OIT.

Many of the responders added remarks from their successful experiences with OIT. One thankful parent of a child who completed OIT wrote, "I am grateful for the peace of mind that OIT has granted me." Another recipient shared that their family is "Incredibly blessed for the success we had in OIT. My son doesn't even have to think about his allergy." Another memorable quote was, "OIT has been life-changing for my son."

Works Cited

- AAAAI Group. (2023, November 13). Allergic Reactions. American Academy of Allergy Asthma & Immunology.
<https://www.aaaai.org/tools-for-the-public/conditions-library/allergies/allergic-reactions#:~:text=Your%20immune%20system%20overreacts%20by,stomach%20or%20on%20the%20skin>
- AAAAI Group. (2024, January 10). The Current State of Oral Immunology. American Academy Of Allergy Asthma & Immunology.
<https://www.aaaai.org/tools-for-the-public/conditions-library/allergies/the-current-state-of-oral-immunotherapy#:~:text=Peanut%2C%20egg%20and%20milk%20OIT,efficacious%20for%20every%20food%20allergy>
- Allergies. (2015, September). Asthma & Allergy Foundation of America.
<https://aafa.org/allergies/>
- Allergies and the immune system. (2021, August 8).
<https://www.hopkinsmedicine.org/health/conditions-and-diseases/allergies-and-the-immune-system>
- Allergies and the Immune System. Johns Hopkins Medicine. (2024, May 23).
<https://www.hopkinsmedicine.org/health/conditions-and-diseases/allergies-and-the-immune-system#:~:text>
- Burbank, A. J., & Wood, R. A. (2016). Oral immunotherapy. Oral Immunotherapy - an overview | ScienceDirect Topics.
[https://www.sciencedirect.com/topics/medicine-and-dentistry/oral-immunotherapy#:~:text=During%20oral%20immunotherapy%2C%20food%20is,tolerated\)%20dose%20occurs%20at%20home.](https://www.sciencedirect.com/topics/medicine-and-dentistry/oral-immunotherapy#:~:text=During%20oral%20immunotherapy%2C%20food%20is,tolerated)%20dose%20occurs%20at%20home.)
- Epinephrine auto-injector and anaphylaxis. Illinois Department of Public Health. (n.d.).
[https://dph.illinois.gov/topics-services/prevention-wellness/epinephrine-autoinjector-and-anaphylaxis.html#:~:text=Epinephrine%20\(adrenaline\)%2C%20a%20self,anaphylaxis%20to%20be%20most%20effective.](https://dph.illinois.gov/topics-services/prevention-wellness/epinephrine-autoinjector-and-anaphylaxis.html#:~:text=Epinephrine%20(adrenaline)%2C%20a%20self,anaphylaxis%20to%20be%20most%20effective.)
- FARE. (n.d.-b). Oral immunotherapy (OIT). FoodAllergy.org.
<https://www.foodallergy.org/resources/oral-immunotherapy-oit>
- Food Allergy Research & Education. (2024b). Food Allergy Awareness Week. FoodAllergy.org.
<https://www.foodallergy.org/our-initiatives/awareness-campaigns/food-allergy-awareness-week#:~:text=In%202024%20We%20Celebrated%20Allies,patient%20to%20the%20emergency%20room.>
- Food Allergy Treatment Research. National Institute of Allergy and Infectious Diseases. (2024, May 23). <https://www.niaid.nih.gov/diseases-conditions/food-allergy>
- Mayo Clinical Staff. (2021, October 2). Anaphylaxis. Mayo Clinic.
<https://www.mayoclinic.org/diseases-conditions/anaphylaxis/symptoms-causes/syc-20351468>
- Mayo Clinic Staff. (2022, August 5). Allergies—Symptoms and causes. Mayo Clinic.
<https://www.mayoclinic.org/diseases-conditions/allergies/symptoms-causes/syc-2035149>

- Philadelphia, T. C. H. of. (2021, April 28). Oral immunotherapy 101: Learning module [Text]. <https://www.chop.edu/centers-programs/oral-immunotherapy-program/oral-immunotherapy-101-learning-module>
- What is Allergy? (2024, March). Australasian Society of Clinical Immunology and Allergy. <https://www.allergy.org.au/patients/about-allergy/what-is-allergy>
- Wilson, B. E., Meltzer, E. C., & Wright, B. L. (2023, August 10). Ethical implications of continuing oral immunotherapy after the development of eosinophilic esophagitis. *The Journal of Allergy and Clinical Immunology: In Practice*. <https://www.sciencedirect.com/science/article/abs/pii/S2213219823009121#:~:text=Oral%20immunotherapy%20is%20associated%20with%20gastrointestinal%20complications%20and%20may%20induce%20EoE.&text=Eosinophilic%20esophagitis%20is%20a%20chronic,prevent%20fibrostenosis%20and%20esophageal%20stricture>.

An Analysis of the Orthodontic Section of a Dental Clinic in Ho Chi Minh City By Khoi A. Lam, Thuy T. Nguyen, Thu N. Y. Nguyen, Trung V. Lam, Kevin Lam

Abstract

This study aimed to analyze different aspects of a dental clinic in Ho Chi Minh City, Vietnam, particularly in orthodontics. The study was carried out from 25 June to 8 July 2024 at Dental Clinic 79, a high-end dental clinic in Ho Chi Minh City, Vietnam. All orthodontic patient profiles from 2020 to 2023 were analyzed, tallying up to 111 profiles. Using data from these profiles, this study mainly explores trends and relationships between age groups, treatments, and check-up rates. Data were sorted out for easier analysis in the form of tables, line graphs, and bar graphs. Several chi-square tests were conducted to figure out the correlation between age groups and treatments. The result shows that, from 2020 to 2023, there is a significant increase in the total number of patients, particularly in the children age group, and the number of types of orthodontic treatments used. The COVID-19 pandemic, increasing parents' awareness, social media, softer food, and environmental factors might have led to these changes. At this clinic, teenagers and adults use braces and aligners at the same rate, and teens use face masks more often. Braces require the most check-ups, followed by aligners. Combining different treatments solves not only multiple problems at the same time but is also likely to reduce the number of check-ups required. Other orthodontic tools for children demand much fewer check-ups because children's developing teeth and jaws can be easily adjustable. The pros and cons of this clinic are analyzed through observation, interaction with dentists, and the experience of being a patient at this clinic. The clinic did a good job retaining patients by providing exceptional care and communicating professionally, as evidenced by the increase in the number of referred patients throughout the years. A variety of orthodontic treatments, a calming clinic environment, and a good dentist's reputation are some of this clinic's strengths. This study will give readers an overview of the age and treatment distribution of patients at a dental clinic in Ho Chi Minh City. Dentists who are planning to open a clinic can evaluate the pros and cons to have a better start.

Keywords orthodontics, Vietnam dental clinic, ho chi minh, age groups, dental treatments

Introduction

The data on orthodontically related hearings or those involving orthodontic treatment providers from 2015–2018 from the General Dental Council (GDC) website of the U.K. states that the total number of registrant cases involved in orthodontics is 8% of the total number of dental cases [1]. Meanwhile, in Vietnam, the prevalence of a moderate need for orthodontic treatment among 18-year-olds was 24%, while the definitive need was 30.5%. Among 12-year-olds, the moderate need was 21%, and the definitive need was 19% [2]. This shows that dental clinic practices, particularly orthodontics, are in high demand in Vietnam. However, there aren't many thorough investigations that provide an overview of these practices. The insights in this research will help readers resolve some of their questions about orthodontics, provide

experts with an overview of the private dentistry practices in Ho Chi Minh City, and guide dentists at this clinic in seeking improvements. This research paper will examine different aspects of patients' records to identify trends and relationships between age groups, treatments, and check-up rates. Additionally, it will point out what causes these trends using recent events. We will also highlight the pros and cons and aspects that stand out in this clinic based on the data collected from observations and interactions with dentists at this clinic.

Methodology

The first part of this section will provide an introduction to this clinic. This clinic opens from Monday to Saturday from 5 p.m. to 8 p.m. because dentists and dental assistants at this clinic also work at local hospitals during the day. The clinic has two floors, one for reception and the other has a dental treatment room with two dental chairs. This dental clinic offers a wide range of treatments, from regular tooth cleaning, extraction, and filling to specialized treatments like dental implants and orthodontics. The pricing for orthodontic treatment at this clinic is considered average and reasonable. In 2023, the clinic handled a total of 364 cases, comprising 14 dental implant cases and 34 orthodontics cases. The total number of orthodontic cases is 9.3% of the total number of dental cases. This information suggests that the orthodontic section of this clinic is relatively stable. Using 2020 to 2023 orthodontic patients' records, data on patient gender, age, what kind of orthodontic treatment they used, whether their treatment was finished or not, the number of check-ups they had, the fee of their treatment, the reason for visit, and whether they received a 5% discount are collected. Then, unnecessary data are filtered out, and the relevant ones are organized into tables and graphs for easier comparison and analysis, utilizing only line and bar graphs. To examine the correlation between braces, aligners, and face masks with different age groups, several chi-square tests are conducted to prove whether there is a significant difference between the number of people from certain age groups using certain treatments. The pros and cons of this clinic are analyzed through observation, interaction with dentists, and the experience of being a patient at this clinic.

Results

- 1) Representation of the distribution of age, treatment method frequency, patients' awareness of the clinic, and check-up rate from 2020 to 2023:**

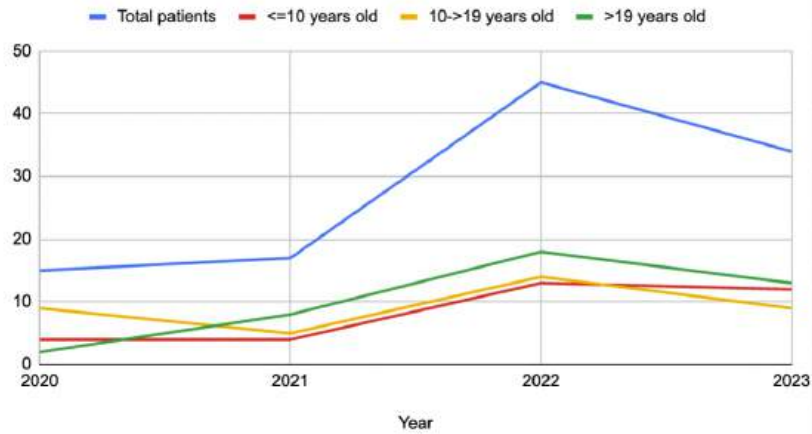


Figure 1: Yearly Distribution of Patients by Age Group from 2020 to 2023

This graph allows for the observation of correlations between age and treatment usage and helps identify whether any trend is present. We can see a massive increase in patients starting in 2022, with the number of child patients rising from third place to nearly first place by 2023.

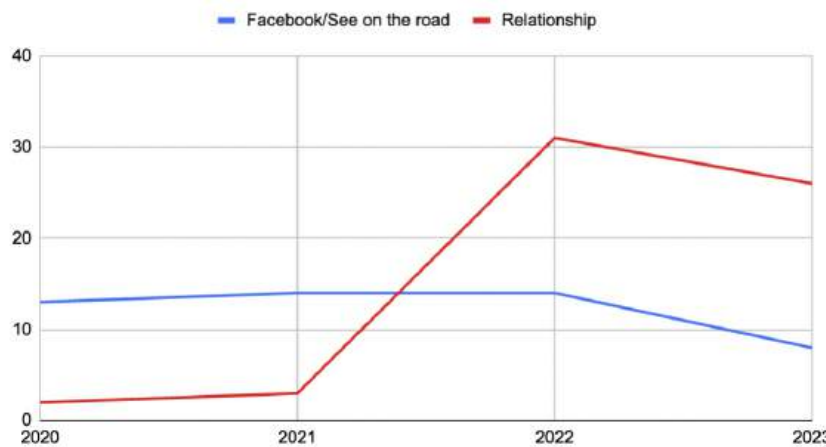


Figure 2: Sources of Clinic Awareness by Year from 2020 to 2023

This graph shows what brought patients to this clinic. We can see how this trend might affect other trends in other graphs and also point out an advantage of this clinic. Toward 2023, the number of referred patients has increased significantly.

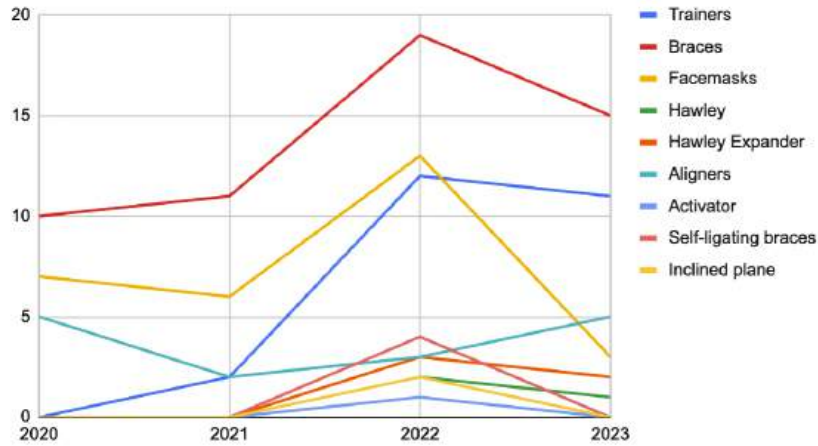


Figure 3: Yearly Usage of Orthodontic Treatments from 2020 to 2023

The usage of trainers has increased dramatically over the years. Braces have been used predominantly over the years, followed by facemasks. There is a huge decline in both braces and facemasks in 2023. Facemasks and braces are used together 19 times over the course of 4 years.

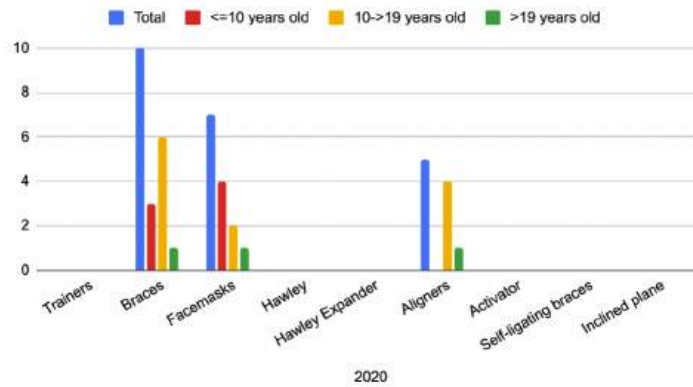


Figure 4: Distribution of Orthodontic Cases by Age Group in 2020

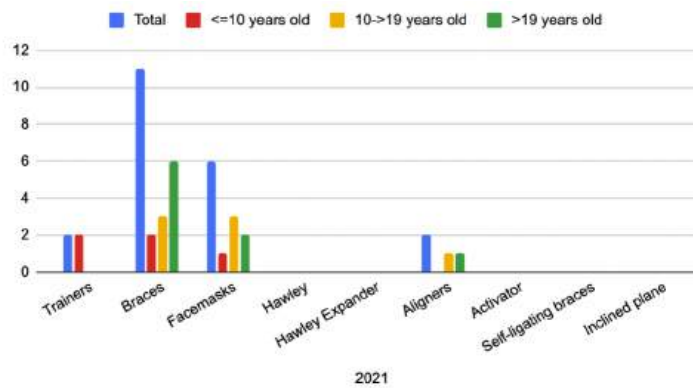


Figure 5: Distribution of Orthodontic Cases by Age Group in 2021

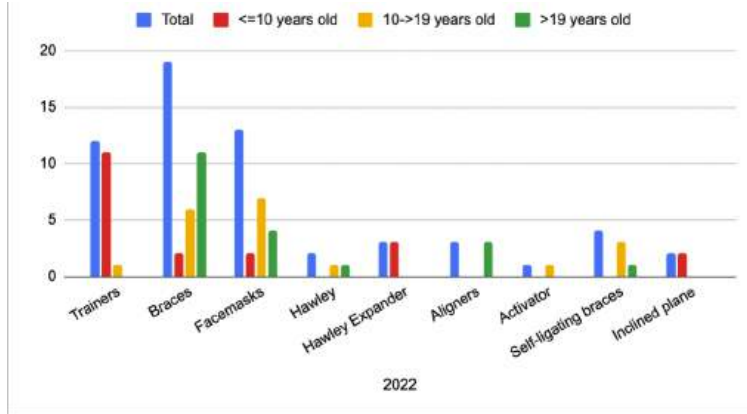


Figure 6: Distribution of Orthodontic Cases by Age Group in 2022

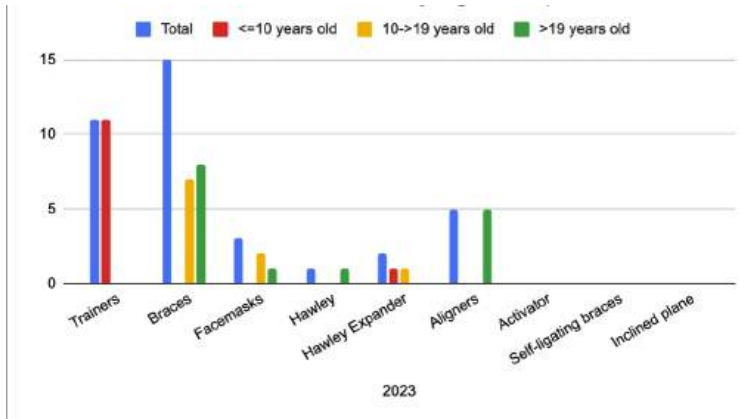


Figure 7: Distribution of Orthodontic Cases by Age Group in 2023

These four bar graphs allow for the observation of the changes in the amount and variety of each orthodontic treatment as the distribution of age and total number of patients change over the years. Many other treatments were introduced in 2022 and 2023 compared to 2020 and 2021, like Hawley, Hawley Expander, Activators, etc.

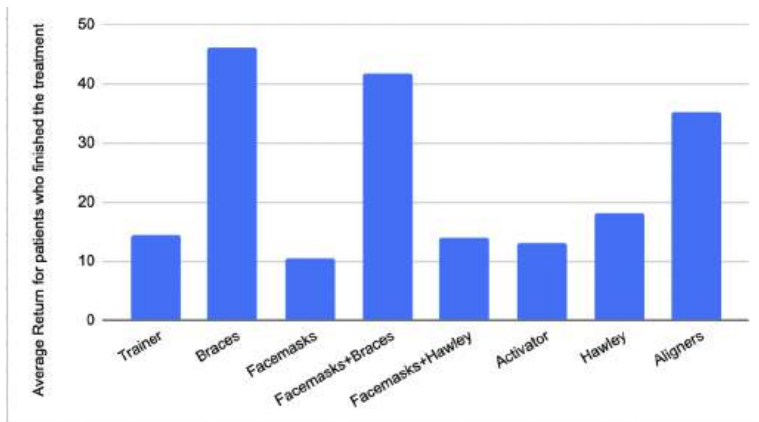


Figure 8: Average Return for Patients by Orthodontic Treatment Type

The patient's average number of check-ups for each type of orthodontic treatment, including combined treatments, is presented. This allows for the interpretation of the pros and cons of each treatment. Patients treated with braces return most often.

2) **Analysis of the correlation between age groups and different orthodontic treatments:**

Table 1: Orthodontic Treatment Distribution by Age Group

	<=10 years old	10->19 years old	>19 years old	Total
Braces	7	22	26	55
Aligners	0	5	10	15
Facemasks	7	14	7	28
Trainers	24	1	0	25
Hawley	0	1	2	3
Hawley Expander	4	1	0	5
Activator	0	1	0	1
Self-ligating braces	0	3	1	4
Inclined plane	2	0	0	2

Table 2: Total number of patients by age group

	<=10 years old	10->19 years old	>19 years old	Total
Total patients	33	37	41	111

The number of different orthodontic treatment (every treatment) cases for each age group is presented in this table.

In this clinic, there is no case when one patient uses both aligners and braces, so these two variables are independent of each other. Note that many treatments are used together, so the total number of patients won't be equal to the number of treatments, but that won't affect the result of the chi-square tests. This table allows for the comparison of the usage of each treatment among these age groups.

(The work for each statistical test below is listed in the supplementary materials section.)

a) **Braces:**

Null Hypothesis (H_0): There is no significant difference between the number of 10 to 19-year-old patients and more than 19-year-old patients using braces.

Alternative Hypothesis (H_A): There is a significant difference between the number of 10 to 19-year-old patients and more than 19-year-old patients using braces.

(p-value<0.05 is considered statically significant)

Table 3: Braces vs. Other Treatments by Age Group

Age Group	Braces	Other Treatments	Total Patients
10-19 years old	22	15	37
>19 years old	26	15	41
Total	48	30	78

Chi-square test result: There is no significant difference between the number of 10 to 19-year-old patients and more than 19-year-old patients using braces.

b) **Aligners:**

Null Hypothesis (H_0): There is no significant difference between the number of 10 to 19-year-old patients and more than 19-year-old patients using aligners.

Alternative Hypothesis (H_A): There is a significant difference between the number of 10 to 19-year-old patients and more than 19-year-old patients using aligners.

(p -value <0.05 is considered statically significant)

Table 4: Aligners vs. Other Treatments by Age Group

Age Group	Aligners	Other Treatments	Total Patients
10-19 years old	5	32	37
>19 years old	10	31	41
Total	15	63	78

Chi-square test result: There is no significant difference between the number of 10 to 19-year-old patients and more than 19-year-old patients using aligners.

c) Facemasks:

Null Hypothesis (H_0): There is no significant difference between the number of 10 to 19-year-old patients and more than 19-year-old patients using facemasks.

Alternative Hypothesis (H_A): There is a significant difference between the number of 10 to 19-year-old patients and more than 19-year-old patients using facemasks.

(p -value <0.05 is considered statically significant)

Table 5: Facemasks vs. Other Treatments by Age Group.

Age Group	Facemasks	Other Treatments	Total Patients
10-19 years old	14	23	37
>19 years old	7	34	41
Total	21	57	78

Face mask cases of children below ten years old are lower than those of teenagers.

Chi-square test result: There is a significant difference between the number of 10 to 19-year-old patients and more than 19-year-old patients using facemasks.

3) Results from observation, interaction with dentists, and the experience of being a patient at this clinic:

Now, let's look at the pros and cons of this clinic. The manager did a good job promoting the clinic to more people by offering a 5% discount on referred patients to encourage current patients to introduce people they know to the clinic. According to Dr. Thuy, she sees her patients as her family, so she always takes good care of them. The vast increase in the number of patients, especially referred ones, while the marketing level is low at this clinic (only having a Facebook page) shows that the clinic did a good job curing patients and delivering excellent service with reasonable prices; that's why current patients feel comfortable introducing their friends and families to come to the clinic for dental treatment. Referring patients from public hospitals to

private clinics is also an essential factor contributing to the annual number of patients. This is not something any private clinic could do and is a great strength of this clinic. Many dentists working at this clinic are professionals and are well-known for their specialty at local public hospitals. That's why many patients trust their skills and come to this private clinic for their caring treatment. This action is not considered unethical as these doctors are specialists with decades of experience, capable of giving patients the best care. Furthermore, since the services at public hospitals are not so great but hold many talented dentists, patients will have a much better experience by going to private clinics like the one in this research [3]. The clinic did a good job organizing patients' profiles into clear binders and categorized sections on a bookshelf. Patients' general data are also presented on a spreadsheet for quicker search.

Moreover, this clinic did a good job combining different orthodontic treatments professionally to cure multiple problems simultaneously while ensuring different treatments won't interfere with each other's efficiency. This was proven by the lower amount of check-ups established by patients who used multiple treatments. This also highlights the variety of orthodontic treatments this clinic possesses, ensuring the patients remain at this clinic for the whole treatment process, not having to be transferred anywhere. To make sure patients adhere to the treatment plan and achieve the desired results, the clinic asks the patient to pay the full fee in advance and stick with the plan. This is why every single patient returns for multiple check-ups, and no one quits in the middle. Through observation, the clinic did a good job estimating patients' visits and controlling appointments. That's why, even though there are only two dental chairs, patients rarely have to wait, and if they do, there is a calming and relaxing waiting area with air conditioners, comfortable seats, and soothing Ghibli music to allow patients to chill out while waiting for their turn; this also applies to the setting of the patient sitting on the dental chair. The communication between doctors and patients is natural, friendly, and professional, allowing patients to feel comfortable expressing their feelings without fear. Dentists treat patients like their families, and this is reflected in the simple, easily comprehensible language and the warm, attentive manner in which these professionals interact with patients. These interactions, professionalism, and calming environment encourage patients to return for check-ups regularly. The pricing is reasonable and aligns with the overall market, even when the skills of dentists here are top-tier in Ho Chi Minh City, particularly in the field of periodontics. Dentists' reputations also play a role. Dr. Thuy, the Deputy Head of the Department of Dentistry, University of Medicine and Pharmacy HCMC, highlights the credibility of this clinic, attracting patients with her esteemed position.

Discussion

I'm an undergraduate student at the University of the Pacific, majoring in Pre-Dentistry. I've been collecting data from 25 June to 8 July 2024 at this dental clinic in Ho Chi Minh City, Vietnam. The rising trend in child patients observed in [Figure 1](#) explains why the number of trainees increased exponentially over the years and peaked in 2023, as observed in [Figure 3](#). There could be some reasons for this trend. With the growth of social media and globalization,

facial beauty is undeniably essential, encouraging people to seek orthodontic treatments [4]. As parents expanded their knowledge and became more aware of the importance of early orthodontic treatment for their children, parents started to bring their children to dental clinics for orthodontic treatments to enhance their lifestyle and beauty [5]. Recent new technology like Invisalign® eliminates the shame of teenagers and adults wearing braces, further promoting orthodontic treatments to more people in need [6]. Education can also play a role, as more oral health education at schools helps increase children's and parents' awareness of dental care [7]. The evolutionary change in children's teeth nowadays also plays a role in this trend. Food is getting softer as there are more processed foods than before, and today, parents have various ways to allow their children to eat food more easily by breaking down the food. This makes children's jaws weaker and underdeveloped, leading to smaller jaw sizes. Since the teeth size is always the same, this causes overcrowding of teeth and malocclusion as they compete for space, increasing the need for orthodontic care [8]. Environmental pollution also affects children's airways, leading to craniofacial changes like narrow jaws and dental arches [9]. The COVID-19 pandemic could also play a role, causing many people to struggle with their finances through job loss and preventing them from accessing effective high-end dental treatments like orthodontics. In 2020 and 2021, during the pandemic, the nation is still recovering, so the number of patients is not that high, especially children. However, starting in 2022 and 2023, the number of patients has skyrocketed as people have already recovered from the pandemic [10]. The 5% discount policy also explains the increase in the total number of patients as, according to Figure 2, toward 2023, the number of referred patients has increased significantly.

The dominance of braces and facemasks, observed in Figure 3, is likely due to their effectiveness and reasonable price—braces cost approximately 40 million VND, and facemasks cost around 10 million VND. The reason for the increase in trainers is that they are used primarily on children under ten years old, and there has been a massive increase in young children over the years. The introduction of a variety of orthodontic treatments in 2022 and 2023 could be explained by the increase in patient numbers over the years, which has led to more diversity in dental issues and promotes other types of treatments like Hawley and inclined planes, as well as combined treatments. We see a decline in both facemasks and braces in 2023 because as one declines, the other also declines since they are usually used together to address both teeth and face/jaw alignment simultaneously. The number of times these two treatments are used together is also the highest compared to any other type of combination, explaining why these two have such a huge correlation with each other. The cause of this decline is probably because of the decrease in patients in 2023 and the increase in aligner cases, presented in Figures 1 and 3. As addressed before, COVID-19 could play a role in increasing aligner cases as people start to recover financially after the pandemic and are more likely to afford costly but convenient and effective orthodontic treatments like aligners [10].

The result of the first chi-square test states that there is no correlation between being in the teenage age group or the adult age group and the likelihood of having braces as your orthodontic treatment. The popularity of braces in both age groups shows that this treatment is

effective for both teenagers and adults and that this treatment is affordable. Many fewer people use aligners than braces because of the huge difference in price. The second chi-square test proves that there is no significant difference between the number of patients aged 10 to 19 and those over 19 years old using aligners. However, if the amount of samples collected is much higher, there will definitely be a significant difference. The reason for this interpretation is that aligners are so expensive compared to braces that most of the time, only adults with greater financial means can afford them. Parents would be less likely to spend that much on their teenage children's teeth since it's too costly and unnecessary, as braces are an effective and affordable treatment for teenagers [11]. A crucial factor driving adults to choose Invisalign® as their orthodontic treatment is that many people, particularly public speakers or businessmen, have to interact with others in a formal setting and would prefer not to smile with visible braces. Invisalign® is nearly unnoticeable when patients put it on, making it a suitable option for many adults [6]. Furthermore, looking solely at this small sample and result, with approximately the same total patients, the number of adults using aligners is double the number of teens using aligners. This is good evidence that if the sample size is larger, there would be a significant difference in aligner usage between the age groups. Lastly, looking at the third chi-square test, the result shows that face masks are used to treat teenagers more often than adults. A unique thing about this clinic is that, according to Dr. Thu, she applies facemasks to patients of all ages. Although facemasks are said to be less effective for mature individuals, Dr. Thu's results suggest otherwise, so she continues to incorporate this treatment in her patients widely. According to the article called "Skin Irritation in Children Undergoing Orthodontic Facemask Therapy," [12] the age at the initiation of facemask therapy ranged from 5.58 to 13.5 years. However, at this clinic, the number of children under ten years of age using face masks is just about the same as adults. Back to the chi-square test, even though Dr. Thu stated that she widely uses this treatment for all ages, it seems like teenagers establish the greatest effectiveness, especially when combined with braces, because their developing facial bones are flexible [13] and still have space for adjustment while adults don't. A notable point about the data result is that the face mask cases of children below ten years old are even lower than those of teenagers. This could be because there are better orthodontic treatments for children, like trainers and inclined planes.

According to [Figure 8](#), it's clear that braces, with the highest average check-up required, demand the most effort and patience to come to the clinic for check-ups regularly. Combined treatments are shown to have a lower amount of check-ups compared to other uncombined treatments, suggesting the high effectiveness of using multiple treatments at a time. The evidence is shown by looking at the combination of braces and face masks compared to braces alone and by comparing the combination of Hawley and face masks and Hawley alone. Both of these combinations require fewer check-ups than the corresponding individual treatments. Looking at [Figure 3](#), we can see that trainers, Hawley expanders, and inclined planes are used primarily for children under ten years old, whose teeth and jaws are still developing and are quick to adjust [13]. This explains why the number of check-ups required for these treatments is lower than that of other treatments.

Conclusion

From 2020 to 2023, there is a significant increase in the total number of patients, particularly in the under ten years old age group, and the number of types of orthodontic treatments used. Many factors might have contributed to these changes, including COVID-19, increased parental awareness, social media, softer food, and environmental factors. Braces require the most check-ups, followed by aligners. Combining different treatments solves not only multiple problems at the same time but is also likely to reduce the number of check-ups required. Other orthodontic tools for children require much fewer check-ups because their teeth and jaws are still developing and can be adjusted easily. The clinic did a good job retaining patients by providing exceptional care and communicating professionally. A variety of orthodontic treatments, a calming clinic environment, and a good dentist's reputation are some of this clinic's strengths.

Limitation of the Study

A limitation during the data-collecting process is that the success rate can't be measured, especially for orthodontic cases. According to Dr. Thuy, whether a patient's case is successful or not depends on whether their teeth stay where they are after a few years of using retainers. That is why there is no "unsuccessful" case just by looking at current patients' profiles. The sample size of some of the targeted sample groups is not large enough, so some data from the statistical test might not be representative. There is no feedback section for patients, so it's impossible to evaluate the professionalism of dental practices or whether there are any clinical limitations. With the chi-square tests, we were able to see that teenagers and adults use braces and aligners at the same rate, and teens use face masks more often. However, for the aligners case, if the sample was larger, I suspect that many more adults would use aligners than braces for various reasons.

Author Contributions

Data access permission, Thuy N.; data collection, Khoi L.; data analysis, Khoi L. and Thu N.; conceptualization, Khoi L., Thuy N., and Thu. N; clinic observation permission, Thuy N.; writing—original draft preparation, Khoi L.; methodology, Khoi L.; statistical analysis, Khoi L.; discussion, Khoi L.; clinic information check, Thuy N.; patients' data provider, Thuy N.; review and editing, Thuy N., Thu N.; Trung L., Kevin L.. All authors have read and agreed to the published version of the manuscript.

Works Cited

1. "Dental Data from Clinic 79." Google Sheets, Google, docs.google.com/spreadsheets/d/1xyLcp3CtNuE_4NFkGyXtFHYZ34kOruv35O5rOw7Nejc/edit?usp=sharing.
2. "Clinic 79 Data Access Permission." Google Drive, Google, https://drive.google.com/file/d/1R4aUEBAoqZIRH7WsmR5yLQHbTJzhXO6q/view?usp=sharing.
3. "Chi-Square Test." Google Drive, Google, drive.google.com/file/d/1GyF8B9tBgsq93hgZzAeFA7oJw9RRW7J/view?usp=sharing.
4. "Table 5." Google Drive, Google, drive.google.com/file/d/1M6uteZ5DUctayGg6bb7HMAZJyFQ08Tbn/view?usp=sharing.
5. "Table 6-9." Google Drive, Google, drive.google.com/file/d/1GUZ6sd5UpP2Q1502D7yglOx8VOBwXINz/view?usp=sharing
6. "Table 10." Google Drive, Google, drive.google.com/file/d/1wfU6w_gaefJ6lcYCmap6Rz7NWyXSnZty/view?usp=sharing.
1. Ahmad, Z., & Singh, P. (2023). A four-year review of orthodontic cases appearing before the General Dental Council Professional Conduct Committee. *J Orthod*, 50(4), 391-399. doi:10.1177/1465312520954219.
2. Nguyen, Son Minh, et al. (2014). The Need for Orthodontic Treatment among Vietnamese School Children and Young Adults. *International Journal of Dentistry*, vol 2014, 5 pages.
3. Barnea R., Tur-Sinai A., Levtzion-Korach O., Weiss Y., Tal O. (2014). Patient Preferences and Choices as a Reflection of Trust-a Cluster Analysis Comparing Postsurgical Perceptions in a Private and a Public Hospital. *Health Expectations : An International Journal of Public Participation in Health Care and Health Policy*, vol 25, 2340–2354.
4. Merino, M., Tornero-Aguilera, J. F., Rubio-Zarapuz, A., Villanueva-Tobaldo, C. V., Martin-Rodriguez, A., & Clemente-Suarez, V. J. (2024). Body Perceptions and Psychological Well-Being: A Review of the Impact of Social Media and Physical Measurements on Self-Esteem and Mental Health with a Focus on Body Image Satisfaction and Its Relationship with Cultural and Gender Factors. *Healthcare (Basel)*, 12(14). doi:10.3390/healthcare12141396.
5. Aldweesh, A. H., Ben Gassem, A. A., AlShehri, B. M., AlTowaijri, A. A., & Albarakati, S. F. (2022). Parents' Awareness of Early Orthodontic Consultation: A Cross-Sectional Study. *Int J Environ Res Public Health*, 19(3). doi:10.3390/ijerph19031800.
6. Sharma, R., Drummond, R., Wiltshire, W., Schroth, R., Lekic, M., Bertone, M., & Tate, R. (2021). Quality of life in an adolescent orthodontic population. *Angle Orthod*, 91(6), 718-724. doi:10.2319/062820-592.1.
7. Eun, H. R., Park, J. T., & Jang, J. H. (2022). Factors Related to the Intention to Use Dental Care by Industrial Workers Due to COVID-19: Application of Anderson Model and Planned Behavior Theory. *Int J Environ Res Public Health*, 19(19). doi:10.3390/ijerph191912883.

8. Anand, T., Garg, A. K., & Singh, S. (2022). Effect of socioeconomic, nutritional status, diet, and oral habits on the prevalence of different types of malocclusion in school-children. *Acta Biomed*, 93(3), e2022161. doi:10.23750/abm.v93i3.13027.
9. Brennan, L. C., Kirkham, F. J., & Gavlak, J. C. (2020). Sleep-disordered breathing and comorbidities: role of the upper airway and craniofacial skeleton. *Nat Sci Sleep*, 12, 907-936. doi:10.2147/NSS.S146608
10. Mallah, S. I., Ghorab, O. K., Al-Salmi, S., Abdellatif, O. S., Tharmaratnam, T., Iskandar, M. A., . . . Al-Qahtani, M. (2021). COVID-19: breaking down a global health crisis. *Ann Clin Microbiol Antimicrob*, 20(1), 35. doi:10.1186/s12941-021-00438-7
11. Ke, Y., Zhu, Y., & Zhu, M. (2019). A comparison of treatment effectiveness between clear aligner and fixed appliance therapies. *BMC Oral Health*, 19(1), 24. doi:10.1186/s12903-018-0695-z
12. Kim, H., Kim, J. S., Kim, C. S., Becker-Weimann, S. Y., Cha, J. Y., & Choi, S. H. (2023). Skin irritation in children undergoing orthodontic facemask therapy. *Sci Rep*, 13(1), 2200. doi:10.1038/s41598-023-29253-0
13. Sophia Matos, Matthew D. Johnson. Pediatric Craniomaxillofacial Fractures: A Review. *Facial Plastic Surgery Clinics of North America*, Volume 32, Issue 1, 2024, Pages 1-12, ISSN 1064-7406, ISBN 9780443131059, doi:10.1016/j.fsc.2023.06.008.

How did the COVID-19 Pandemic Impact the Mental Development of Children Across Different Socioeconomic Backgrounds? By Zhiyuan Liu

Introduction

Starting from 2019, the COVID-19 became a global pandemic impacting countries around the world. In March 2020, the World Health Organization (WHO) had declared COVID-19 as a pandemic due to the unpredictable impact of it on humans around the world (Saggiro de Figueiredo et al., 2021). The COVID-19 spreaded to over 216 countries by August, 2020. It not only caused both economic and social recession, but also disrupted the health and education systems. One of the biggest concerns during the pandemic was its impact on children's mental development, within the various socioeconomic statuses (SES) (Shah et al., 2020). Due to major changes such as school closure, family financial instability, parent's stress level, and limited resources, children were especially vulnerable during this period of time. With these critical factors, it is important to address them and accommodate children to relieve their burdens and also prepare for future possible pandemics. This literature review aims to present ways the COVID-19 pandemic had severely impacted children's mental development through different SES, focusing on the aspect of family environment, school disruptions, and biological stressors.

Family Environment and Quarantine Situation

The COVID-19 pandemic has influenced children's developmental environments worldwide from different perspectives. During quarantine, parent's and their children spent most of their time within their households. A lot of parents lost their jobs or were asked to work remotely, in the meantime, they also had to take care of their children. This leads to an increase in a parent's daily stress level. It is important to understand and examine how these stresses can affect children's mental health development, with the connection to SES, to come up with better solutions. This section looks at how the pandemic influences children's emotional and cognitive development within the quarantine family setting, with a specific focus on the difficulties encountered by families with lower SES.

Due to job insecurity and the increase of unemployment rate during the pandemic, parents struggled with financial situations to support the family. Many of them faced mental health issues led by stress, which can directly impact their children's well being. A study by Hendry et al. (2023) revealed that parents who multitasked with remote work and childcare faced much higher pressure and anxieties. This limited their ability to provide emotional support and quality family time with their children. A study by Saggiro de Figueiredo et al. (2021) found correlation with poor parental mental condition to increase in behavioral and emotional difficulties for their children. Stress that came from unemployment, reduced income, and the uncertainty of the pandemic worsened parents' mental health conditions. These factors lead to increased reported cases of domestic violence, and create a dangerous household environment for children to live in. Domestic violence causes long term mental trauma for children. (Ashok et al., 2021). As the tension in the family increased, children graduated and lost their sense of

security at home, they often felt uncertain about their future. However, not every child had unstable experiences during COVID-19. The amount of impact depends heavily on the situation of different families and the resources available to them. SES played a key role in determining the extent of children being affected.

These impacts vary among different family circumstances and SES. Children from lower SES families had particular greater challenges due to economic instability and limited access to resources. In contrast, children from higher SES families had more access to resources available to help them cope with the negative impact from the pandemic. Parents from higher SES families are less likely to become unemployed and face financial difficulties. They often had access to a stable household environment, internet, and less mental stress which all helped to provide their children a better environment to stay (González et al., 2022). In contrast, children from lower SES families face significant barriers that further impact their development. These families have limited access to stable internet and even have a shortage of food resources. The parents faced great challenges, therefore had less time spent supporting their children, which increased children's stress and uncertainty (George et al., 2021). Many parents from the lower SES families worked in high-risks jobs which can increase their exposure to COVID-19 and bring it back home to their children (González et al., 2022). These parents had less time to focus on their children's physical and mental health. Children's experience during the pandemic strongly depended on their family's SES. The challenges faced by lower SES families may produce a long-term impact on children's development.

Understanding the different impacts on children's development is crucial for developing appropriate strategies for supporting their physical and mental well-being. There need to be solutions to specifically address the challenges faced by children especially from lower SES families around the world. Improving access to resources, family and community support are also important to contribute for children's development during and after the pandemic. Aside from the issues with parent's mental health and home environments, school recession and the switch to virtual learning mode are also key factors that impact children's life during the pandemic.

Impact of School Closures and Online Learning on Children's School Readiness During the Pandemic

The COVID-19 led to school closures across the world, with a great impact on education over billions of students from various countries. These disruptions harmed several areas on children's development, including school readiness, emotional well-being, social skills, and physical health (Pfefferbaum, 2021). Prior to the pandemic, children across different SES families already had educational inequalities. When school shifted to online learning, many students faced difficulties to accommodate this new learning environment and it was not an effective learning method for everybody. Learning system change made the existing education gap between students from lower to higher SES even further apart. While online courses allowed school to continue, students from lower SES families faced challenges for not having access to

suitable devices and safe learning environments (George et al., 2021). Online learning slowed down student's academic progress and limited their social interactions with their peers and instructors. Social interactions are significantly important for children's cognitive development. School closures disrupted the original learning structure and daily routine that school provided, leading to more confusion and negative impact on children. A study by Zimmerman and Anderson (2021) provided insight on how school closure reduced the amount of resources to access, such as playgrounds in the parks, library, museums, and community events that all are essential during children's development. Aside from community resources, schools also provide health services especially to accommodate disadvantaged children who might not have these services at home. During the pandemic, these students lost the access of those resources, which can be both impacting to them and their families (Pfefferbaum, 2021).

Additionally to extracurricular activities and school supporting programs, students' school readiness were severely impacted during the pandemic. A study by González et al. (2022) revealed that Uruguayan children from lower SES families experienced significant decreases in both cognitive and motor skills. The pandemic has led to a severe impact on students' educational progress. The Child Development Inventory (INDI) was used for this study to measure school readiness among Uruguayan children. It discovered that children from higher SES families faced less negative effects than those from lower SES backgrounds. The educational gaps between SES further lead to learning losses during the pandemic. A study by Bailey et al. (2021) used a survey of 200 educational researchers and predicted that the academic performance from reading and math skills would worsen significantly in the years after the pandemic. Especially for those children from lower SES families. The inequalities between children across different SES were expected to last for at least a few more years after the pandemic is over. These findings imply that lower SES families are more likely to have children to fall behind academically due to the potential factors such as limited access to educational resources and a stable living environment. School readiness is very essential for a future academic career. Both caregivers and educators should work on supporting their children's education and address their emotional needs. Communities should help children especially from lower SES families to overcome challenges of school closure and online learning to ensure they will not fall further behind their peers. The changes brought by COVID-19 disrupted the daily routine and increased stress levels for children, these can impact their brain and neurological development. Which linked back to the reasons for school readiness and mental development.

Biological impacts

During the COVID-19 pandemic, children faced biological challenges due to the disruptions from their education and home environment. The pandemic increased more stressors for children's daily life, including issues such as food insecurity, lack of family support, household and neighborhood safety, and possibility of domestic violence (Graves & Nowakowski, 2017). These factors contribute to emotional stress and can trigger several biological responses that increase the effect on their development. To understand these biological

impacts, it is important to recognize how SES can lead to increased stress and affect the brain development. Particularly in areas such as the prefrontal cortex (PFC) and hippocampus. These brain regions are essential for social, emotional, and cognitive functions (George et al., 2021).

The PFC is responsible for executive functions such as paying attention, making decisions, and self regulation. The hippocampus is responsible for memory formation. Both of these areas were easily affected by long-term stress. When these brain areas were damaged, children began to struggle with performing academic tasks such as completing homework assignments, paying attention in classes, and performing well on exams. All of these tasks were crucial for school readiness and academic success. A study done by George et al. (2021) showed that repeat exposures to stress and emotional changes disrupted the hypothalamic-pituitary-adrenal (HPA) axis, which was a stress response system in our brain.

Disruption led to further consequences by impacting the immune system and cognitive function, which resulted in poor mental and physical health outcomes that negatively interrupted children's progress to academic success. Furthermore, prior history of stress from childhood experience caused neuroinflammation, which was linked to a risk of psychological disorders such as depression and anxiety (Saggiore de Figueiredo et al., 2021). Children were already considered the vulnerabilities during the pandemic, and with the biological impact, they were exposed to more psychological disorders like ADHD, mood disorders, and more behavioral issues (Shah et al., 2020). These challenges led to long-term effects on children's health, impacting various aspects of their academic performance and social interactions with those around them.

Understanding the biological impact was important for caregivers and educators to help children with the challenges they faced during the pandemic. Children, especially those from lower SES families, already faced financial problems and limitations to access resources, and they were more likely to fall behind in school and experience more stress. These findings highlighted the need to focus on both the educational and emotional needs for the development of children, especially those that faced the most risk. It was important to reduce their long-term effect of stress led by the pandemic. Educators and the community should have worked together to prioritize the emotional well-being and development for children and made plans to help them overcome these challenges and also come up with actions in case of future pandemics.

Community and Family Strategies to Prevent Negative Impacts on Children's Mental Health

Children from higher SES families had a higher chance to recover from the negative impact on their daily life led by the pandemic, since they had more available resources around them. While students from lower SES families don't have these advantages, which effected both emotional and educational outcomes (George et al., 2021). To help these children, specific strategies are needed for them to get back on track. Social isolation is one of the biggest issues for them which can lead to post-traumatic stress disorder (PTSD) and anxiety. While in-person schools are unavailable, online activities and educational support become very important for

student's development. However, as mentioned in previous sections, a lot of lower SES families don't have access to the internet or devices to participate in these online activities or courses. To accommodate this, the U.S. Congress passed the CARES Act, which included about \$67 billion in flexible funds for K-12 schools to help them reopen and form programs to support students (George et al., 2021). These funds can be used to shorten the educational gaps between high and low SES students by providing access and equipment for online learning, and create accessible clinic programs that focus on social and emotional health recovery for children.

Local schools can help by offering summer learning programs and other activities outside of normal school hours. It can make sure that every student has access to online education, since it is crucial to expand educational opportunities among the community. One way that it can be done to help low SES families is to create community Wifi hotspots, and offer stable Wifi within specific areas in the community to reduce internet cost for families. Families that need internet access can go to the community internet hotspot and their children can also be able to access online education. Other than educational support, mental health services such as online counseling and local clinics can work together to help children cope with stress led by the pandemic. SAMVAD (Support, Advocacy & Mental health interventions for children in Vulnerable circumstances And Distress) is a program that is helpful to address mental health, education, and safety issues for children especially when they face difficult situations during the pandemic (Ashok et al., 2021). This program is useful to help low SES families and their children to keep up with their peers both socially and academically and prevent them from being lost and falling behind.

Aside from school and local government support, parental care and involvement is crucial for helping children to build resilience and stable their mood during the pandemic. A study by Hendry et al. (2023) showed that children who participated in family bonding activities during the first couple months of the pandemic performed better prosocial behaviors after several months after the pandemic. In order to take care of their children's mental well-being, parents were suggested to take approaches from Psychological First Aid (PFA) guidelines to observe and reduce their children's stress and negative mood. It is important for parents to show their children accurate information about COVID-19 situation from trusted sources. It is useful to teach children why quarantine is needed at this time, and how to protect themselves and the people around them (Shah et al., 2020). These supporting information sources served to help lower SES families, to lessen the impact of the pandemic to their family and especially for children's development. By combining strategies from government, school, and family support, they all contribute to address the long-term effect of the pandemic on children's education and promote their physical and mental well-being.

Limitation

A limitation of this literature review is that it does not include studies and participant samples from every country and culture background around the world. It is possible to have outliers and certain countries that are not included while some others are discussed in this review.

Reader's understanding might be limited on how the COVID-19 pandemic has affected children's mental health development in different countries, and within those countries, different SES backgrounds. Additionally, this review mentioned multiple mental health issues that might be resulting from the pandemic, it is difficult to generalize these findings to the broad impact from the pandemic on mental health. Impact to individuals vary broadly depending on personal and cultural experiences. This review might not be able to capture every aspect within the range of mental and developmental challenges that children faced during the pandemic worldwide. Lastly, this review did not specifically address those families that fall between high and low SES.

Conclusion

The COVID-19 pandemic had extremely impacted children's mental development, with disproportionate differences between high and low SES. Children from lower SES families faced greater challenges such as limited educational resources, high stress levels in families, and lack of mental health support. All of them had impacted children's cognitive and emotional development. In contrast, children from high SES families had more opportunities to access resources, which helped to reduce some negative effects and difficulties during the pandemic. These experience differences between SES families pointed out the need for support such as improving access to online learning resources for children. It is also important to provide community resources such as mental health counseling more available to meet the needs of families with children from different SES backgrounds. Efforts need to be made for supporting children's mental development from all aspects particularly during this special period of time. Addressing children from different backgrounds ensures that they are able to overcome the challenges from the pandemic and prepare them for future possible pandemic situations. It is important to ensure children get equal access to educational and healthcare services, which can be supported by local governments and community programs. These are essential to encourage positive development and to build resilience for children. Understanding the importance of children's mental development allows caregivers to create effective solutions for possible future worldwide challenges. More research is needed in the future to focus on designing specific programs to target the needs of children from different SES backgrounds, especially during global crises.

Works Cited

- Ashok, S. S., Ramaswamy, S., Kulkarni, A., Singh, D., Quadros, D., Dzuwichu, K. J., Tikku, K., Rizvi, M., Kandula, P., Tank, P., & Seshadri, S. (2021). Responding to children's mental health, protection and psychosocial care concerns in the COVID crisis. *Asian Journal of Psychiatry*, 65, 102831. <https://doi.org/10.1016/j.ajp.2021.102831>
- Bailey, D. H., Duncan, G. J., Murnane, R. J., & Au-Yeung, N. (2021). Achievement gaps in the wake of COVID-19. *Educational Researcher*, 50, 266–275. <https://doi.org/10.3102/0013189X211011237>
- George, G., Dilworth-Bart, J., & Herringa, R. (2021). Potential socioeconomic effects of the COVID-19 pandemic on neural development, mental health, and K-12 educational achievement. *Policy Insights from the Behavioral and Brain Sciences*, 8(2), 111-118. <https://doi.org/10.1177/23727322211032248>
- González, M., Loose, T., Liz, M., Pérez, M., Rodríguez-Vinçon, J. I., Tomás-Llerena, C., & Vásquez-Echeverría, A. (2022). School readiness losses during the COVID-19 outbreak: A comparison of two cohorts of young children. *Child Development*, 93, 910–924. <https://doi.org/10.1111/cdev.13738>
- Graves K. Y., Nowakowski A. C. H. (2017). Childhood socioeconomic status and stress in late adulthood: A longitudinal approach to measuring allostatic load. *Global Pediatric Health*, 4. <https://doi.org/10.1177/2333794X17744950>
- Hendry, A., Gibson, S. P., Davies, C., McGillion, M., & Gonzalez-Gomez, N. (2023). Toward a dimensional model of risk and protective factors influencing children's early cognitive, social, and emotional development during the COVID-19 pandemic. *Infancy*, 28(1), 158–186. <https://doi.org/10.1111/infa.12495>
- Pfefferbaum, B. (2021). Challenges for child mental health raised by school closure and home confinement during the COVID-19 pandemic. *Current Psychiatry Reports*, 23, 65. <https://doi.org/10.1007/s11920-021-01279-z>
- Saggiaro de Figueiredo, C., Sandre, P. C., Portugal, L. C. L., Mázala-de-Oliveira, T., da Silva Chagas, L., Raony, Í., Ferreira, E. S., Giestal-de-Araujo, E., dos Santos, A. A., & Oliveira-Silva Bomfim, P. (2021). COVID-19 pandemic impact on children and adolescents' mental health: Biological, environmental, and social factors. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 106, 110171. <https://doi.org/10.1016/j.pnpbp.2020.110171>
- Shah, K., Mann, S., Singh, R., Bangar, R., & Kulkarni, R. (2020). Impact of COVID-19 on the mental health of children and adolescents. *Cureus*, 12(8), e10051. <https://doi.org/10.7759/cureus.10051>
- Zimmerman, F. J., & Anderson, N. W. (2021). Association of the timing of school closings and behavioral changes with the evolution of the coronavirus disease 2019 pandemic in the US. *JAMA Pediatrics*, 175(5), 501–509. <https://doi.org/10.1001/jamapediatrics.2020.6371>

The Impact of Fluctuating Housing Prices on Socioeconomic Groups in Singapore By Jungwoon Song

Abstract

This paper examines the fluctuating housing prices in Singapore and their disparate impacts on various socioeconomic groups within the context of the nation's unique housing market. Singapore's diverse demographic landscape necessitates tailored housing policies to address the varying economic capabilities of its residents. While the private property market has experienced volatility, the public housing sector has maintained relative stability, largely due to government intervention. This stability has primarily benefited lower- and middle-income citizens, ensuring their access to affordable housing remains largely intact. In contrast, low-income migrant workers and expatriates have faced significant challenges due to their exclusion from public housing schemes, exacerbating existing inequalities and perpetuating cycles of poverty. Meanwhile, the wealthy have often capitalized on market fluctuations, increasing their wealth through strategic investments in real estate. The paper argues for the necessity of targeted government policies aimed at enhancing housing affordability for all socioeconomic groups. Recommendations include refining existing policies to better support those most affected by current economic conditions, ultimately promoting a more equitable housing landscape in Singapore.

Introduction

The COVID-19 pandemic altered the global housing market, resulting in fluctuating demand, a sudden change in buyer preferences, and shifts in urban living trends. As a city-state known for its robust public housing policies and diverse population, Singapore presents a unique case study on how economic disruptions impact the housing situation of different socioeconomic groups. Singapore's housing market is unique due to the dominance of public housing and the government's close monitoring and regulation of the private housing market. Additionally, with the government's long-term planning, innovative financing, and emphasis on home-ownership, it seems like Singapore's housing market is unshakeable upon first glance. However, the pandemic introduced new challenges, showing shortcomings in the housing policies and exacerbating existing inequalities among citizens, non-citizens, and various income groups. Thus, the purpose of this paper is to discuss how COVID-19 changed the housing market in Singapore, and how housing market fluctuations impacted Singapore's low-income families, middle-income families, migrant workers, and expatriates. The first section of this paper discusses the defining qualities of Singapore's housing market and the government's housing policies. Then, changes during the pandemic will be discussed. The final section analyzes how the fluctuations in the housing market impacted the financial circumstances and quality of life for various socioeconomic groups in Singapore.

Singapore's Public Housing

Singapore's government structure is central to understanding the dynamics of its housing market. Singapore is a unitary parliamentary republic with a strong central government which "plays an extensive and multi-dimensional developmental role in the economy" (Phang). For example, the government owns about 80% of the land and determines how a significant portion of domestic savings are spent (Phang). Thus, the public sector is a driving force behind the country's development, infrastructure, and housing market. The Housing Development Board (HDB), established in 1960, is the cornerstone of Singapore's public housing policy. It is responsible for planning, developing, and managing public housing estates that accommodate over 80% of Singapore's resident population (Housing & Development Board). HDB flats are heavily subsidized by the government, making them affordable for a wide range of citizens, particularly lower- and middle-income households. Because the HDB emphasizes home ownership, not just renting, about 90% of Singaporean residents in HDB flats own their home (HDB).

One of the key mechanisms that support homeownership in Singapore is the Central Provident Fund (CPF). The CPF is a mandatory savings scheme in which both employers and employees contribute a portion of the employee's salary. A significant portion of these savings can be used for housing purposes, including the down payment and monthly mortgage payments for HDB flats. This integration of the CPF with the housing system has been instrumental in achieving high homeownership rates in Singapore. The availability of public housing and high homeownership rates mean that Singaporean citizens who live in public housing are somewhat insulated, but not immune, as I will later discuss, from the volatile shifts in the housing market.

However, some groups, such as citizens in private housing, migrant workers and expatriates living in Singapore, do not have the same cushion. For example, non-citizens cannot buy the government-subsidized public housing units sold by the HDB in Singapore as HDB housing can only be purchased by Singaporean citizens of at least 35 years of age. Because Singapore has such a diverse population, with about 30% being non-citizens, this means that a substantial portion of the population is especially vulnerable to changes in the housing market, especially those as severe as those during COVID-19 (National Population and Talent Division).

COVID-19's impact on Housing Policy

The COVID-19 pandemic had a profound impact on Singapore's housing market, leading to a decrease in private housing prices due to economic uncertainty and a temporary decline in demand. The economy and job market in Singapore were heavily impacted by the COVID-19 pandemic, with GDP contracting by 5.4% in 2020, marking the country's worst recession since independence. Key sectors like tourism, retail, and hospitality were severely disrupted due to global travel restrictions and domestic lockdown measures, leading to widespread job losses and economic downturns (CGTN). In response, the government implemented several measures, including subsidies and interest rate adjustments, to mitigate the effects on homeowners and potential buyers.

For example, the Singaporean government provided cash grants like the Solidarity Payment and the Care and Support package to Singaporean residents over the age of 21. These

grants, of at least \$600 or more depending on taxable income and number of dependents, helped reduce the financial burden during the pandemic (Ministry of Finance, Singapore). The government also reduced installment plans for property loans and worked with banks to make it easier for residents to refinance or reprice investment property loans or pause mortgage payments without repercussions (Monetary Authority of Singapore). Moreover, the Enhanced Housing Grant was expanded to help aspiring homeowners afford their purchases. Finally, tenants in HDB flats were granted rental waivers and rebates for a certain span of time depending on their eligibility for the Rental Relief Fund and the Additional Rental Relief Fund; landlords were also provided support for the impact the waivers had on their income (IRAS).

With these measures in place for residents, public housing in Singapore, mostly the prices of HDB flats, was resilient in the face of the COVID-19 pandemic. In addition to the measures listed above, very instrumental to this was the integrated HDB-CPF framework in sustaining the rate of home ownership regardless of private housing prices. The HDB provides the cheapest possible housing, while the CPF allows citizens to tap into the savings they have accumulated in order to finance such purchases, including down payments and mortgage repayments. This framework was one mechanism that kept the sector of mortgage functional, hence cushioning the homeownership rates from experiencing a sharp drop despite the economic challenges resulting from the pandemic.

However, these changes by the government did not reach every person living in Singapore equally, as people working in harder-hit industries were more affected by the economic downturn than others. Additionally, many benefits were limited to Singaporean citizens, meaning that migrant workers and expatriates were more affected by changes in the housing market, especially those living in private housing.

Impact on Socioeconomic Groups

Public housing in Singapore offers several benefits, including affordability, accessibility, and security of tenure. The same cannot be said of the private housing sector, however. Singapore's private housing market experienced significant price volatility, with sharp increases in some areas and declines in others. For example, the Urban Redevelopment Authority (URA) reported that private residential property prices rose by 2.2% in Q4 2020, despite the economic downturn. Thus, there are also significant challenges for foreigners trying to secure housing in Singapore. Their exclusion from the public housing sector creates a cycle where low-income foreign workers are stuck in temporary, often substandard rental accommodations, unable to build equity or secure long-term housing.

Low-income foreign workers are usually low-skilled laborers with work permits (WPs) and foreign domestic workers (FDW). The employers who have hired the migrant workers usually provide housing for them in private company-owned dormitories. The living conditions in dormitories are often far below acceptable standards, with issues such as inadequate access to basic amenities and insufficient nutrition from catered meals, further exacerbating their hardships. For example, it is not uncommon for 12 to 20 workers to share a single room, often

with poor ventilation and minimal personal space. This leads to uncomfortable living conditions that negatively impact their health and well-being, as seen during the COVID-19 pandemic, when migrant workers accounted for over 90% of Singapore's confirmed cases due to these cramped environments (HOME). In the case that a foreign worker becomes a permanent resident of Singapore, they are eligible to rent a unit in some public housing. However, “the maximum proportion of flats in each building that can be rented out to non-citizens is capped at 11%”, so options are incredibly limited (Yi Fan). Combined with the fact that this demographic’s already precarious was hit harder by the economic downturn during COVID-19, low-income migrant families were the ones who were most likely forced to remain in rental accommodations or downsize, further entrenching their economic struggles (Yi Fan).

Other foreign workers in Singapore, such as those with the Employment Pass (EP) and S Pass (SP) for professionals and skilled workers, were affected differently. These workers must satisfy a higher salary requirement to obtain their visas and therefore are more likely to enter the private housing market (Yi Fan). During COVID-19, companies aiming to cut costs in the face of the economic downturn were forced to reduce salaries or the number of employees, and expatriates often bore the brunt of that (Hayes). Since the Singaporean government was only offering aid to citizens and permanent residents, skilled workers were left to fend for themselves. In the face of such uncertainties, many faced financial instability or chose to return home instead of remaining in Singapore. Their exodus created a decreased demand for higher-end private housing and decreased rental prices as landlords scrambled to replace expatriate tenants with local ones.

Thus, the fluctuating housing market prices actually benefited some groups in Singapore, namely high-income households and investors. Investors, whether local or foreign, found new opportunities in the volatile private market, purchasing properties at lower prices during downturns and selling them at a profit during price surges. For example, in 2022, private property prices in Singapore surged by nearly 30%, so individuals who acquired properties at a reduced cost in the private housing market during 2020 could capitalize on the appreciation of their property values, leading to greater wealth accumulation (“Singapore Real Estate is Soaring”). Those expatriates who left because of the financial difficulties during the pandemic and wish to return to Singapore would find it much more difficult to find affordable housing on the private market now. And middle-income locals will also find it more challenging to enter the private housing market with the current state of affairs.

Conclusion

This paper explored the varying impacts of fluctuating housing prices on different socioeconomic groups within Singapore’s unique market context. The nation's diversity leads to a wide range of economic capabilities, highlighting the need for differentiated housing policies. While private property prices experienced a decline, the public housing sector remained relatively stable, largely due to proactive government intervention. This stability benefited lower- and middle-income citizens who depend on public housing, ensuring their access to affordable

homes was not significantly disrupted. For citizens, robust support from government-subsidized public housing through HDB and CPF schemes has generally shielded them from the most severe market fluctuations. In contrast, low-income migrant workers and middle- to high-income expatriates have faced significant challenges, as they do not have access to public housing. This disparity perpetuates cycles of poverty and underscores deep-rooted inequalities in housing access. To promote equitable housing affordability, targeted government intervention is essential. Policies must be refined to better support low- and middle-income groups, addressing the specific challenges posed by the current economic climate.

Works Cited

- CGTN. "Singapore's GDP Contracts 5.4% in 2020, on Path to Recovery from COVID-19 Pandemic." CGTN, 15 Feb. 2021, <https://news.cgtn.com/news/2021-02-15/Singapore-GDP-contracts-5-4-percent-in-2020-XTrb3qCB4A/index.html>. Accessed 12 Sept. 2024.
- Hayes, Trudy. "Expats in Singapore Fearing for Jobs." Forth Capital Insight, n.d., <https://www.forthcapital.com/articles/expats-in-singapore-fearing-for-jobs>.
- Housing & Development Board. "About Us." *HDB*, <https://www.hdb.gov.sg/cs/infoweb/about-us>. Accessed 28 September 2024.
- Humanitarian Organization for Migration Economics (HOME). "Singapore Must Re-Think How It Treats Migrant Workers." HOME, 9 May 2020, <https://www.home.org.sg/letters-to-the-press/2020/5/9/singapore-must-rethink-how-it-treats-migrant-workers>. Accessed 12 Sept. 2024.
- IRAS. "Rental Relief Framework." *IRAS*, 20 July 2023, <https://www.iras.gov.sg/news-events/singapore-budget/covid-19-support-measures-and-tax-guidance/tax-guidance/for-companies-self-employed-partnerships/rental-relief-framework>.
- Ministry of Finance. "Press Releases." *MOF | Press Releases*, 18 June 2020, <https://www.mof.gov.sg/news-publications/press-releases/budget-2020-care-and-support-cash-to-be-paid-from-18-june-2020>.
- Monetary Authority of Singapore. "Supporting Individuals through COVID-19." *Monetary Authority of Singapore*, 31 July 2024, <https://www.mas.gov.sg/covid-19/supporting-individuals-through-covid-19>.
- National Population & Talent Division. "Overview - Overview." *National Population and Talent Division*, 2023, <https://www.population.gov.sg/our-population/population-trends/overview/>. Accessed 28 September 2024.
- Phang, Sock-Yong. "The Singapore model of housing and the welfare state." *Housing and the New Welfare State*. Routledge, 2016. 15-44.
- "Singapore Real Estate Is Soaring: Check YoY Returns, Rental Income | Can Indians Buy Property in Singapore? DECODED." Times Now, 22 June 2023, <https://www.timesnownews.com/business-economy/markets/singapore-real-estate-is-soaring-check-yoy-returns-rental-income-can-indians-buy-property-in-singapore-decoded-article-99578520>. Accessed 12 Sept. 2024.
- URA. "Release of 4th Quarter 2020 real estate statistics - Singapore." *URA*, 22 January 2021, <https://www.ura.gov.sg/Corporate/Media-Room/Media-Releases/pr21-02>. Accessed 28 September 2024.
- Yi Fan. "The Road to Integration: Post-Migration Experience and Migrant Housing Behavior in Singapore." ResearchGate, Mar. 2021, https://www.researchgate.net/profile/Yi-Fan-9/publication/349984402_The_Road_to_Integration_Post-Migration_Experience_and_Migrant_Housing_Behavior_in_Singapore/link/s/604db421a6fdcccfee7d24aa/The-Road-to-Integration-Post-Migration-Experience-and-Migrant-Housing-Behavior-in-Singapore.pdf. Accessed 12 Sept. 2024.

An Analysis of Housing Affordability in Vancouver By Sophia Huang

Abstract

This research paper addresses the relationship between diverse income groups within all neighbourhoods and housing affordability in Vancouver. By leveraging two comprehensive census and housing datasets, the objective is to predict the housing prices by analyzing which factors influence housing prices. Given the current circumstances of the ongoing housing crisis in Vancouver, this paper provides insights for decision making in real estate. We conducted comprehensive data preprocessing and then applied regression techniques over the demographic and housing features. We found that geographic location has a strong impact on housing prices, with strong correlations between locations and housing costs. Our analysis revealed that some neighbourhoods have more severe housing unaffordability than others, even beyond what is expected according to income distributions.

1. Introduction

Vancouver is known for its high home prices. According to the 2023 housing report from Canada Mortgage and Housing Corporation, the ratio of housing unaffordability in the Greater Vancouver Area is more than 14.1, which ranked the worst in Canada. Understanding the affordability of housing is significant for both economic policy and individual home buying decisions.

Vancouver's housing market has been the subject of extensive research due to its notorious affordability crisis. What we know about housing affordability and pricing in Vancouver is substantial. Studies have shown that Vancouver consistently ranks as one of the least affordable housing markets in North America (Gordon 3). The city's housing prices have far outpaced local income growth, creating a significant affordability gap (Dahms and Ducharme 2). Research has identified several contributing factors, including foreign investment, limited land supply due to geographic constraints, and local zoning policies that restrict housing density (Gordon 5-7).

However, there are still gaps in our understanding of Vancouver's housing market dynamics. While we have a broad picture of the city's affordability crisis, we lack detailed insights into how housing affordability varies across different neighborhoods within Vancouver. Additionally, the relationship between neighborhood-specific factors (such as local amenities, transportation access, and community characteristics) and housing prices is not fully understood. Furthermore, there's limited research on how different income groups are affected by affordability issues in specific areas of the city. These knowledge gaps hinder the development of targeted policies to address housing affordability at a neighborhood level. In this research, we used linear regression and an ordinary least square (OLS) model to predict each neighbourhood's expected housing price and average income. Two datasets were analyzed: Vancouver census data from 2016 and House Sale Data. The census data includes comprehensive demographic numerical data from Vancouver's Open Data Portal, while House Sale Data contains both

numerical and categorical data for us to conduct statistical analysis and make predictions (2016 Census of Population; Ghorbanian).

2. Background

Housing affordability has been continuously monitored by the National Bank of Canada. Based on The National Bank Housing Affordability Report in the first quarter of 2024, the report measures the number of months a median-income household required to save for the minimum cash down payment and mortgage payment as a percentage of income (MPPI) calculated by the monthly mortgage with the assumption of a 25-year amortization period and a 5-year term (Dahms, et al.). They used sophisticated economic and mathematical formulas to obtain their results, including:

1. Mortgage Payment as a Percentage of Income (MPPI): $MPPI = (\text{Monthly Mortgage Payment} / \text{Monthly Median Household Income}) * 100$
2. Minimum Down Payment: Calculated based on current regulations (5% for homes up to \$500,000, 10% for the portion between \$500,000 and \$1 million, and 20% for homes over \$1 million)
3. Qualifying Income: $(\text{Monthly Mortgage Payment} + \text{Property Taxes} + \text{Heating Costs}) / 0.32$
4. Months of Saving Required: $\text{Minimum Down Payment} / (\text{Saving Rate} * \text{Monthly Income})$
5. Standard mortgage payment calculations incorporating current 5-year mortgage rates

The report analyzed different types of dwellings using these formulas along with current data on median house prices, median household incomes, and prevailing mortgage rates. However, the results did not include detailed geographic information within each city.

Another study “Vancouver’s Housing Affordability Crisis: Causes, Consequences and Solutions” analyzes various reasons that lead to the housing crisis in Vancouver, such as foreign investment, the local economy, and ownership patterns of high end homes (Gordon). This paper comprehensively analyzes this problem from a rigorous perspective of economic and social science, with the measure of income discussed. Numerous comparisons of Vancouver and other major cities were presented in this paper. Nevertheless, the paper lacks a discussion of neighbourhood impact on the local housing market.

3. Dataset

1. Census data (2016): This data was gathered from Vancouver’s Open Data Portal based on the 2016 Census Local Area Profiles. This dataset gives a thorough view of demographic indicators for all 22 Vancouver neighbourhoods (2016 Census of Population).

We cleaned the data by handling missing/incomplete values. We also filtered unnecessary features in order to focus on the most impactful ones in this study. The two key features we used in this analysis were:

- **‘Total Population’**: the number of people living in each neighbourhood.
- **‘Income Groups’**: the population count in each income bracket for each neighbourhood.

To simplify the model, we excluded additional components. Other attributes (e.g. education level, age distribution) were omitted from this study since they were not directly related to housing affordability. In addition, we did not split the census data into training and testing sets because it was not used for direct model training. Instead, it was merged with the housing dataset.

2. House Sale Data: The Ghorbanian (2022) dataset consists of house sales in Vancouver ("EDA for House Sale Price in Vancouver."). It contains a total of 1,302 records obtained through various sources. Along with each house sale, it includes information on property characteristics (floor area, year built, lot size), price, and address.

We processed the dataset by removing duplicate/incorrect rows and geocoding the longitude and latitude of house addresses. Key features used in this analysis are:

- **‘FloorArea’**: the total floor area of the house, measured in square feet.
- **‘Year’**: the year the house was built.
- **‘LotSize’**: the lot size, measured in square feet.
- **‘Neighbourhood’**: the neighbourhood the house is situated in.
- **‘Price’**: the sale price of the house, measured in Canadian dollars.
- **Geocoded coordinates**: the longitude and latitude of house addresses.

Our analysis excluded qualitative and information not pertaining to (e.g. number of days on the market, number of bathrooms).

4. Methodology/Models

Overview of experimental approach

The purpose of this research was to predict relationships between housing prices in Vancouver. In particular, we explore relative income and housing affordability. This is conducted using linear regression techniques through ordinary least squares regression. The prediction is based on not only the physical characteristics but also the situational factors of the properties.

Dataset split

Before we feed our dataset into the models, we split the housing data into random training and testing subsets. The purpose is to prevent overfitting and improve its generalizability, where a model performs well on the training data but poorly on the unseen testing data. In this project, 80% of the data was used for training the models, and the remaining 20% for testing.

Geocoding and neighbourhoods

To obtain the geographic and housing information, we first converted the street addresses from the housing data into longitude and latitude coordinates to create a geocoded dataset. Then we merged the geocoded data with the housing data only if both datasets have the same length. The merged dataset should be filtered with only valid latitude values. In this study, the latitude values less than or equal to one are considered invalid. We iterated over each coordinate pair to determine the neighbourhood based on the latitude and longitude. Finally, the average housing prices are calculated for each neighbourhood.

Linear regression model

We used linear regression to model the relationship between price and the house features. First, we fed the training set into the linear regression model, as seen below. Then, we used the model to predict the price using the testing set. Next, we evaluated the model by calculating the mean absolute error (MAE), which is the average absolute difference between the ground truth and the predicted prices. We also employed the Pearson correlation coefficient technique to measure the linear correlation between the predicted and actual prices, with values falling between one and negative one, with zero representing no linear relationship.

Models

$$\text{HousePrice} = \beta_0 + \beta_1 * \text{FloorArea} + \beta_2 * \text{Year} + \beta_3 * \text{LotSize} \quad (1)$$

$$\text{HousePrice} = \beta_0 + \beta_1 * \text{FloorArea} + \beta_2 * \text{Year} + \beta_3 * \text{LotSize} + \beta_i * \text{Neighbourhood}_i \quad (2)$$

* i in {1,22} represents each neighborhood

Ordinary Least Squares

Ordinary least squares is a machine learning method for estimating the parameters of a linear regression model. The underlying methodology of OLS is that it works by minimizing the sum of the squared residuals (difference between the observed and predicted values); the difference between the true values and the model's predicted values.

Coefficients for predictors in the linear models

Other than using a linear regression model to analyze the relationship between the features, we also included the categorical data for neighbourhoods in the OLS model to study if geographic locations affect the correlation. The OLS model generates one coefficient for each variable included in the model.

5. Results and Discussion

Table 1: Model Coefficient Estimates

Model	Coefficient	Coefficient Estimates (95% CI)	p-value
Model 1	β_0	-3.111e+06	0.0001573
	β_1	353.4	1.185e-46
	β_2	1842	1.467e-05
	β_3	91.68	4.583e-15
Model 2	β_0	-6.792e+06	1.837e-33
	β_1	280.5	2.644e-72
	β_2	4233	1.039e-47
	β_3	62.36	7.509e-16

*Neighborhood specific coefficient estimates and p-values from Model 2 can be seen in Appendix

In Table 1, we see that all predictors in Model 1 are statistically significant with housing price, as seen by their p-value $< \alpha=0.05$. However, in Model 2 floor area and year are not statistically significant in predicting housing price. This is because geographic location (neighborhood) plays a more significant role in determining housing prices than the physical characteristics of the house.

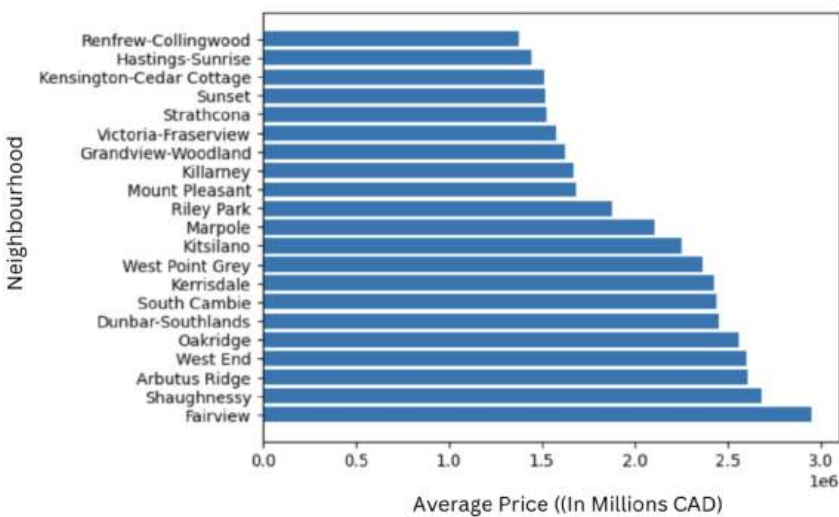


Figure 1: Average Housing Prices for Each Neighbourhood

Figure 1 compares the average housing prices across neighbourhoods. The y-axis is neighbourhoods, listed from lowest to highest average price from top to bottom, and the x-axis shows the average price (in millions). Fairview has the highest housing price (\$2,950,000.00), whereas Renfrew-Collingwood has the lowest (\$1,373,856.77). The range of housing prices is \$1,576,143.23.

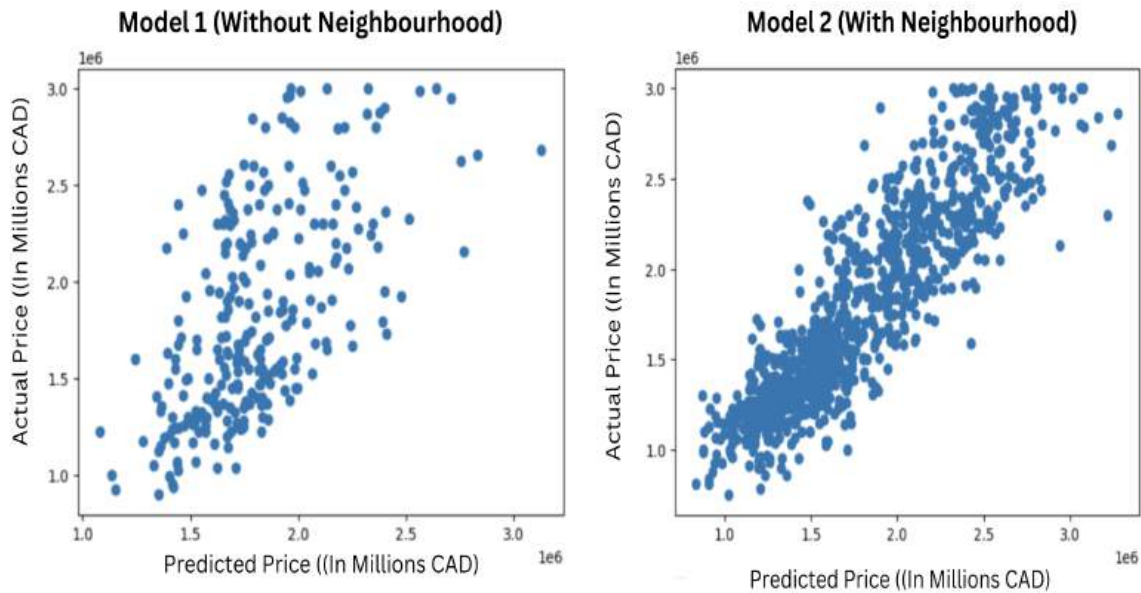


Figure 2: Predicted v.s. Actual Housing

Prices. In the left panel, only non-location variables (lot size, year built, and floor area) are included. In the right panel, all variables (including neighbourhood) are included. Left panel: Model 1; Right panel, Model 2.

In Figure 2, the x-axis shows the predicted housing prices (in millions) derived from the linear regression model based on the selected features, while the y-axis represents the actual prices (in millions) of the houses from the test dataset. These are the true house pricing values that the model predicts.

Ideally, accurate predictions would align the points along a diagonal from the bottom left to the top right, showing a strong positive correlation between predicted and actual prices. According to the above scatter plots, it is obvious that the distribution for Model 1 is much more dispersed compared to Model 2, showing a weaker correlation. Statistically, Model 1 yields an MAE result of 351366 and the Pearson correlation coefficient of 0.61.

The significant deviations on the non-location features panel (Model 1) suggest discrepancies and areas for Model improvement, indicating that factors beyond the three influence house prices. The graph for Model 2 has a smaller MAE value of 188933, compared to

Model 1 (MAE value of 351366), and a stronger positive correlation, showing how much geographic locations affect housing prices.

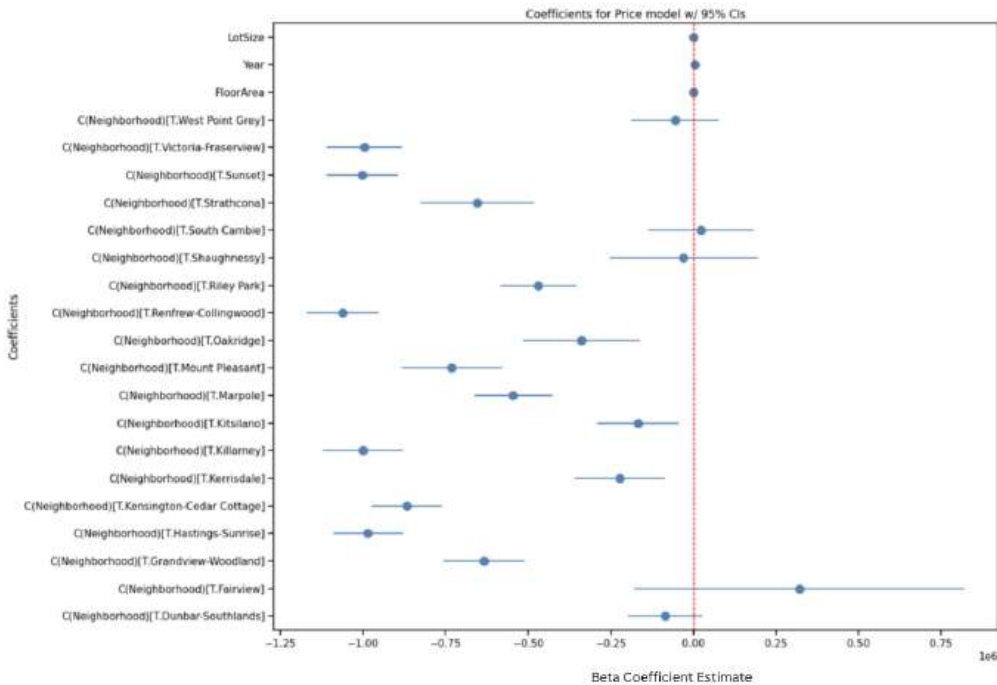


Figure 3: Coefficient plot of each variable in Model 2, with 95% confidence intervals

In Figure 3, Model 2 coefficients to the right of the zero line (dotted red line) indicate a positive effect (increasing prices compared to the baseline), while those to the left indicate a negative effect (decreasing prices). The graphs include error bars representing the 95% confidence interval (CI) of the coefficient. We used the 95% CI as it provides a measure of the uncertainty or reliability of an estimate. Using these confidence intervals, we determined the statistical significance of including each feature in the model. The shorter the bar is, the smaller the CI, therefore the more accurate the prediction is.

This graph displays the estimated effects of including various neighbourhoods on the overall Vancouver house price, compared to the baseline neighbourhood (Arbutus Ridge). It's useful for identifying areas that might be undervalued or overvalued based on neighbourhood effects alone. This information can guide investment decisions, policy making, and further economic analysis of housing markets.

The graphs include error bars representing the 95% confidence interval (CI) of the coefficient. We use these confidence intervals to determine the statistical significance of including each feature in the model.

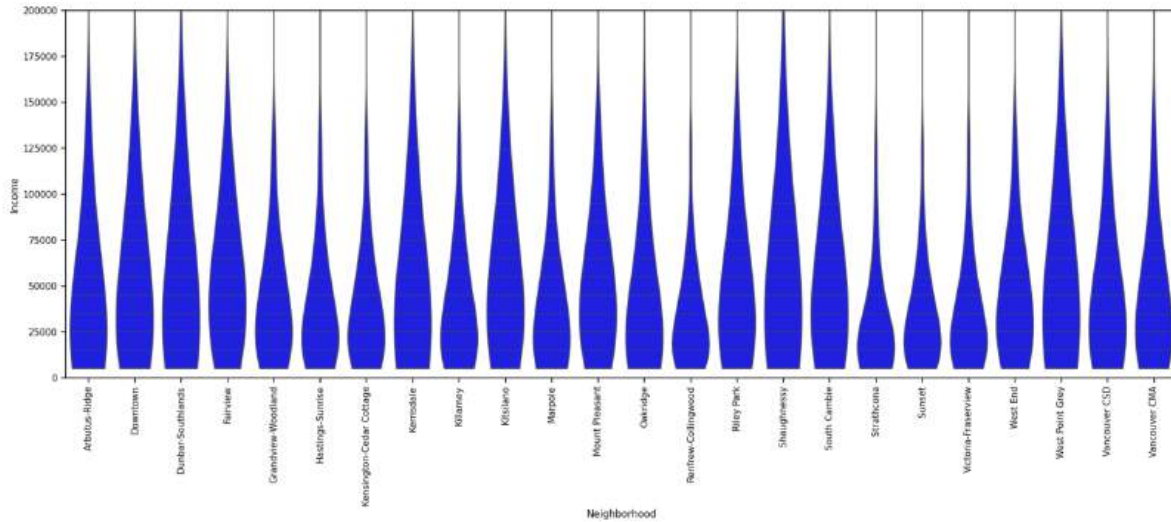
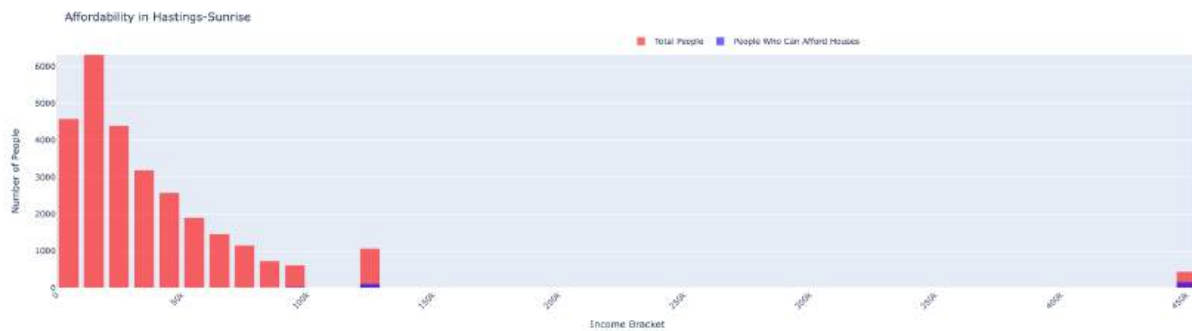


Figure 4: Income Distribution in Neighbourhoods

Figure 4, a violin plot illustrates the distribution of incomes across various neighbourhoods. Note that the scale shows income from 0 to slightly over 400,000. Each "violin" shows the density of the data at different income levels, with the width of the shape indicating the concentration of residents. A wider section means more people earn around that income level in the neighbourhood. The graph offers a clear depiction of income distributions across neighbourhoods, highlighting economic diversity and the range of incomes present in each area. This figure reveals that Shaughnessy and West Point Grey are relatively economically homogeneous neighbourhoods, whereas Strathcona and Sunset have more unequal distributions. This information can help targeted socio-economic planning and policy development.



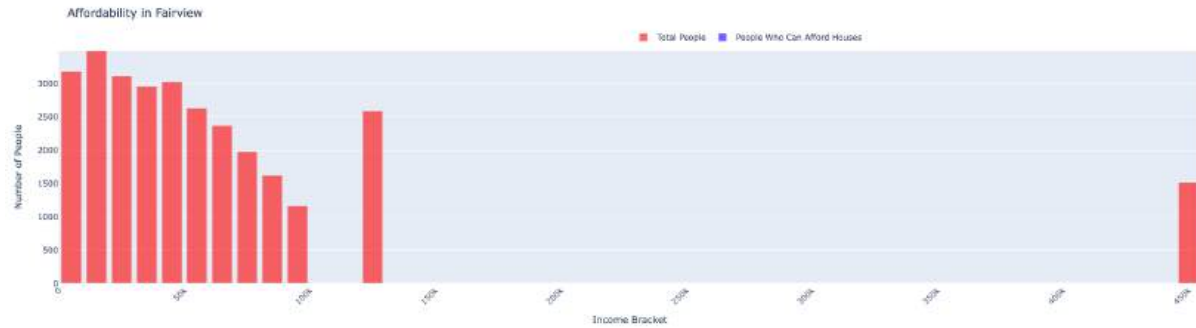


Figure 5: Average Affordability for Each Neighbourhood

The x-axis represents the income bracket and the y-axis represents the number of people within this neighbourhood. According to Figure 5 and our previous Figure 1, it is obvious that houses in the Hasting-Sunrise neighbourhood are more affordable than the houses in the most expensive neighbourhood, Fairview, as seen by the proportion of people that can afford Hasting-Sunrise but not Fairview. People with an income of \$125k can buy a property in Hasting-Sunrise; however, a \$450k income family is not even guaranteed to afford a house in Fairview.

We can make connections between the housing affordability landscape in Vancouver by looking at the east-west neighborhood divide. Eastern neighbourhoods like Hastings-Sunrise and Renfrew-Collingwood offer more affordable options, while western areas such as Fairview, Shaughnessy, and West Point Grey remain prohibitively expensive for many families despite their higher-income residents. This pattern highlights how neighbourhood desirability and investment trends significantly influence housing disparities across the city.

The Neighbourhood Price Map reveals that eastern neighbourhoods like Renfrew-Collingwood and Hastings-Sunrise are more affordable, while western neighbourhoods such as West Point Grey and Shaughnessy are more expensive. According to the Average Prices Bar Chart, the highest average prices are found in Fairview, Shaughnessy, and Arbutus Ridge, with homes costing around \$2.5 million. In contrast, Renfrew-Collingwood, Hastings-Sunrise, and Kensington-Cedar Cottage offer more affordable options at approximately \$1 million. The Price Predictions scatter plot indicates a moderate correlation between predicted and actual housing prices, though significant deviations suggest that additional variables are needed to enhance the model. The Coefficient Plot highlights that areas like West Point Grey and Shaughnessy significantly raise housing prices, while Renfrew-Collingwood and Hastings-Sunrise lower them, indicating greater affordability. The Income Distribution data shows that wealthier neighbourhoods like Shaughnessy and West Point Grey have higher income levels, whereas Renfrew-Collingwood and Hastings-Sunrise are economically diverse. In summary, there is a noticeable disparity in housing prices and income levels across these neighbourhoods, with more expensive areas having higher incomes and less economic diversity.

6. Discussion

While western neighbourhoods like West Point Grey and Shaunessy hold significantly higher prices, eastern neighbourhoods such as Renfrew-Collingwood and Hastings-Sunrise are much more affordable. Thus, our findings demonstrate a strong relationship between location and price of Vancouver houses. In addition, disparities across neighbourhoods are not only present in income distribution, but also neighbourhood desirability and investment patterns. The inclusion of both physical and situation factors of the houses provide a nuanced analysis.

A limitation in the use of my datasets is that some neighbourhoods did not have a surplus of observations, which could have led to inaccurate representations of those neighbourhoods.

7. Conclusion

The study's goal was to analyze housing affordability across Vancouver's neighbourhoods by examining income distribution and housing prices. Two datasets were used: the 2016 Vancouver census and Vancouver house sale. Using linear regression technique (including ordinary least squares regression), we predicted housing prices with various factors. Including geographical data in our model, categorizing each house by their neighbourhood, significantly improved the accuracy of housing price predictions. This confirms the strong influence location has on housing prices in Vancouver.

Ultimately, our results reveal significant discrepancy in housing affordability across neighbourhoods. Neighbourhoods like Fairview and Shaunessy exhibit notably higher average house prices compared to Renfrew-Collingwood, a neighbourhood with relatively affordable options. In addition, we observed income distribution patterns, where expensive neighbourhoods have higher incomes and less economic diversity. However, when considering these results, it is important to note that some neighborhoods (Downtown and Fairview) lacked a substantial amount of housing samples, thus not fully representative of the neighbourhood's housing price.

Nonetheless, this study provides a nuanced understanding of Vancouver's housing affordability crisis at a neighbourhood-level. By providing a detailed picture of variations across the city, it sets a foundation for targeted and effective approaches to address this urban challenge. Future research could expand on this by incorporating other factors such as differences in housing policies as well as access to transportation and amenities. A longitudinal study could also provide insights into how neighborhood affordability trends evolve over time.

Appendix

Model	Coefficient	Coefficient Estimate (95% CI)	p-value ($\alpha=0.05$)
Model 1	β_0	-3.111e+06	0.0001573
	β_1	353.4	1.185e-46
	β_2	1842	1.467e-05
	β_3	91.68	4.583e-15
Model 2	β_0	-6.792e+06	1.837e-33
	β_1	280.5	2.644e-72
	β_2	4233	1.039e-47
	β_3	62.36	7.509e-16
	β_4	-8.535e+04	0.1387
	β_5	3.213e+05	0.208
	β_6	-6.334e+05	3.482e-23
	β_7	-9.847e+05	4.49e-65
	β_8	-8.669e+05	1.228e-51
	β_9	-2.222e+05	0.001603
	β_{10}	-1e+06	9.877e-54
	β_{11}	-1.676e+05	0.00787
	β_{12}	-5.453e+05	5.273e-19
	β_{13}	-7.312e+05	3.119e-20
	β_{14}	-3.386e+05	0.0002069
	β_{15}	-1.061e+06	3.004e-70
	β_{16}	-4.696e+05	3.245e-15
	β_{17}	-2.998e+04	0.7925
	β_{18}	2.294e+04	0.7775
	β_{19}	-6.539e+05	2.092e-13
	β_{20}	-1.002e+06	8.656e-63
	β_{21}	-9.956e+05	7.013e-60
β_{22}	-5.516e+04	0.4124	

Works Cited

- Dahms, Kyle, and Alexandra Ducharme. Housing Affordability Monitor. National Bank of Canada, Financial Markets, 2024,
<https://www.nbc.ca/content/dam/bnc/taux-analyses/analyse-eco/logement/housing-affordability.pdf>
- Ghorbanian, Darian. "Vancouver Home Price Analysis - Regression." Kaggle, 2022,
www.kaggle.com/datasets/darianghorbanian/vancouver-home-price-analysis-regression/data.
- Gordon, Josh. Vancouver's Housing Affordability Crisis: Causes, Consequences and Solutions. Simon Fraser University School of Public Policy, Center for Public Policy Research, 2016.
- "EDA for House Sale Price in Vancouver." Kaggle, 2023,
<https://www.kaggle.com/code/jake021/eda-for-house-sale-price-in-vancouver/notebook>.
- 2016 Census of Population, Statistics Canada Catalogue no. 98-316-X2016001. Vancouver [CMA], British Columbia. Government of Canada, 2016,
<https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/page.cfm?Lang=E&Geo1=CMACA&Code1=933&Geo2=PR&Code2=59&SearchText=vancouver&SearchType=Begin&SearchPR=01&B1=All&type=0>.

Energy Justice in Global Climate Change Policy By Saral Gupta

Abstract

In a post-COP28 world, questions regarding the role of global cooperation in climate change policy have grown ever more crucial. Fears about the increasing urgency and danger of climate change on the globe have caused many to take the perspective that all nations should be held accountable for carbon emissions, no matter the differences between nations. However, while this claim is correct in its assessment of the urgency of climate change action, it does not adequately address inequities between developing and developed nations. This inequity represents long-standing economic and social conflicts between members of the developed and developing world, and these conflicts have caused long-standing differences between the developed and developing world that make it impossible for the 2 sides to adopt an identical climate change policy to each other. Here, we argue that these inequities should be factored into climate change policy by allowing industrializing nations carbon space to industrialize, and to put greater pressure on the nations of the West in their climate change action.

The effectiveness of industrialization in eliminating poverty

One of a nation's main concerns when looking at effective energy sources for their populations is the effect they have on mitigating poverty. Poverty has come to be a significant issue nations must face as they develop. The phase at which countries decide to start eliminating this problem is when greater economic goals are being set to hopefully turn a developing nation into one that is developed. However, most world leaders who set out to eliminate poverty on a nation-wide scale quickly realize its unattainability without technological advancement.

Through analyzing global historical data, a decrease in extreme poverty on a nationwide scale has shown to be in large part caused by the technological innovation of the Industrial Revolution (Ravallion 2016). The creation of new jobs and cheaper goods through productivity growth & capital accumulation for businesses, and the creation of new industries alongside technological innovation has historically been displayed as a must for eliminating mass poverty on a nationwide scale. Economists have long associated innovations in advanced sources of energy alongside the growth of the quantity of energy to be essential in economic growth and technological advancement (e.g. Cleveland et al., 1984; Ayres and Warr, 2005, 2009, Hall et al., 2003), and so eliminating poverty in developing nations requires advanced energy sources such as either Renewable Energy Technologies (RETs) or advanced fossil fuels for technological advancement and economic growth. Therefore, advanced energy use in all nations is intrinsically tied to the developmental phase in which that nation currently occupies, and so to not only ensure compliance with climate change policy but to also combat inequalities between nations at different developmental phases, climate change policymakers must consider the necessity of advanced energy sources in industrialization.

While it has been established that some form of advanced energy sources are necessary in developing regions, empirical evidence and decades of data have shown that there are a wide

variety of issues adopting RETs in industrializing regions of the globe. Industrializing regions such as China, the Indian subcontinent, and Eastern Africa suffer harsh penalties for adopting RETs over fossil fuels that present an unfair trade-off in productivity and effective industrialization, which presents the clear idea that despite the environmental consequences, an industrializing region's only feasible, efficient, and sustainable option for economic advancement is the use of carbon-based fossil fuels.

Issues Adopting RETs in Eastern Africa

As a major sector of industrialization in the world, the possibilities of implementing RETs into East Africa's industrialization process present a remarkable case study into the possibilities of implementing these technologies across the entire developing world. As a region with some of the poorest nations in the world yet conversely some of the highest amounts of natural resources in the world, shifting from agriculture to manufacturing is a necessity in East Africa that would create massive amounts of jobs, reduce poverty, and create rapid sustained growth in the region that is so desperately needed. (McMillan and Albert Zeufack 2022). However, this industrialization process has only been shown to be hampered by attempts to use RETs as a major power source for this industry change.

RETs suffer from a wide variety of issues in the region, not least of which is the lack of technical skills in their implementation. Implementation of RETs requires advanced technical knowledge that is lacking in East Africa due to a continuing shortage of qualified and educated personnel (Baguant and Manrakhan, 1994). In most of East Africa, the operation of energy sources is largely mechanical. This means that electrical energy technologies such as RETs are difficult to grasp for users in the informal and many formal economic sectors, as well as end users, contributing to an overall inefficiency and unsustainability in the use of RETs as a major power source for industrialization (Karekezi 2003). Additionally, many of the individuals with knowledge of operating these technologies reside in mostly urban areas, making RET use in East Africa inefficient and ineffective. Integrating renewable energy in East Africa also faces infrastructure issues, as regions with limited technical expertise and infrastructure often make RETs unscalable and inefficient for their total use (IEA, 2021).

These issues are only further exacerbated when looking at the benefits of using fossil fuel-based energy. The current economic and logistical realities in East Africa mean that the current energy infrastructure is better suited for fossil fuel-based power generation. Coal, for instance, is seen as a more reliable and scalable energy source, especially given the existing infrastructure and the region's pressing need for consistent and high-output energy to support industrial activities (IRENA, 2022). Furthermore, the reliability of coal and other fossil fuels ensures that energy demands can be met without the interruptions and inefficiencies that currently plague renewable energy systems in the region (Eberhard et al., 2022).

Not only is the concept of RET use in industrializing nations such as East Africa flawed from a perspective of efficiency and usability, but it is also flawed from the perspective of energy justice. These nations in East Africa benefit from a plethora of resources such as fossil fuels

found in their borders, and losing the ability to use these resources negatively affects these nations in an immoral way. Denying these nations the opportunities to exploit their fossil fuel resources presents a perpetuation of the current state of poverty that these nations are in, and presents a form of energy injustice that needs to be addressed (Sovacool et al., 2020). Sovacool et al. (2020) emphasize that African nations should have the autonomy to use their fossil fuel resources to combat energy poverty and support industrialization. This is particularly crucial as the transition to renewable energy must be balanced with the immediate developmental needs of the continent (Eberhard et al., 2017), and the growing poverty in East African nations presents a powerful call to the inherent morality in allowing for fossil fuel use/allotting carbon space in developing regions such as East Africa.

Issues adopting RETs in India

As a developing nation with the 5th largest economy and yet debilitating issues surrounding poverty, India represents a significant middle ground and example to the argument of RET use in developing nations. Despite its status as the 5th largest economy, about 60% of India's nearly 1.3 billion people live on less than \$3.10 a day, the World Bank's median poverty line (Quintero 2019), and 57% of its rural areas are energy impoverished, with little access to electricity, running water, or clean gas (Khandker et al., 2010). This contrast in India presents a powerful energy crisis that has remained one of the top priorities of the Indian government. If, as many claim, RET use in developing countries was economically feasible, it would be the most significant potential solution to the poverty and energy crisis in India. However, renewable energy has a multitude of drawbacks in developing nations, including uneven development, high cost of integration, and lack of infrastructure to support these technologies. By looking at these problems through the lens of attempted implementation in India, the necessity of carbon space and fossil fuel use becomes much clearer on a global scale.

One of these issues remains the extremely high cost of RET integration into India's grid infrastructure. The uneven development across regions and the vast distances between renewable energy production sites and demand centers in India means that the development of necessary infrastructure, such as transmission lines to connect renewable energy sources in remote areas to urban centers, is a significant financial burden (Bhattacharya 2019) that a nation in its developing stages cannot afford, and is unfairly positioned against cheaper fossil fuel alternatives. In addition to this high cost of integration is frequent curtailment of India's renewable energy due to the lack of adequate storage solutions and grid flexibility, which are critical for effectively integrating intermittent renewable energy sources into the grid (IEA 2021). This means that India, alongside other developing nations, lacks the infrastructure to efficiently use renewable energy compared to fossil fuel alternatives, the lack of adequate storage solutions, and grid flexibility, which are critical for effectively integrating energy sources like solar and wind into the grid. Solar and wind energy in particular are also extremely intermittent in their power production, as well as being extremely unscalable. Singh and Jamasb (2020) highlight that the current storage technologies in India are not yet developed enough to handle large-scale

renewable integration, and the lack of affordable and scalable energy storage solutions means that India must rely on conventional power sources to ensure a stable and reliable energy supply. These issues are exacerbated when compared to fossil fuel alternatives. While India has made strides in renewable energy adoption, the economic and policy frameworks are still heavily skewed towards coal and other fossil fuels, which continue to dominate the energy mix due to their lower costs and higher reliability (World Bank 2021). This is due in large part to the high initial capital costs associated with renewable energy projects due to a lack of infrastructure to sustainably develop RETs, especially in poorer areas. Rural and poorer areas of India, in particular, struggle with the high costs of renewable energy systems because of similar reasons as East Africa. Poorer regions in developing nations such as India continue to suffer from a lack of infrastructure, lack of technical expertise, and limited access to finance these renewable energy programs (Chaudhary et. al., 2021), claiming that renewable energy should be the foundation of every nation's energy systems, a claim that is unattainable and unjust to the impoverished in these nations.

Issues Adopting RETs in China

As the world leader in adopting renewable energy, China has put heavy investment into ensuring its various energy needs are met by sources such as hydropower, wind power, and PV (photovoltaic) technology. However, while these energy sources have been effective at curbing much of China's pollution power, their very nature conflicts with China's status as an industrializing nation, leading to various issues arising from their adoption. The utilization cost of RETs such as wind power and PV technology remains extremely high due to issues surrounding curtailment and infrastructure construction prices (Zhao 2022). China's status as an industrializing nation with uneven development across its borders means that building the ultra-high-voltage transmission lines needed to connect the renewable energy developing regions of north and northwest China to the energy-dependent regions of east and south-east China becomes untenably expensive (Chen 2024). This expensive procedure leads to frequent bottlenecks of renewable energy and high levels of curtailment of renewable power by Chinese companies, leading to extremely inefficient renewable energy use caused by the problems associated with China's developmental stage. These issues are further worsened when comparing the efficiency of China's renewable energy industries to its coal and fossil fuel industries. Coal is much cheaper in terms of operational costs than renewable energy, which is incredibly important for China's growing economy. A large reason why this remains so is because China has a well-established coal infrastructure, including mining, transportation, and power plants. Transitioning to renewable energy requires significant changes to the existing infrastructure, which is both time-consuming and costly in a way that a nation still reliant on producing goods for economic gain cannot afford (Zaremba 2024). These inefficiencies and the superiority of coal are best shown in the installment capacity between the 2 industries. The gap between the installed capacity of China's coal industries in producing electricity versus China's renewable energy continues to make coal the preferred choice in electricity generation (Chen 2024). This

installed capacity gap means that 70% of China's electricity is still generated by coal. This coal production remains incredibly important, especially in China's poorer Western regions, which due to China's status as an industrializing nation, remain incredibly poor and underdeveloped. In these regions, the costs of implementing RETs as an electrical source are much too high to implement due to the bottlenecks, added costs, and overall inefficiency of China's RET industry. In these regions especially, energy security means coal use, and for China to emerge from its industrializing phase to a modern nation in all of its regions, coal use is still necessary (Hilton 2024).

These large issues adopting renewable energy in China underscore significant trends that emerge in all developing regions when attempting to adopt RETs as a source of energy needs. As the leader in overall renewable energy use in the globe, as well as the nation that has put the most investment into renewable energy use (IEA 2022) through a centralized government with large control over its economy, China above all developing nations is in the best position to adopt a majority renewable energy economy over others. However, even with these advantages, China is still plagued by extreme issues that make adopting renewable energy extremely detrimental to eliminating poverty and emerging from its developmental cocoon. If this is the case for a country in the prime position to adopt renewable energy, then it is safe to assume that forcing developing nations in worse conditions than China to adopt renewable energy would be catastrophic for their overall health and security. The technical, economic, and policy-related obstacles that China faces would be even more pronounced in less developed countries, where the infrastructure and economic stability required to support such a transition are often lacking (Hu et. al 2020) due to the high startup costs for infrastructure combined with the need for cheap energy to promote economic growth as shown in the Industrial Revolution. Thus, it is highly unfair and unwise to promote 100% or majority renewable energy projects in these nations, and major renewable energy policymakers should instead focus on allowing these nations "carbon space" that could instead be made up of developed nations in a prime position to adopt RETs.

Overall Trends Displayed from these Examples

Looking at the similarities in RET adoption between these 3 developing regions of the globe helps conclude renewable energy use in developing nations as a whole. In all of the areas analyzed, a clear trend that emerged was that no matter the cohesiveness of the region's economy, or the potential investment put into renewable energy, RETs were simply not feasible due to a lack of infrastructure. As nations still in the industrializing phase, developing nations simply do not have the technological buildup to have adequate infrastructure for renewable energy, thus making renewable energy an inefficient and insufficient main power source for industrialization. These RETs are also simply much too expensive to all implement at once. Transitioning a nation from carbon-based energy sources to renewable ones requires a slow buildup of RET infrastructure, one that requires at least a period of carbon space. The second trend that emerged was that even when using renewable energy, that energy was ineffective at providing power to poorer areas of these developing regions and that this energy was ineffective

at lifting those areas out of poverty. RETs do not function as a sufficient advanced energy source to lift an area out of poverty because poorer areas often lack adequate financing for expensive RET buildup, and lack the technical knowledge to adequately operate these technologies consistently. These factors make renewable energy not nearly as good of an option for advanced energy use as fossil fuels for developing countries and display the need for allotting carbon space for developing nations as a whole.

The efficiency of Renewable Energy Technology use in developed nations

In contrast to developing nations, post-industrial societies such as those of North America and Western Europe have been presented with a clear advantage and benefit to adopting RETs. Not only are these regions more able to implement these energy sources into their economies, but they also are presented with clear benefits of economic and energy independence from the unstable economies from which a large amount of carbon-based energy sources are produced. Because of these reasons, ensuring the adoption of RETs in post-industrial nations should be the top priority of modern climate change policy, and presents a clear balancing force to the necessary increase in carbon emissions from industrializing nations.

Feasibility of the EU to Adopt Renewable Energy Technologies (RETs)

The EU's capacity to integrate innovations in renewable power technologies is supported by a favorable legal and political environment, adequate funds and financial commitment, and a multi-stakeholder approach. The recurrent revision of the RED II for 2023 has set a binding target for renewable energy which is set to account for 42 percent. At least 5% of the total EU energy consumption through renewable energy by the end of 2030 with a 45% target aim (European Commission 2023). This audacious goal is supported by comprehensive enabling policies that have cut across the permitting uncertainties thereby making it easier and clearer for renewable energy projects ("Renewable Energy Policies"). The Just Transition Fund is essential while the Innovation Fund assists, and the European Investment Bank is a very significant financier for renewable energy projects (European Investment Bank 2022). The large-scale investments on renewable energy infrastructure and research is attributed to the financial capacity of the EU economy. This financial capacity has provided for big funding of projects such as Horizon Europe that aims at promoting innovation in things like the energy sector. Having capital from the public and private sectors means that research, innovation as well as implementation of valid and credible projects can be fully accomplished, and thus the EU is one of the leading regions in the development of renewable energy sources (International Renewable Energy Agency 2022). Also, the EU seeks to invest a lot of money in research and innovation, for instance through the Horizon Europe programs that sponsor the evolution of advanced energy technologies (European Commission 2022). This approach does guarantee that the usage of renewable energy impacts many sectors such as power, transport, heating, and cooling thus decomposing the overall security and sustainability of energy in the EU (Eurostat 2023). The embracement of renewable energy in public and private sectors is boosted by public-private

partnerships, which equally involve expertise, resources, and risks among others as per the International Energy Agency.

The ability of the US to Adopt Renewable Energy Technologies (RETs)

The adoption path of renewable energy technologies in the United States has shown a high potential with federal and state support and demand for the technologies along with technological development. The passed Inflation Reduction Act of 2022 presents massive incentives for renewable energy sources where they proposed to extend tax credit facilities for solar and wind projects (U. S. Department of Energy 2023). Renewable energy is backed by good federal programs and projects aimed at enhancing this resource with the DOE at the helm of funding, research, and development (National Renewable Energy Laboratory 2022). Both state and federal policies are important here; for example, California's Renewable Portfolio Standard (RPS), which requires that 60% of the state's electricity come from renewables by 2030; or Texas, which leads the way in wind energy as a result of favorable policies and geological conditions (California Energy Commission 2023; Texas Renewable Energy Industry Association 2023). The constantly developing and highly competitive structure of the energy market in the US spurs investment and innovation: REC and PPAs as market instruments provide for the structuring and financing of renewable energy projects (Lawrence Berkeley National Laboratory 2022). Given the fact that the US is a technological powerhouse, especially in technological development, the country is highly developed in research on renewable energy technologies with the national laboratories, universities, and private companies constantly enhancing the new technologies through more research (MIT Energy Initiative 2022). The U. S. economy which has significant financial resources is also particularly important to fund the expansion of renewable energy. They are healthy economically thereby enabling the provision of large fiscal support in the framework of infrastructural developments and in addition encourages innovative advancements through research (U.S. Energy Information Administration 2023). This increases the possibility of large scale investments in renewable power projects whose development would be propelled by the availability of capital (U. S. Energy Information Administration 2023). High public concern and approval extend to environmental conservation and thus the promotion of renewable energy through influencing policies and funding to the renewable energy sector (Kennedy et al. 2023).

General Advantages Developed Nations Have in Adopting RETs

Developed nations possess several inherent advantages that facilitate the adoption of renewable energy technologies. Higher levels of economic development enable substantial investments in renewable energy projects, with the financial capacity to fund large-scale infrastructure and advanced research initiatives (World Bank 2021). Advanced technological capabilities and robust infrastructure support the integration of renewable energy, including well-established grid infrastructure, advanced manufacturing capabilities, and a skilled workforce (Organisation for Economic Co-operation and Development 2023). Strong regulatory

frameworks and supportive policies create a conducive environment for renewable energy adoption, providing stability and predictability for investors and developers (International Energy Agency 2022). Established markets in developed nations attract investments and foster competition, driving innovation and efficiency in renewable energy technologies (BloombergNEF 2023). Higher levels of public awareness and support for environmental issues enhance the political and social acceptance of renewable energy projects, with public demand for sustainable energy solutions influencing policy decisions and encouraging the development of renewable energy infrastructure (United Nations Environment Programme 2023).

Economic Strength of RETs in Developed Nations

Renewable energy technologies in developed countries are economically important, providing new jobs, economic development, and security of supply. According to Eurostat, the renewable energy sector is the biggest part of the green economy in the European Union in terms of job creation and economic growth. This focus on renewable energy forms is also in line with the EU's long-term aim of attaining climate neutrality by 2050, with these projections also likely to unlock more economic prospects and market development (European Commission 2023). The renewable energy industry is among the first-growing industries in the United States as evidenced by thousands of new job slots created and billions of investment dollars (U. S. Department of Energy 2023). Federal and state policies give the appropriate motivations to renewable energy projects that give the proper long-term financial stability and energy security (Brookings Institution 2023). Renewables are also part of the economy which drives new industries, supports innovation, and delivers more stable and secure energy to the economy (American Council on Renewable Energy 2023).

Addressing Historical Imbalances and Inequities in Climate Change Policy

The developed countries have several initial conditions to help adopt renewable energy technologies, which include the following. Economic development also offers enhancement of capital to finance more and bigger renewable energy system infrastructures and research (World Bank 2021). There is an established grid connection, high technology capacity for manufacturing equipment, and a skilled workforce in the nation to enable the integration of renewable energy (Organisation for Economic Co-operation and Development 203). Nonetheless, comprehensive and effective regulatory measures make the deployment of renewables conducive and safe from policy interference, legal uncertainties, and political risks for investors and operators (International Energy Agency 2022). Affordable solar products in developed countries such as the USA receive investment and also encourage competition in renewable energy technologies (BloombergNEF 2023). Larger public concern and appreciation of environmental concerns improves the political and social acceptability of renewable energy projects, and the public pressure on power producers fuels policy-making on sustainable energy provision and creating renewable energy systems probably beyond possible government policy frameworks (United Nations Environment Programme 2023).

Comparison

When considering the possibilities of whether or not developing nations should adopt for growth while simultaneously being hindered by the necessity of keeping a low carbon profile, a comparison must be made in-depth. This comparison allows for the convergence of various ideals that stem from a country's fundamental issues with RETs to discover a blueprint for ideal carbon modernization.

Technologies or systems in its industrialization process, with special emphasis on the affordability crisis. As for East Africa, it is endowed with vast natural resources but to transit from the agriculture sector to the manufacturing sector we need reliable energy sources which could not be provided by RETs. One of the most significant challenges is the absence of specialized knowledge in the adoption and management of these sophisticated technologies and the comparatively poor development of infrastructure across the area. Over eighty percent of the energy operations in most East African countries are mechanical, so shifting to electrical technologies such as RETs is especially difficult. This has been compounded by the fact that qualified staff tend to work in urban areas leaving RETs as institutions in rural and other poor economic areas ineffective. Therefore, the existing energy structures mostly correspond to the requirements of fossil energy sources that are considered more reliable and easily scalable, responding to the imperative needs of the East African region in stable, high-power energy for industrial development.

However, China, the ultimate developing nation of renewable energy, does experience some inefficiencies in terms of expense to incorporate RETs into the energy integration because the country is in an industrially developing State. There are high costs involved in the process of building infrastructure such as it is costly to build the Ultra High Voltage Direct Current transmission system that is needed to transport renewable electricity production in northern and northwest regions to the energy-starved east and southeast regions. These logistic complications result in numerous opportunities for congestion and high levels of curtailment in the renewable energy sector in China; thus, a dysfunctional system is created, which further emulates the necessity for more affordable energy options. Coal still appears to be a lot cheaper and more dependable as a result of the transport, mining, and generation infrastructure that the Chinese possess. This cost advantage of coal is important to China's growing economy given that energy security is of considerable importance in the region. Thus while China has embarked on a massive drive for the development of renewable energy, the country still depends on coal, especially in the impoverished regions of the western part of the nation where the adoption of RETs is impossible due to their high costs and technical complexity.

Poverty is one of the biggest issues that face India which makes it difficult for the nation to embrace Renewable Energy Technologies (RETs) as a solution to energy problems. India is the fifth largest economy in the world, but the poverty index is still high in the country and the energy poverty index is particularly high in rural areas of the country. The main challenge of integrating RETs into India's load is that the cost of developing these technologies in remote and developing areas is very expensive. This is further compounded by the problem of storage and

flexibility in the grid; two critical factors needed to balance intermittent renewable power sources such as solar and wind power. As a result, presently also the demand for fossil fuels particularly coal remains high in India due to their low costs and high availability. Such a dependency on traditional energy sources reveals the existing imbalance in the development of power infrastructure in the country where the deprived regions cannot afford the initial costs of investment in renewable energy projects. Consequently, the assumption that RETs have to be the cornerstone of energy infrastructures of any country for it to be sustainable just seems unreachable, especially for the suffering populace in a third-world country like India.

The economically and technologically advanced nations of the world especially those in North America and Western Europe can foster the adoption of Renewable Energy Technologies (RETs). These nations reap numerous advantages from RETs which include; a boost of energy independence out of geopolitically sensitive regions that offer fossil fuels and the creation of business opportunities in a green economy and job opportunities. The EU alone is an excellent example; It has established strict plans with effective policies and a robust framework to support renewable energy sources and has provided more than sufficient funding for these goals and programs on renewables. Likewise, the use of federal and state incentives has been provided where America, for example, utilizes the Inflation Reduction Act of 2022 to prolong the tax credits for sustainable energy projects, thereby promoting change towards sustainability. These goals are supported by advanced technology from continuing research and development efforts at national laboratories and universities in the country towards renewable energy. However, as seen with the benefits outlined and the laid down roadmap for implementation of RETs there is one issue that developed nations often ignore; the challenges that developing countries encounter in this transition. Hence developed nations have worthwhile and powerful assets like an economic muscle, a skilled workforce, and a well-developed physical infrastructure that provides flexible settings for easy integration of renewable energy. The developing nations on the other hand are faced with severe challenges such as poverty, lack of infrastructure, and very high implementation cost of renewable energy. These dynamics are seldom taken into consideration by the push for a global shift to the use of renewable sources of energy, which tends to disregard the specifics of the development of industrializing countries. For instance, the EU and the US can easily commit and achieve high REN standards, while nations in eastern Africa and South Asia have immense struggles in doing the same without affecting their GDP and ES. This is the case where established countries dominate the climate change policies and agenda and do not consider the capacities of developing nations. The insistence on RETs in the least developed countries often excludes the harsh realities of poverty, poor resource infrastructure, and the economic role of fossil fuels. Thus climate policies that predominantly focus on the expansion of RETs without taking necessary measures to address these disparities are likely to widen the existing gap between the developed and the developing states. For a fair and effective climate plan that can be implemented across the world, there must be more focus on optimization of the process of shift towards renewable energy and simultaneously addressing the need for development of the poor industrializing countries while acknowledging the rights of these

nations to industrialize without having to worry much about climate change since the developed nations can afford to invest more for the reduction of carbon emissions.

To appropriately develop a plan for developing nations to maintain stable growth patterns, it is important to observe other countries that may have shifted from the developing to the developed stage relatively recently. One of these countries is Norway, having undergone its “industrialization” to become developed in the last three decades. This statement can be made because sustained economic growth is a sign of industrialization, so using Norway’s Gross Domestic Product (GDP) can allow for a determination to be made about the country’s modernization and how it developed in thirty years. Norway started its industrialization process only in around 1950 as an oil exporting country and hence at the initial stage of industrialization, the carbon emission of Norway is higher. But as the nation grew it also got the capacity to transform to renewable energy sources that is the real picture. So thus, during the years between 1990 and 2020 Norway’s GDP, a standard for development and modernization, has risen while the carbon emissions dropped in the same manner (Hestad and Stensgård 2023). This pattern indicates that the increase in the carbon emission rate in developing countries is a short-term problem. As these nations undergo the process of industrialization, they acquire the ability to harness RETs and therefore cut down on emissions. This success story can therefore prove helpful to other nations that seek to change their energy mix to renewable sources as Norway did. Upon learning what CNET’s Daniel Van Boom has to say about this, I learned that countries like Norway and Iceland have been able to meet almost all their energy demands from renewable sources such as geothermal or hydropower. This is not just environmentally positive but it also brings better energy security as these countries are not soaring high in the/global market and politics for energy. As such, Norway’s experience can be replicated in countries in the developing world, particularly in South Asia, East Asia, and Africa. These regions should be given the carbon space they need for industrialization to mimic Norway ultimately. At the same time, post-industrial economies should bear the responsibility to cut down their carbon emissions which they are capable of doing. Such a two-step approach may guarantee a fair and efficient climate mitigation framework for the entire world by letting developing countries become industrialized while the developed countries set the bar for emissions reduction (Nordhaus and Shellenberger 45).

Potential Solutions, Plans, and Overall Implementation

These findings present a clear climate change policy for the future. If RET use is unfeasible for developing nations, while sustainable and beneficial to developed nations, then adopting a universal renewable energy policy for all nations is unwise. Playing to each region’s economic and developmental state is extremely important, and regions that are still developing should be allotted carbon space, in which fossil-fuel-based energy sources continue to be able to be used until the region reaches a state of development from which transitioning to RET use is possible and efficient. In contrast to this policy, post-industrial societies should have much more pressure placed on them to adopt RET use in their nations to counteract this permitted carbon

space for developed nations. Not only is RET use efficient in these nations, but it allows them economic independence from unstable developing nations while also benefiting the globe and global climate change policy.

Possible implementation most likely resides in the UN. As the world's premier method for international agreement, as well as the platform from which other climate change agreements have passed, the UN represents the most successful method from which to implement this renewable energy goal and carbon space implementation. The UN could facilitate the establishment of binding international agreements that set specific renewable energy targets for all nations, similar to the Paris Agreement's targets on climate change. These agreements would include definitions of what constitutes a developing and developed nation, as well as outlining plans for transition to renewable energy as a main power source following a move from developed to developing nation. Additionally, financial incentives supported by organizations such as the IMF for developing nations to become developed nations could be implemented to promote this renewable energy transition. The UN could also assist countries in developing legal frameworks and regulatory mechanisms that encourage the adoption of renewable energy, such as feed-in tariffs, tax incentives, and renewable energy credits.

Conclusion

While climate change has been a pressing issue for decades on end, the current climate change policy is ineffective because it directly conflicts with the economic goals of developing nations. In their strive to eliminate poverty and modernize, developing nations must industrialize, and can only do so efficiently with fossil fuels. This makes developing nations unable to adopt RETs as a successful method of energy production, making implementing RET use in these countries a failed climate change policy. Developed nations, on the other hand, are in a prime position to adopt renewable energy as their primary source of energy, due to the stability associated with RETs, and the technological infrastructure and knowledge in developed nations to adopt these technologies. This juxtaposition between developed and developing nations shows that an effective global climate change policy allows developing nations “carbon space” to industrialize while putting greater pressure on developed nations to adopt renewable energy. This allows for a slight decrease in carbon emissions across the globe until the majority of developing nations have industrialized, allowing them to switch to renewable energy and create a near 100% renewable globe like many in the UN have been advocating for, in a way that does not conflict with any nation’s economic goals. Implementation of this goal allows for economics to blend with climate change policy, creating the most effective climate change policy possible, and hopefully leading to a greener world.

Works Cited

- American Council on Renewable Energy. "Renewable Energy and Economic Growth." 2023.
- Ayres, R.U., Warr, B. Energy Efficiency and Economic Growth: the 'Rebound Effect' as a Driver. 2009
https://link.springer.com/chapter/10.1057/9780230583108_6
- Baguant, J. The Case of Mauritius. In: Bhagavan, M.R., Karekezi, S., (Eds.), Energy Management in Africa. ZED Books and African Energy Policy Research Network (AFREPREN/FWD). 1992.
- Bhattacharyya, S. C. "Grid Integration of Renewable Energy in India: Challenges and Opportunities." *Energy Policy*. 2019
- BloombergNEF. "Renewable Energy Investment Trends." 2023.
<https://www.bloomberg.com/professional/insights/webinar/asia-pacific-energy-transition-investment-trends-for-2023/>
- Brian Kennedy, Cary Funk and Alec Tyson. "What Americans Think about an Energy Transition from Fossil Fuels to Renewables." Pew Research Center, 28 June 2023,
www.pewresearch.org/science/2023/06/28/what-americans-think-about-an-energy-transition-from-fossil-fuels-to-renewables/.
- Brookings Institution. "Economic Impact of Renewable Energy in the United States." 2023.
- California Energy Commission. "Renewable Portfolio Standard." 2023.
<https://www.energy.ca.gov/news/2023-05/new-data-shows-growth-californias-clean-electricity-portfolio-and-battery>
- Cleveland, Cutler & Costanza, Robert & Hall, Charles & Kaufmann, Robert. "Energy and the U.S. Economy: A Biophysical Perspective. *Science*." 1984
<https://www.science.org/doi/10.1126/science.225.4665.890>.
- Chaudhary, Anjali, et al. "*Barriers to Renewable Energy Development in Rural India: Insights from Policy and Implementation*." *Energy for Sustainable Development*, vol. 62, pp. 136-145. 2019
- Eberhard, Anton, et al. *Africa's Fossil Fuels: Energy Security and Economic Growth in the Age of Climate Change*. Routledge. 2022.
<https://www.sciencedirect.com/science/article/pii/S0973082622000928>
- European Commission. "Horizon Europe." 2022.
https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe_en
- European Commission. "REPowerEU Plan." 2023.
https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal/repower-eu-affordable-secure-and-sustainable-energy-europe_en
- European Commission. "Renewable Energy Directive 2023 Revision." 2023.
https://energy.ec.europa.eu/topics/renewable-energy/renewable-energy-directive-targets-and-rules/renewable-energy-directive_en
- European Investment Bank. "Financing the Green Transition." 2022.

<https://www.eib.org/en/events/financing-the-green-transition-in-times-of-covid-19>

- Eurostat. "Renewable Energy in the EU." 2023.
- Hu, Zhaoguang, et al. *Renewable Energy in China: Resources, Policy, and Competitiveness*. Elsevier, 2020.
- Intergovernmental Panel on Climate Change. "Climate Change 2022: Impacts, Adaptation, and Vulnerability." 2022.
<https://www.ipcc.ch/report/ar6/wg2/>
- International Energy Agency (IEA). "African Energy Outlook" IEA, 2022.
<https://www.iea.org/events/iea-at-cop27-africa-energy-outlook-2022-implementing-and-financing-the-iaa-s-sustainable-africa-2030-scenario>
- International Energy Agency (IEA). "Renewables Integration in India: Policy Recommendations and Key Issues." IEA, 2021.
- International Energy Agency. "Renewable Energy Market Update." 2022.
- International Energy Agency. "Renewable Energy Policies in a Time of Transition." 2022.
- Joshua Mugisha, et. al. Assessing the opportunities and challenges facing the development of off-grid solar systems in Eastern Africa: The cases of Kenya, Ethiopia, and Rwanda. *Energy Policy*, Volume 150. 2021
- Kumar. J, C.R., Majid, M.A. Renewable energy for sustainable development in India: current status, future prospects, challenges, employment, and investment opportunities.(2020).
- Lawrence Berkeley National Laboratory. "Power Purchase Agreements and Renewable Energy Certificates." 2022.
- Hestad, Dagfinn, and Marianne Stensgård. *Norwegian Industrialization and Its Effects on Carbon Emissions*. Oslo University Press, 2021.
- Mukherjee, Ishani, and Benjamin K. Sovacool. "The Role of Fossil Fuels in Poverty Alleviation and Energy Access in Rural India." *Energy Policy*, vol. 132, pp. 536-544. 2019
- National Renewable Energy Laboratory. "Advancing Clean Energy Technologies." 2022.
- Nordhaus, Ted, and Michael Shellenberger. *Break Through: Why We Can't Leave Saving the Planet to Environmentalists*. Houghton Mifflin Harcourt, 2007.
- Organization for Economic Co-operation and Development. "Green Growth and Sustainable Development." 2023.
- Ravallion, Martin, *The Economics of Poverty: History, Measurement, and Policy*, Oxford Academic. 2016.
<https://doi.org/10.1093/acprof:oso/9780190212766.001.0001>
- Sovacool, Benjamin K., et al. *Energy Justice and Transitions to Low Carbon Energy in Sub-Saharan Africa*. Palgrave Macmillan, 2020.
- Singh, A., & Jamasb, T. "Renewable Energy in India: Barriers, Opportunities, and Economic Impacts." *Renewable and Sustainable Energy Reviews*. 2020.
- Weiss, Charles, and William B. Bonvillian. "Innovations in Energy Technology." 2022.
<https://doi.org/10.7551/mitpress/8161.001.0001>

WINDEXchange, "Wind Energy in Texas." WINDEXchange, 2023.
<https://windexchange.energy.gov/states/tx>

U.S. Department of Energy. "Inflation Reduction Act 2022." 2023.
<https://www.energy.gov/lpo/inflation-reduction-act-2022#:~:text=The%20President's%20Inflation%20Reduction%20Act,energy%20manufacturing%2C%20and%20putting%20the>

U.S. Energy Information Administration. Annual Energy Outlook 2023: With Projections to 2050. U.S. Department of Energy, March 2023

United Nations Environment Programme. "Global Renewable Energy Status Report." 2023.

United Nations Framework Convention on Climate Change. "Climate Change and Developing Countries." 2023.

World Bank. "India's Energy Transition: Scaling Up Renewable Energy and Energy Efficiency." 2021

Van Boom, Daniel. "How Norway and Iceland Lead the World in Renewable Energy." CNET, 13 Jan. 2022,
www.cnet.com/science/how-norway-and-iceland-lead-the-world-in-renewable-energy/.

Xin Zhao, Xiaowei Ma, Boyang Chen, Yuping Shang, Malin Song, Challenges toward carbon neutrality in China: Strategies and countermeasures. 2019
https://www.researchgate.net/publication/355140643_Challenges_toward_carbon_neutrality_in_China_Strategies_and_countermeasures

Beliefs: a Psychological Analysis By Zhen Xu

Introduction

Humans are long known to have been privileged to be able to exercise free will. However, we really have little control over how we think: our brains, genes, and upbringings dominate over the wills we have as humans. How we think unquestionably leads to what we believe. A belief is a subjective attitude toward a proposition. We may falsely assume that we have the power to choose our beliefs, since it is merely a subjective thought; however, our subjectivity in belief formation is not entirely under our control. For a belief to be voluntary, the formation of the belief must involve conscious and aware decision making during thinking, and the owner of the belief must have total control over how they think. And since it's known that we lack this control, beliefs are not voluntary.

Definition

A belief can only be voluntary if the owner of that belief deliberately chooses to believe what they believe; in other words, the individual must be aware of what exactly the brain does in every step of the formation of the belief. However, beliefs are naturally derived conclusions from one's mind: our brains make decisions for us in the subconscious, beyond our conscious control where we are unaware of the processes and therefore cannot influence (Jung et al., 1960). Human brains assimilate and analyze information and evidence, wherein the process of analysis follows a natural pathway guided by the inherent dynamics of the unconscious, and the individual undergoing analysis does not have conscious control over the content that arises.

A Case of Involuntary Belief Formation

The rubber hand illusion serves as a profound example of how the brain forms natural assumptions that can occur paradoxically despite our conscious intentions to avoid them. Matthew Botvinick and Jonathan Cohen conducted an experiment where each of the subjects' left arm was hidden from the subject's view on a table, and a rubber model of left hand and arm was placed on the table directly in front of the subject (Botvinick and Cohen, 1998). Two paint brushes simultaneously stroked the rubber hand and the subject's hidden hand. After approximately ten minutes, the subjects cannot resist to experience an illusion in which they begin to perceive the rubber hand as if it were their own, even feeling a tactile sensation when just the rubber hand is touched (Botvinick and Cohen, 1998).

The subjects were fully aware of the procedure of the experiment, meaning that they acknowledge that the rubber hand placed in front of them is not their real hand, but their brain still tricked them into believing an ownership of the rubber hand. The rubber hand illusion demonstrates the brain's ability to amalgamate and interpret multisensory information, i.e. one of many forms of information we assimilate, and decide our beliefs. To conclude, the phenomenon of the rubber hand illusion revealed that the participants' minds constructed an erroneous perception that forced them to believe in their ownership of the rubber hand, despite full

consciousness of the fact that the hand placed in front of them was a mere rubber model. The rubber hand illusion is only one of many examples of illusions, which convincingly proves beliefs involuntary: an individual can not help *but* to perceive things in a specific way – even if one does not wish to.

The Subconscious

Brain processes in the subconscious are outside of our control: in other words, our sense of making choices or decisions is just an awareness of what the brain has already decided for us. Subliminal perception, for example, shows that the brain reacts and processes information that is presented below the threshold of conscious awareness (“Subliminal Definition & Meaning”). We autonomously and constantly process information in the subconscious; these subliminal perceptions obscurely affect our values and knowledge, which play significant roles in the brain’s judgment and analysis of evidence, e.g. confirmation biases, and therefore determine our beliefs.

Nature vs. Nurture

Every individual’s brain thinks in a specific way, which paves that natural pathway in the formation of beliefs. According to Eric Turkheimer, both genes and the environment shape cognitive abilities within individuals (Turkheimer, 2000). Sandra Scarr’s work also highlights how genetic factors influence cognitive processes through their impact on environmental factors. Nature, i.e. genetic factors, impacts the structure and function of the brain, thus influencing the cognitive abilities of the brain. Genetic variations, which are not voluntarily selected by their owners, can affect neural plasticity, how we perceive sensory information e.g. visual and auditory stimuli, and the neurotransmitter system, all of which impact how humans interpret information and form their own opinions, and therefore beliefs. Nurture, on the other hand, modifies the way an individual thinks as the individual accumulates experience. The modification process of one’s cognitive abilities occurs predominantly during childhood and adolescence (Piaget, 1954). As children, humans are not capable of choosing their own environment, so by extension what alters how we think and consequently our beliefs is not chosen by us.

Our genes and the environment in which we develop as human beings conjointly construct that natural pathway an individual’s brain undertakes when assimilating and analyzing information. The so-called logical thinking or pondering where we seemingly evaluate evidence from all perspectives and arrive at a conclusion which becomes our belief is in reality is not actually under our full conscious control. For example, heuristics are mental shortcuts our brains take to save cognitive loads. Heuristics are often based on previous experiences (nurture), e.g. availability heuristic and representativeness heuristic. The purpose of thinking is to obtain an unbiased conclusion with logical evidence, but humans are more often than not unaware of utilizing heuristics during processes of thinking, causing beliefs to be biased, which is against our will. The usage of heuristics is usually not under one’s conscious control, therefore heuristics

can distort one's beliefs without one's voluntary consent. How we think, before forming a specific belief, has already been predetermined by genetics and the environment of our nurture, so our beliefs, as a result of how we think, are not voluntary.

Some may argue that beliefs are voluntary because no outside force is coercing an individual into holding a certain belief, but it is undeniable that the process of belief formation is influenced by a variety of factors that are not entirely under our control or awareness, which are a form of external forces that we humans cannot repel. A widely accepted definition of belief that opposes my claim is "belief is a persistent readiness to act in a certain way toward a certain object, and is a selective interpretation of what an object is" (Dewey, 1929). How an individual's brain assesses the available options for selection, though, is formed by the combination of accumulative environmental influence and human genetics, where neither of which our free will has the dominance over.

Conclusion

Beliefs are not conclusions by human choice or will; they are not voluntary. A vast amount of things, like nature, the environment, the subconscious, the brain itself, and many more, rule over the free will of human beings. In that sense, nothing is necessarily voluntary because there is always an existence of a stronger outside force that influences one's actions or thoughts. We, as humans, believe in what we believe, but we do not voluntarily choose what we believe.

Works Cited

- Bennett, O. "Why Is Belief Involuntary?" *Analysis*, vol. 50, no. 2, 1990, pp. 87–107, <https://doi.org/10.1093/analys/50.2.87>. Accessed 10 May 2023.
- Botvinick, Matthew, and Jonathan Cohen. "Rubber Hands 'feel' Touch That Eyes See." *Nature News*, 1998, www.nature.com/articles/35784/. Accessed 08 May 2023.
- Dewey, John. *The Quest for Certainty: A Study of the Relation of Knowledge and Action*. Allen, 1929.
- Jung, Carl, et al. *The Structure and Dynamics of the Psyche*. Pantheon Books, 1960.
- "Nature vs Nurture Debate." *Encyclopedia of Child Behavior and Development*, 2011, pp. 993–993, https://doi.org/10.1007/978-0-387-79061-9_5485. Accessed 10 May 2023.
- Piaget, Jean. "The Construction of Reality in the Child." *American Psychological Association*, psycnet.apa.org/record/2006-09595-000. Accessed 10 May 2023.
- Piaget, Jean. *The Psychology of Intelligence*. Taylor & Francis, 2015.
- Schwitzgebel, Eric. "Belief." *Stanford Encyclopedia of Philosophy*, 3 June 2019, plato.stanford.edu/entries/belief/. Accessed 28 April 2023.
- "Subliminal Definition & Meaning." *Merriam-Webster*, www.merriam-webster.com/dictionary/subliminal. Accessed 15 June 2023.
- Turkheimer, Eric. "Three Laws of Behavior Genetics and What They Mean." *Current Directions in Psychological Science*, vol. 9, no. 5, 2000, pp. 160–164, <https://doi.org/10.1111/1467-8721.00084>. Accessed 10 May 2023.

Revolutionizing Justice by Addressing Racial Bias in Capital Sentencing

By Santo Raggiri

Abstract

This paper critically examines the discrimination argument against the death penalty, with a particular focus on the concepts of comparative and noncomparative justice. Through an analysis of Michael Cholbi's argument for a moratorium, the paper explores how racial disparities in capital sentencing constitute a profound injustice against African Americans, undermining both their equal legal status and protection under the law. Cholbi's argument suggests that these disparities are so deeply ingrained that they require a moratorium to allow time for systemic reform. The paper then addresses the counter arguments posed by Patrick Lenta and Douglas Farland, who argue that a moratorium would lead to a loss of noncomparative justice, as it would result in murderers facing less severe penalties than they deserve. The central argument of this paper challenges Lenta and Farland's critique by redefining the concept of desert, proposing that the sentencing of white offenders, who are less affected by racial bias, should serve as a more accurate standard. By adopting this standard, the paper demonstrates that a moratorium would not only enhance comparative justice by addressing racial disparities but also strengthen noncomparative justice by ensuring that punishments more accurately reflect what individuals deserve. Ultimately, the conclusion asserts that a moratorium on the death penalty is justified as it leads to substantial increases in both forms of justice.

Introduction to the Discrimination Argument

When participating in a discourse surrounding the death penalty, arguments that oppose the death penalty are classically separated into two different categories. The first type of argument is known as a substantive argument – an argument that challenges the death penalty's moral, ethical, or philosophical justification. Substantive arguments against the death penalty assert that it is inherently wrong and cannot be justified under any circumstances.

For example: a proponent will argue that if someone kills another person, they deserve to die. This is often the most basic path through which all retributivist proponents of the death penalty justify the death penalty. An opponent, on the other hand, will argue in opposition using principles such as the sanctity of human life, the immorality of state-inflicted death, the natural progression of society and suffering, an inability to properly scale punishments proportionally, or the belief that capital punishment is overly severe.

Separate from substantive discussions surrounding death penalty, is the discussion surrounding whether the death penalty can be justly imposed. These arguments are known as proceduralist arguments, and they focus on the fairness, consistency, and reliability of the processes used to impose the punishment. Proceduralist arguments claim that there are unfixable flaws in the processes used to determine what smaller subset, out of a large set of murderers, deserves to be executed. Thus, proceduralist arguments do not inherently challenge the moral legitimacy of the death penalty itself but rather scrutinize the way it is applied. In fact, the

purpose of arguing against the death penalty through a proceduralist lens is that the substantive discourse can be completely yielded – that is, a proceduralist can agree that the death penalty is morally permissible in theory but can base their argument on how that theory gets put into practice. For the purposes of this paper, I will employ this method and the arguments that follow will not question the substantive permissibility of the death penalty.

Two types of proceduralist arguments are often made to oppose the death penalty. The first type is known as the arbitrariness argument. It compares large groups of people who have been sentenced to death and others who have not, arguing that the difference in severity of the crimes committed in these groups is indistinguishable. Thus, the death penalty is administered arbitrarily, since the decisions leading to its imposition are made without clear, consistent, or rational standards, resulting in unpredictable and unjust outcomes. This argument emphasizes that the application of the death penalty is fraught with inconsistencies that undermine its legitimacy and underscore its inability to be rationally imposed by any given metric. From a legal perspective, the arbitrariness argument often scrutinizes the criteria used to determine who receives the death penalty, claiming that the criteria are often vague and open to interpretation. In capital trials, the difference between life in prison and capital punishment often comes down to descriptive qualities of the murder such as "heinous," "atrocious," or "depraved". Those terms lack clear definitions, allowing for overly subjective judgments by prosecutors, judges, and juries. This vagueness can lead to disparate outcomes for similar crimes, depending on how these terms are interpreted in different cases.⁵⁷

The second type of proceduralist argument is known as the discrimination argument. This argument is distinct from the arbitrariness argument. While the arbitrariness argument claims that there is no rational distinguishability between murderers who are executed and murderers who are not, the discrimination argument claims that there is an unjust bias that creates this discrepancy. This bias is against racial minorities, specifically African Americans, and the discrimination argument claims that they are executed in such a frequency compared to whites that is disproportionately unjust. Justifications for the discrimination argument begin with "empirical studies dating back to the 1940s, [indicating] that, all other things being equal, racial minorities, particularly African Americans, are disproportionately more likely to receive the death penalty for murder than are convicted whites".⁵⁸ Furthermore, empirical studies delineate a bias in defendants. The 1987 Baldus study found that murderers of whites were more likely to be executed than murderers of African Americans, irrespective of any differing degree of severity.⁵⁹

Proponents of the discrimination argument will say that this stark data indicates that the death penalty cannot be fairly applied. However, for their argument to hold, they must also argue that the biases that plague the criminal justice system are irremediable. If the biases are irremediable either abolition or a moratorium is required. Thus, opponents will point to specific places in the criminal justice system where bias is permanently entrenched to support their arguments. Some argue that judges and juries either consciously or unconsciously allow racial

⁵⁷ Dolinko, David. "Foreword: How to Criticize the Death Penalty.", (p.560).

⁵⁸ Cholbi, M. Race, Capital Punishment, and the Cost of Murder., (p.255).

⁵⁹ Cholbi, (p.267).

biases to influence their decisions, leading to harsher penalties for minority defendants. Others will say that errors in the legal process—such as inadequate legal representation, prosecutorial misconduct, or flawed jury instructions—can lead to outcomes at unjust rates for African Americans.⁶⁰ Some might argue that issues like prosecutorial misconduct and inadequate legal representation seem remediable because they can be addressed through reforms such as stricter oversight, better training, and increased funding for public defenders. However, imposing more restrictions in the legal system to make it more fair for African Americans may not fully address deeper, systemic problems. These reforms might fail to eliminate entrenched biases in areas such as prosecutorial discretion, jury selection, or sentencing, which are often subtle, deeply ingrained, and influenced by societal prejudices. For example, in many jurisdictions, African American defendants are disproportionately excluded from juries due to peremptory challenges, a practice that allows prosecutors to dismiss potential jurors without stating a reason, often masking racial biases. Additionally, prosecutorial discretion often results in African American defendants being charged with more severe offenses than their white counterparts for similar conduct, leading to harsher plea deals or sentences, which perpetuates systemic inequality even when legal reforms are in place. As a result, these biases are considered by many to be irreparable because they stem from broader societal issues that cannot be easily corrected through changes to legal procedures alone.⁶¹

After establishing that at least some level of bias in the legal system is irreparable, the proponent of the discrimination argument will explain why these biases should be dealt with very seriously. A bias at any stage contributes to an unequal application of justice that develops an argument for why the death penalty cannot be justly imposed. Bias in the legal or criminal system is generally understood to be against notions of justice and equality. Most justice systems are founded on principles of punishing those proportionately to the crimes they have committed, regardless of who they are or personal opinions of them. It would be inconceivable for two people to commit the same crime, yet one of them receive twice the time in jail because that person wore red clothes on their court date. Factors that have to do with anything except for details of the crime committed should have no place in determining the weight of justice on an individual – the same applies for race.

In this paper, I will critically examine the debate surrounding the death penalty, focusing specifically on the proceduralist arguments and the concepts of comparative and noncomparative justice. We will begin with a detailed summary of Michael Cholbi's argument for a moratorium on the death penalty, which posits that racial disparities in sentencing constitute a significant political and social injustice against African Americans as a class. Cholbi argues that these disparities undermine the equal legal status and equal protection of African Americans under the law, and he contends that a moratorium on the death penalty would give time to address these systemic biases.

⁶⁰ Dolinko, (p.577).

⁶¹ Cholbi, (p.257).

Next, we will explore the counterarguments presented by Patrick Lenta and Douglas Farland, who critique Cholbi's position by focusing on the potential loss of noncomparative justice that a moratorium might entail. Lenta and Farland argue that while a moratorium might achieve greater comparative justice by reducing racial disparities, it could simultaneously result in a loss of noncomparative justice, as African Americans would face a lower cost for murder than they should be facing, noncomparatively speaking. This loss in noncomparative justice disrupts Cholbi's mathematical calculation, eliminating his grounds for a moratorium

The central argument of this paper challenges Lenta and Farland's critique by considering using the white American population as a standard for desert. Since this population is less impacted by bias, we should trust that white Americans are more accurately sentenced. By reassessing desert for Black individuals in this context, we can show that a moratorium on the death penalty would not lead to a loss in noncomparative justice, as they claim, but rather to a gain. While Cholbi and Lenta and Farland's arguments are mathematical, my strictly retributivist argument eliminates this weighing process entirely. My approach supports Cholbi's position, revealing that a moratorium not only enhances comparative justice by addressing racial disparities but also strengthens noncomparative justice, leading to a more just and equitable outcome overall.

Cholbi on the Cost of Murder

In his paper "Race, Capital Punishment, and the Cost of Murder," Michael Cholbi poses an intriguing version of the discrimination argument, arguing that racial disparities constitute a political injustice against African Americans as a class, not just a judicial wrong against individual defendants. This foundational premise supports his call for a moratorium on the death penalty in the United States.

Cholbi begins by contextualizing the legal backdrop of racial discrimination in capital sentencing in the United States. Since the Supreme Court's decision in *Gregg v. Georgia* in 1976, which reinstated the death penalty, opponents of capital punishment have faced significant challenges in court. In 1987, a defining statistic was revealed: African American defendants are "1.1 times as likely as defendants of other races to be sentenced to death for murder, while defendants of all races are 4.3 times more likely to be sentenced to death when the victim is white than when the victim is of another race".⁶² However, later that year the Court ruled in *McCleskey v. Kemp*, demanding specific evidence of racial discrimination in individual cases, effectively dismissing broad statistical evidence as insufficient for overturning death sentences. This requirement created a narrow and often unachievable path for defendants to prove racial bias, despite substantial empirical evidence indicating pervasive racial disparities.⁶³

Thus, one of Cholbi's major obstacles is ensuring the relevance of his argument despite the Supreme Court's assertion that broad empirical evidence is not sufficient to claim bias in any individual case. Cholbi's goal is to argue for a moratorium on the death penalty; however, instead

⁶² Cholbi, (p.267).

⁶³ Cholbi, (p.256).

of trying to narrow the scope of his argument to individual cases, he expands the scope of his argument, ultimately stating that racial discrimination in the application of the death penalty plagues all African Americans, even those who are not charged with crimes. By expanding his argument, he bypasses the Supreme Court's objection which concerns only African Americans who are charged with crimes and examines all African Americans as a class.

Cholbi begins his argument by explaining that the racial discrimination in capital sentencing leads to a distorted "market" for murder, where African Americans face higher costs for committing murder due to a greater likelihood of receiving the death penalty. This discrepancy generates rational expectations among African Americans that they will be treated more harshly by the criminal justice system, which undermines their equal status under the law. Then, Cholbi postulates a principle that he calls the Principle of Equal Legal Status:

Individuals do not enjoy equal legal status relative to a given crime if (a) the law or legal practices provide different individuals with different expectations about the likely costs of committing that crime, and (b) the differences in expectations are best explained by a factor other than differences in individuals' desert.⁶⁴

Cholbi writes that this principle is violated as African Americans face different expectations about the legal consequences of their actions solely based on their race.⁶⁵ Therefore, African Americans do not enjoy equal legal status in the United States, going against some of the founding principles of equality and justice that are the basis its legal system. It's problematic to face different expectations about legal consequences because that impacts your ability to enjoy life and feel protected by the state, at least relative to the enjoyment and protection of white Americans.

Cholbi makes a similar argument postulating another principle he writes out as the Principle of Equal Protection:

Individuals do not enjoy equal protection relative to a given law if (a) the costs of violations of that law vary depending on who the victim of the violation is, and (b) these cost differences are explained by factors unrelated to the intrinsic seriousness of the violation.⁶⁶

This principle posits that individuals do not enjoy equal protection under the law if the costs of violating the law vary based on irrelevant factors, such as race.⁶⁷ Empirical studies such as the Baldus study mentioned above confirm that murderers who kill African Americans are less likely to receive the death penalty compared to those who kill whites. This disparity implies that African American lives are valued less, violating their right to equal protection. In total, both

⁶⁴ Cholbi, (p.263).

⁶⁵ Ibid

⁶⁶ Cholbi, (p.268).

⁶⁷ Ibid

African Americans' rights to equal protection and equal legal status are violated, and Cholbi contends that such racial justify a moratorium on the death penalty until those biases are eradicated.

To strengthen his argument, Cholbi engages with an objection to the discrimination argument popularized by Ernest van den Haag. Van den Haag argues that the discriminatory sentencing of capital punishment does not affect its justice to those who deserve it. He says that that as long as some murderers receive their deserved punishment, their sentences are just , even if as a whole the punishment is administered in a racially discriminatory manner. For example: imagine a courtroom with a racist judge overseeing an egregious and heinous mass murder case committed by an African American defendant. Notwithstanding the judge's racism, the defendant deserves the death penalty because of the severity of the crime in question (if we accept that death is a just penalty for murder). Van den Haag sees that because the person deserves it, it doesn't matter if there is racial bias involved. The logic used to govern this extreme case should apply for all cases in which the defendant deserves the death penalty. Cholbi challenges this view by emphasizing the broader social injustice inflicted on African Americans as a class. He posits that the cumulative costs and reduced liberties for African Americans, stemming from racially biased capital punishment, outweigh the justice achieved by executing a few deserving murderers. He does an estimated mathematical calculation:

On average about 3500 prisoners were held on Death Row in the U.S. over the past five years. Let us dub this Group One. Under a moratorium, few if any of these individuals would be executed "on schedule", a loss from the standpoint of justice (if we accept that death is a just penalty for murder). Historically about a third of those sentenced to die in the U.S. have been African American, and over 1000 of those currently on Death Row are African American. While it is unclear exactly how many of these are slated for execution in part due to racial discrimination, the number is likely to be substantial, probably in the hundreds. Call this Group Two. At the same time, the African American population in the U.S. is nearly 35 million. This is Group Three. For van den Haag to emerge victorious on this question, the injustice of not immediately giving those in Group One precisely what they deserve must be greater than the justice gained (by means of a moratorium) vis-a-vis Groups Two and Three. But a comparison of the numbers involved shows that the injustice done to each member of Group One must be many thousands of times worse than the injustice done to each member of Groups Two and Three, absent a moratorium.⁶⁸

⁶⁸ Cholbi, (p.275).

Net Gain/Loss of Justice

Red=Loss, Green=Gain

Note: Not to Scale

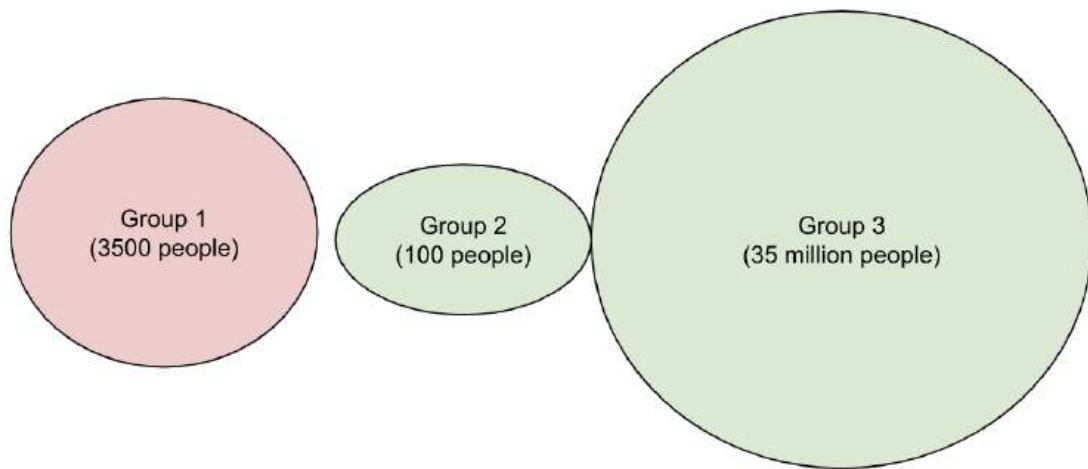


Fig 1: Cholbi's Net Gain/Loss of Justice

To substantiate his call for a moratorium, Cholbi presents this comparative analysis of injustice. After his mathematical estimations, Cholbi writes, “surely it is better to give some murderers less than they deserve than to give the entire African American community less than what it deserves, namely, equal status under, and the equal protection of, the law”.⁶⁹ When he references the murderers who will receive less than what they deserve, this is group one, a mere 3500 people. Though this does not outnumber group 2, the hundred or so African Americans whose sentences under a moratorium would be noncomparatively just, 35 million African Americans from group 3 would get what they deserve in their equal legal status and equal protection under the law. Mathematically, the significant positive impact on African Americans as a class justifies a moratorium.

Lenta and Farland in Opposition:

Patrick Lenta and Douglas Farland wrote their paper *Desert, Justice and Capital Punishment*, aiming to critique recent procedural objections to the death penalty, tackling both "arbitrariness and discrimination" (AD) at once. What they refer to as the AD objection is slightly expanded from what this paper has so far called the discrimination argument as it also includes the arbitrariness argument.

⁶⁹ Cholbi, (p.276).

One of the primary critiques delineated by Lenta and Farland focuses on Cholbi's reconceptualization of the AD objection. To critique Cholbi, Lenta and Farland accept that the first portion of Cholbi's argument – where he argues that all African Americans are treated unjustly by discriminatory death penalty sentencing -- is correct. Having no issue with that part of Cholbi's argument, instead, Lenta and Farland make their objection around Cholbi's mathematical "weighing". Their objection brings us back to the concepts of comparative and noncomparative justice. Before moving on to the specifics of Lenta and Farland's critique of Cholbi, I should give a more in-depth definition for both comparative and noncomparative justice.

Noncomparative justice relates an individual's crime and punishment to what they deserve, operating under the proportionality principle of retributivism. This principle asserts that the severity of punishment should correspond to the severity of the crime. For example, someone who commits a minor theft should receive a lighter sentence than someone who commits a violent assault. However, comparative justice extends beyond simple proportionality, often emphasizing equality in punishment relative to others' actions rather than strictly adhering to the individual's deserts. This raises debates about whether comparative justice aligns with desert-based retributivism, as it sometimes calls for adjustments to achieve equality that might deviate from strictly deserved punishments.

Noncomparative justice is more straightforward as it focuses solely on the individual's crime and the deserved punishment, independent of others' sentences. According to noncomparative justice, if a crime warrants a specific punishment, that punishment remains just irrespective of what others receive. Thus, the focus of noncomparative justice remains solely to ensure that everyone receives the punishment they deserve based on their actions, not on achieving parity among offenders.

For example: suppose ten individuals committed murders of the same severity, all deserving of the death penalty. A world where all ten individuals receive the death penalty for their crimes is both comparatively and noncomparatively just. A world where nine individuals receive the death penalty and one individual receives a life sentence (for no good reason) is slightly unjust, both comparatively and noncomparatively. This world is noncomparatively unjust because one individual did not receive the sentence that they deserve. This world is comparatively unjust because one individual received a more favorable sentence relative to the nine others for no apparent reason. Now, take a world where all ten individuals receive life sentences. Comparatively, this world is completely just since all ten offenders deserving sentences of equal severity received the same sentence. However, noncomparatively, this world is the least just world since not a single offender received the sentence they deserve. Someone like Ernest van den Haag who only considers noncomparative justice in his writing would say that a world where five individuals receive the death penalty and five individuals receive life sentences is far more just – which it is, noncomparatively speaking. Previously, no individuals received the crime they deserve; whereas now five individuals received the crime they deserve.

On the flip side, this scenario is the least just comparatively. With a 50/50 split, half of the individuals received comparatively less severe punishments for no apparent reason.

This overarching discussion becomes crucial in the context of the death penalty and the discrimination argument. The discrimination argument inherently relies on comparative justice, highlighting how racial biases lead to unequal sentencing. Empirical data indicates that African Americans are disproportionately more likely to receive the death penalty for murder compared to whites. By noncomparative standards, if those African Americans deserved the death penalty, their punishment would be considered just regardless of others' sentences. However, from a comparative justice perspective, a significant injustice is evident. The discrepancy in sentencing based on race reveals a systemic bias, undermining the principle that like cases should be treated alike.

Consider two murder cases with similar circumstances but different racial backgrounds of the defendants. If an African American defendant receives the death penalty while a white defendant receives a life sentence, comparative justice would highlight the disparity as an injustice. This disparity suggests that the criminal justice system values the lives of African American defendants less, reflecting deep-seated racial biases. Lenta and Farland agree with Cholbi up to this point; however, they argue that Cholbi's analysis overlooks the noncomparative aspect, where the primary concern is whether each individual received the punishment they deserved. They contend that achieving equality by lowering the punishment for all (a moratorium on the death penalty) may fail to address the core issue of ensuring just deserts for each crime, potentially leading to a loss of justice in noncomparative terms.

More specifically, Lenta and Farland take issue with how Cholbi writes up his mathematical analysis. Lenta and Farland critique Cholbi's mathematical calculation by comparing the number of individuals affected under the death penalty system versus those affected by a moratorium. Cholbi claims that a moratorium would produce a greater net gain in justice. Lenta and Farland challenge this by arguing that Cholbi's calculations are flawed, neglecting a loss of noncomparative justice.⁷⁰ They assert that the loss in noncomparative terms for all black citizens and those sentenced to death would not be justified by the gain in comparative justice for a larger number of individuals. Lenta and Farland write, "[Cholbi's analysis] surely involves a loss in noncomparative terms, since 35 million blacks will now be facing a lower cost than they should, noncomparatively-speaking, be facing for murder, in order to achieve parity between the cost that they are facing and the cost that whites, as a result of discrimination, are confronted with under the capital punishment system".⁷¹ To them, Cholbi fails to account for the moral weight of noncomparative justice adequately and that Cholbi's method of balancing these injustices is overly simplistic and does not hold up under scrutiny.

⁷⁰ Lenta, Patrick & Farland, Douglas. *Desert, Justice and Capital Punishment.*, (p.288).

⁷¹ *Ibid*

Net Gain/Loss of Justice

Red=Loss, Green=Gain

Note: Not to Scale

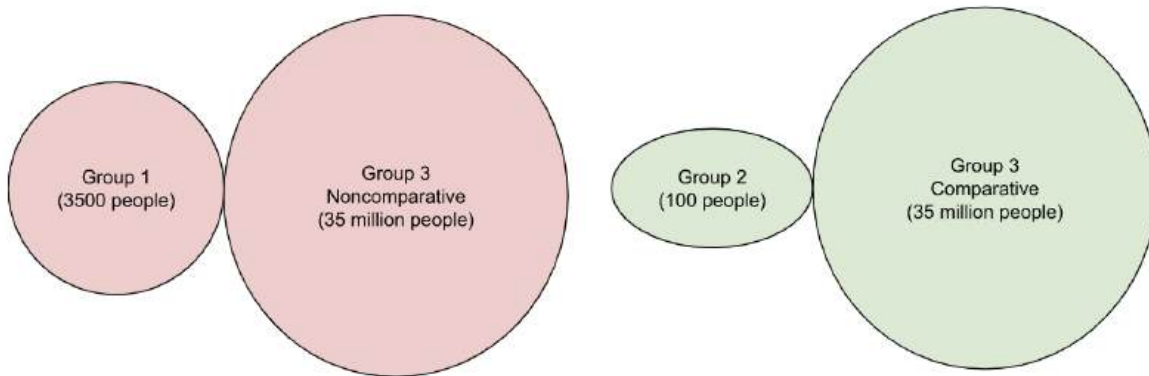


Fig 2: Lenta & Farland's Net Gain/Loss of Justice

As a whole, Lenta and Farland's objection is rather simple: Cholbi's math is wrong. Whereas Cholbi argues that group 1 is outnumbered by groups 2 and 3, Lenta and Farland essentially point out that Cholbi was negligent – group 3 is not as straightforward as Cholbi explained. Within group 3, 35 million African Americans face a gain in comparative justice but also a loss in noncomparative justice, meaning that, group 3 cancels out and a moratorium would net negative in terms of justice. They argue that Cholbi's approach simplifies the issue by focusing primarily on the numbers of those affected by comparative injustice while ignoring the depth and severity of noncomparative injustice that would arise if murderers received less punishment than they deserve. Thus, Cholbi's argument for an easy win is not as straightforward or compelling as he suggests.

In conclusion, Lenta and Farland argue that Cholbi's defense of the AD objection and his call for a moratorium on the death penalty are based on an incorrect balancing of comparative and noncomparative justice. They assert that addressing procedural flaws such as arbitrariness and discrimination should focus on reforming the system to ensure the fair and just application of the death penalty, rather than abolishing it altogether. Now we can turn to the main argument of my paper, which demonstrates how redefining desert within our flawed justice system nullifies Lenta and Farland's critique. This approach supports Cholbi's position by showing that a moratorium on the death penalty actually enhances both comparative and noncomparative justice.

The New Standardized Solution

Lenta and Farland's main critique asserts that Cholbi's analysis involves a loss in noncomparative terms, as a moratorium on the death penalty would mean 35 million African Americans would face a lower cost for murder than they should be facing noncomparatively. However, this critique is based on a skewed definition of desert that doesn't account for what desert means in our flawed criminal justice system. By defining desert – and thereby redefining noncomparative justice – more accurately, we will show that there is actually a noncomparative gain, nullifying Lenta and Farland's argument and supporting Cholbi's position.

Desert is a fundamental concept in discussions of justice and punishment, referring to what a person deserves based on their actions. Initially, desert can be simply defined as the notion that individuals should receive outcomes – punishments in the context of crimes. However, to fully grasp the implications of desert in the context of justice, it is essential to delve deeper into its application, through the principle of proportionality, retributivism, and the definition of noncomparative justice.

The principle of proportionality is central to the definition of desert because together, the two complement each other to define the principles of retributive justice. The principle of proportionality asserts that the severity of punishment should correspond to the severity of the crime. This means that minor offenses should receive lighter punishments, while severe crimes should be met with harsher penalties. This principle ensures that justice is meted out in a way that reflects the moral gravity of the wrongdoing. When we incorporate the principle of proportionality into the concept of desert, we refine our definition: (negative) desert means that a person deserves to be punished in a manner and to a degree that corresponds to the severity and nature of the crime they have committed.

This definition of desert is crucial to our understanding of noncomparative justice. Noncomparative justice posits that each individual should receive the punishment they deserve based solely on their actions, independent of how others are punished for similar actions. This approach focuses on individual culpability and the intrinsic relationship between the crime and the punishment, without reference to the treatment of others. Noncomparative justice is, therefore, another extension of our definition of desert. Desert remains the key component – the part where we as a society determine and define what outcome an individual should receive. Noncomparative justice simply adds that this examination should not consider any factors except the individual's actions.

Now that we understand these definitions and how they relate to one another, let's take another look at the definition of desert – what I have identified to be the key component of our discussion. Generally speaking, it's not easy to assess what people deserve. Even with all of the framework I have presented above, it's still not easy, for instance, to answer this question: do all murderers deserve death. To this day this question is still heavily debated amongst philosophers and lawyers alike. The reality of a criminal justice system adds that these decisions on what individuals deserve are any point, somewhat subjective. Whether it's the person who writes laws or the person who judges criminal acts, these abstract definitions of desert need to be applied.

During that process, those definitions encounter humans and their subjective interests. There is nothing we can do to make our justice system completely free of all bias because no human decision can be entirely objective.

That being said, I have found a method that can ensure that our justice systems are as objective as possible. This method creates a standard for justice, attempting to eliminate all kinds of egregious bias. Accepting that things will never be perfect, if we design and reframe our justice system according to its most unbiased standard, we will optimize the application of all of the abstract principles: desert, proportionality, and noncomparative justice. Now, I will reveal the issue with Lenta and Farland's argument, revealing what should be the new standard for justice in the US criminal-legal system.

As we all agree, including Lenta and Farland, empirical evidence consistently shows that black offenders are punished more harshly than white offenders for similar crimes. This evidence defines the harshness of punishment against black offenders as it related to white people. At first glance, this isn't helpful for identifying what anyone deserves. Take another example: Alex and Charlie both consistently submit work of almost identical quality to their teacher but Alex consistently receives higher grades. Does the teacher have a positive bias towards Alex or a negative bias towards Charlie? With this limited information, either Charlie is getting deserving grades and Alex is getting overly generous grades, or Charlie is getting overly harshly graded while Alex is getting deserving grades. We require more specific information to then determine who is receiving grades that are undeserved. Similarly, for this empirical evidence, are black offenders punished overly harshly or are white offenders punished overly generously? It seems as though it would be more helpful to define the harshness of punishment against some ideal standard of desert.

However, because of the role subjectivity plays in the US criminal-legal system, it is quite difficult to generate an ideal standard of desert. Biases get in the way, and if we agree that biases should be minimized because they contribute to an unequal application of justice, it follows that we should try to find a standard for justice that is least affected by bias and subjectivity, dubbing that our new ideal standard. Clearly, empirical evidence the US criminal-legal system does a poor job determining what black offenders deserve based on an unbiased assessment of their actions. On the other hand, because the white population in the United States does not suffer from racial bias, we can trust that, at least compared to the black population in the United States, white offenders receive punishments that align more closely with what they deserve. Thus, our theoretical definition of desert the concept of desert is more accurately applied to white offenders. This means that the most accurate choice we should trust to be our standard of desert should be that which is imposed on white offenders. Thus, the best comparative standard of desert that we would like to impose on the entire American legal system should, in fact, be framed as it relates to the white population in America – exactly how it is framed above by the empirical evidence. In other words, it is useful to measure the standard of desert based on what white Americans are given as sentences because that demographic is most accurately sentenced. Using our best standard of applicable desert, if we can say that black

offenders are punished more harshly than white offenders for similar crimes without substantive reasoning, we can now say that black offenders – or at least a subset of them – are punished more harshly than they deserve.

In this case, achieving comparative justice can serve as a pathway to noncomparative justice, where individuals receive both precisely what they deserve and equality with reference to others. By setting white offenders' sentencing as the standard, we establish a benchmark that is, as discussed, less influenced by racial biases. This approach helps us identify that black offenders are being punished more harshly than deserve – meaning that lowering the overall severity of their punishments would both lower them to the appropriate standard and the standard on which others are judged. Treating black offenders equally does not just address the disparities in comparative justice; it also moves us closer to achieving noncomparative justice, where each individual's punishment is aligned with their desert. Thus, the final diagram of justice should showcase dual gains in justice for group 3, completely overshadowing the impacts of groups 1 and 2:

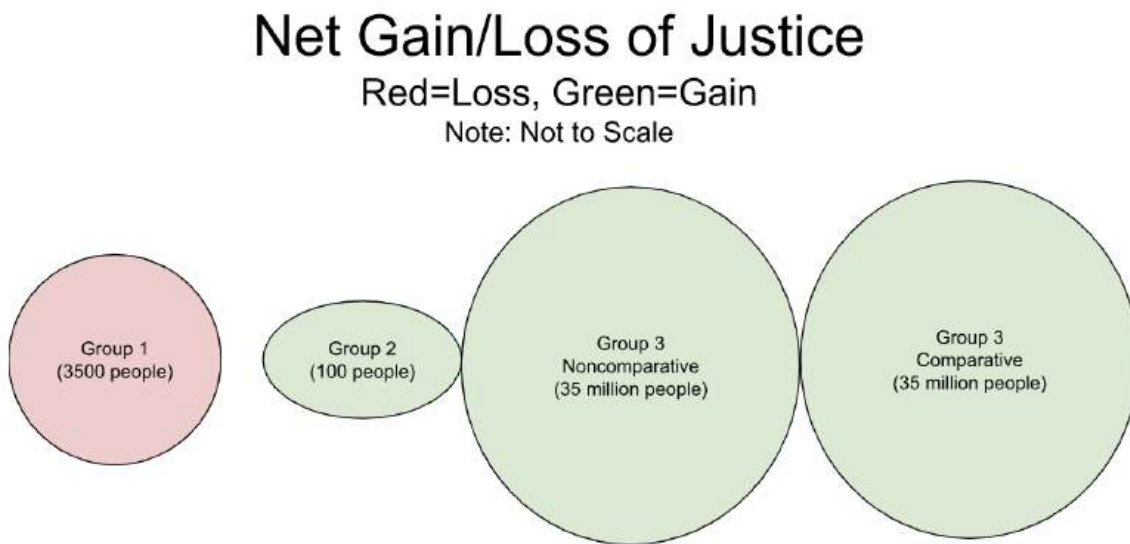


Fig 3: Raggiri's Final, Corrected Net Gain/Loss of Justice

To recap, Lenta and Farland's perspective fails to account for a proper standard of desert in the US criminal-legal system. We should trust that white offenders, due to their lower exposure to racial bias serve as good standard for desert. The punishments that they receive for their crimes are less riddled with biases and subjectivity, making them a good demographic to look to in search of a standard for proportional punishment. Comparatively, by analyzing

empirical evidence, we see that black offenders, due to their higher exposure to racial bias, receive sentences that do not accurately reflect what they deserve. Black offenders – or at least a subset of them – receive overly harsh punishments. Using the white population as our comparative standard, we can address and correct the harsher punishments imposed on black offenders. Ultimately, this approach not only rectifies disparities in comparative justice but also helps us achieve a more objective, noncomparative justice.

Conclusion

Overall, Lenta and Farland's objection fails to undermine Cholbi's position in favor of a moratorium, which not only remains intact but is further supported by a more precise examination of noncomparative justice. As Cholbi explains, this moratorium is being instilled until a solution is found and implemented to rectify these injustices. Considering comparative issues more carefully shows us that black offenders are treated unjustly in a noncomparative sense. Cholbi's original assertion that black offenders are being treated comparatively unjust still holds; however, this argument adds to his analysis that black offenders are also being treated noncomparatively unjustly. Finally, achieving comparative justice, as Cholbi suggests through a moratorium, also helps us achieve noncomparative justice, nullifying Lenta and Farland's mathematical objection.

Works Cited

Cholbi, Michael. "Race, Capital Punishment, and the Cost of Murder." *Philosophical Studies*, vol. 127, no. 2, 2006, pp. 255-282.

Dolinko, David. "Foreword: How to Criticize the Death Penalty." *The Journal of Criminal Law and Criminology*, vol. 77, no. 3, 1986, pp. 546-601. <https://doi.org/10.2307/1143430>.

Lenta, Patrick, and Douglas Farland. "Desert, Justice and Capital Punishment." *Criminal Law and Philosophy*, vol. 2, no. 3, 2008, pp. 273-290.

Associations Between Social Media Use and Adolescent Depression

By Danah AlBahar

Abstract

Since the pandemic, adolescent mental health struggles have increased, and high social media usage is worsening the effects of teen anxiety and depression. This paper examines the harmful effects of social media use on teenagers, focusing on social comparison, cyberbullying, problematic use, and exposure to violent and disturbing online content. While social media does offer community building opportunities, especially for marginalized groups, it is also a platform for addictive behaviors, harassment, and promoting unrealistic lifestyles and body images. Exposure to violent and sexual content further exacerbates anxiety and desensitization. The findings showcase the importance of parental involvement in teaching media literacy and providing open communication with their teens to mitigate these negative effects. By fostering responsible usage of social media, adults can help teens navigate challenges and promote positive mental health habits.

Introduction

Contrary to the messages displayed in music and movies, the kids are not alright, especially in the current post-pandemic landscape. Since the onset of COVID-19, researchers have tracked a marked increase in mental health struggles amongst adolescents, with one study reporting that adolescent anxiety and depressive symptoms have nearly doubled since the beginning of 2020 (Racine et al.). While pandemic-related factors such as lockdowns and virtual learning have undoubtedly contributed to this increase, there is evidence of a correlation between social media usage and depression among teenagers (Vidal et al.). Since gaining traction in 2007, social media use has impacted the psychological development of adolescents in ways that mental health professionals are still struggling to quantify and fully understand. As defined by the American Psychological Association, social media refers to digital communication through which online users create communities to share information and content (Social Media and the internet). The increased prevalence of smartphones has added to the pervasive presence of social media use among adolescents. In 2018, 95 percent of teens had a smartphone, with 45 percent of teens claiming they are now online on a near-constant basis (Anderson). According to the Pew Research Center, 97% of all adolescents between the ages of 13 and 17 regularly use one or more of the following social media sites: YouTube (85% of adolescents), Instagram (72%), Snapchat (69%), Facebook (51%), Twitter (32%), Tumblr (9%) or Reddit (7%) (Anderson).

Not only do social media platforms allow adolescents constant access to unfiltered content streamed directly to their phones, but they have also created questionable social circles that leave kids vulnerable to threats such as cyberbullying and overly sensationalized pseudo news. Traditional media has widely publicized the worst-case scenario impacts of social media use; depression, addiction, and even suicide are presented as very real risks to adolescents. But are these headlines presenting statistical truths? Or are they exploiting outliers to serve as

clickbait for concerned adults? Determining the actual effects of social media platforms on today's youth has been challenging for researchers because the technology is relatively new and studying predictive behaviors requires analyzing years of data collected from a variety of sources. Fears surrounding adolescents and their use of social media platforms are largely based on anecdotal evidence, but few scary stories are enough to concern parents.

However, for the past few years, psychologists across the world have been conducting surveys, experiments, and scouring and analyzing both data and metadata to better understand the long-term impacts of social media use on adolescent well-being and development. The results are a mixed bag of conclusions that are difficult to quantify. Many adolescents receive benefits from social media use; for example, adolescents who identify as part of the LGBTQIA+ community have utilized social media to connect with other teenagers who are navigating the struggles of sexual identity and lessen feelings of marginalization and otherness by building their own online communities (Craig et al.). For other adolescents, however, social media platforms are a trigger for negative emotions and destructive habits.

Regardless of their findings, researchers acknowledge that social media usage is now a deeply pervasive part of adolescent development. Given the wide range of effects that social media can have on teens, adults who play an important role in the lives of teens, such as parents and teachers, need to be aware of the risks and maintain open lines of communication with adolescents regarding these risks. This paper focuses on the deleterious effects of social media use by examining four of the main factors that contribute to depression and anxiety in adolescents: cyberbullying, problematic use, social media comparison, and disturbing content.

Harmful Interactions

A study by the University of Georgia showed that increased social media usage has led to an increase in the prevalence of cyberbullying (Kao). Online interactions through social media allow levels of anonymity and detachment, creating the perfect breeding ground for targeted, ongoing harassment. This harassment ranges from more traditional forms, such as insulting or ignoring an individual while in online spaces, to sharing nude or embarrassing photos, or even telling disparaging stories about a person on social media platforms. Because adolescents are experiencing rapid physiological and neurological development, they are exceptionally vulnerable to these kinds of personal attacks (Nobre et al.); the World Health Organization concluded that upwards of 50% of mental health issues begin before the age of 14 (Comprehensive Mental Health Action Plan 2013–2030). This means that for teens, harassment in any form has the potential to inflict long-term psychological damage.

Cyberbullying is a pervasive form of ongoing, sustained harassment that is carried out utilizing technology (Marín Suelves et al.). Surveys have found that at least one third of teens have experienced one or more instances of cyberbullying (Patchin and Hinduja). Contrary to traditional bullying, which is commonly associated with violence and physical aggression, cyberbullying inflicts emotional damage on individuals over a sustained period. Cyberbullying

also flips the gender divide that is commonly seen in traditional bullying, with girls experiencing the same, if not higher, rates of cyberbullying as boys (Patchin and Hinduja).

Higher social media usage increases the likelihood of an adolescent experiencing cyberbullying, the negative outcomes of which have been well-publicized. In 2017, Sadie Riggs committed suicide to escape relentless cyberbullying (Family mourns loss of teenager who killed herself because of bullying), and in 2022, Nate Bronstein took his life as the result of severe depression that was the result of persistent online harassment (Hickey). While suicide is a rare response to cyberbullying, this increasingly prevalent form of harassment has resulted in a host of negative, detrimental emotional responses, including shame, humiliation, and feelings of marginalization and isolation. These emotions, in turn, have a deleterious effect on the mental health of adolescents. Worldwide studies have revealed a strong correlation between cyberbullying and depression (Nixon). A survey of American middle school students conducted by Hinduja and Patchin revealed a link between cyberbullying and suicidal ideation for both the victims and perpetrators of this form of abuse (Hinduja and Patchin).

Problematic Social Media Usage

A small subset of adolescents spends excessive amounts of time engaging and interacting with others through social media. As early as 2011, the American Academy of Pediatrics coined the phrase “Facebook Depression” to describe the trend that links social media use to an increase in instances of isolation and loneliness (O’Keeffe and Clarke-Pearson). Researchers have divided this subset into two main groups: intense social media users and problematic social media users (Boer et al). While intense social media users spend many hours online, they do not exhibit behaviors that would constitute a dependency on social media’s dopamine rush. However, problematic social media use resembles addictive behavior. Heavy users exhibiting this pattern experience the same negative emotions associated with addiction and display withdrawal symptoms when they are unable to access social media (Boer et al.). This constant engagement with social media and prolonged screen time often leads to problematic social media users experiencing chronic sleep loss. Chronic sleep loss is correlated to a host of debilitating outcomes, from decreased cognitive ability to socio-emotional deficits (Khalaf et al.). Heavy social media users also tend to have low levels of physical activity, another behavior that is correlated to depressive symptoms. “According to data from several cross-sectional, longitudinal, and empirical research, smartphone and social media use among teenagers relates to an increase in mental distress, self-harming behaviors, and suicidality (Khalaf et al.).”

While lack of social media use causes distress to problematic social media users, ongoing use can do the same, with continued engagement being tied to an increase in the likelihood of anxiety, depression, lower self-esteem, and an overall decrease in mental and physical wellbeing (Shannon et al.). How directly these findings relate to social media in and of itself, however, is less clear. Some researchers have posited the theory that problematic social media use is a result of these feelings and behaviors, not the root cause. According to this theory, what we classify as problematic social media use is actually a coping mechanism used by adolescents to escape the

challenges they face in life (Wang et al.). However, it is a maladaptive coping mechanism because it does not allow the afflicted individual to develop the skills required to face these problems, which can contribute to and worsen depressive symptoms. Data used to form theories such as these make it difficult to differentiate between causes and symptoms, because it supports the idea that social media usage may increase negative feelings for problematic social media users, as opposed to being the cause of them (Pellegrino et al.). Further study is needed to determine if there is a deeper issue wherein an individual with anxious or depressive tendencies, when granted unfettered access to social media, engages in behavior that creates a negative feedback loop and exacerbates their own condition. Doing so would require a shift from the quantitative studies that currently dominate research on the topic, to qualitative studies that can attempt to isolate factors that may influence conclusions (Boer et al.).

Social Media Comparison

Experiencing the negative effects of social media use is not a phenomenon relegated only to problematic social media users; even moderate users can be subject to these risks. Another common issue is comparison. It is natural for adolescents to compare themselves to others (Racine et al.), and social media provides them with the opportunity to do so with no restrictions. Prior to the advent of social media, adolescents could only compare their lives and experiences to those of their friends and classmates. In the age of Instagram and TikTok, adolescents can compare themselves to complete strangers, whether it be someone from a nearby city or an influencer using their profile to market the illusion of a perfect life. When constant comparison becomes a part of teenagers' daily life, it can lead to anxiety and depression, especially amongst those who feel pressured to replicate the idealized images from these influencers (Pellegrino et al., 2022). While these accounts only portray a "perfect" version of their users' lives, many adolescents still view them as the standards by which they should measure their own lives (Pellegrino et al., 2022). When their attempts fall short, adolescents find themselves struggling with feelings of disappointment and inadequacy. This can also lead to a more unique form of anxiety termed FOMO, or fear of missing out; this is the result of adolescents worrying that they are not experiencing the same lifestyles that they see others portraying on social media (Alutaybi et al.).

According to a study of individuals aged 18-29, heavy social media users had higher incidences of lower body image and lower self-esteem, along with a higher prevalence of symptoms of disordered eating (Santarossa and Woodruff). Young girls in particular report feeling negative about their body image, which correlates with social media's influence on eating disorders. Algorithms on video-based social media sites, such as TikTok, create a loop where views of one video related to dieting lead a user's feed to be inundated with videos promoting diets and an ideal body type, triggering dysmorphia and eating disorders amongst teens (Nawaz et al.). Dr Vivek Murthy, the United States Surgeon General, has pushed for social media sites to include warnings about their potential to harm adolescents. His stance has garnered support through the Kids Online Safety Act (Barry and Kang), a measure which has yet to be passed.

Disturbing Content

In the past, parents voiced concerns over the content aired on television, but these concerns are nothing when compared to the sheer volume of uncensored content that is present on social media sites. Since online content is generated by users, it is not subject to the same kind of oversight as rated television content. Exposure to graphically violent and/or sexual images and videos has a significant impact on how teenagers perceive the world, and the algorithms behind social media allow this content to quickly spread and appear in adolescents' social media feeds. According to one study, 23% of teenagers have been unwillingly exposed to pornography, with an additional 9% being sexually solicited (Haelle). Another study uncovered higher rates of exposure to disturbing content, with 31% of the subjects reporting that they have been exposed to sexual content, and 32% reporting that they have, voluntarily or involuntarily, viewed violent content. This barrage of violent and sexual content can cause increased levels of anxiety, as well as desensitization (Garcia).

The effects of disturbing content extend beyond social media and can influence how teens consume news from other online sources. In a world where cell phones allow teenagers constant access to news from nontraditional sources, such as social media, teenagers' world views are being influenced by unfiltered, and sometimes unverified, information. Northwestern's Medill School of Journalism conducted a study whose findings showed over 25% of adolescents consume news through social media, including TikTok (35%), YouTube (37%), and Instagram (33%) (Reger).

Although the statistics mentioned above only apply to students who chose to consume news via social media, it is reasonable to assume that many more are involuntarily exposed to news via these sources, due to the near-constant barrage of content attempting to capitalize on sensationalized takes of breaking news stories. The IS-K terrorist plot to bomb a concert in Vienna is a prime example of the type of news headline that targets adolescents by presenting a story at the intersection of popular culture and relevant news. A survey conducted by the Youth Endowment Fund reported that 60% of adolescents have witnessed real-world acts of violence via social media. Violence-related news fills social media feeds every day, showing everything from global wars to ones that hit closer to home, such as school shootings.

Parental Involvement as a Mitigating Factor for Negative Impacts

The best way for parents to address the negative impacts of social media use is to take preventative measures, including monitoring their children's social media activity and having discussions of what responsible and appropriate usage entails. If a child exhibits depression and isolating tendencies, parents should limit their social media usage to help reduce the likelihood of it becoming problematic. It is also important for parents to help their children understand what cyberbullying is and encourage them to speak with a trusted adult if they experience it. Ignoring and internalizing the emotions that could result from such incidents can increase negative feelings such as loneliness and isolation.

Parents can mitigate the impact of negative news cycles by modeling healthy media habits. Platforms like TikTok offer sensationalized versions of news about wars and natural disasters that spread rapidly because content creators seek to maximize emotional impact for clicks and views. While the reality of the content being shared may or may not be significantly different from what is presented, it's been proven that this content still negatively impacts the psychological well-being of adolescents, including causing increased feelings of hopelessness and loss of control. These feelings can lead to both acute and chronic anxiety and depression. Parents can help their teens manage these feelings by having open discussions about current events. These discussions can have a cathartic effect and allow parents to help their children put these events in perspective.

Conclusion

Social media networking sites are a reality of modern life, but their negative effects don't have to be. Studies show that a minority of social media users suffer from depression and anxiety because of the interactions and the content on these sites. However, many people experience a positive effect, joining healthy online communities and strengthening relationships with peers. Despite this, there has been a rise in problematic social media use and the consequences are clear. Cyberbullying, addictive behaviors, body dysmorphia, and increased consumption of graphic content are on the rise. To counteract this, parents, teachers, and doctors should provide guidelines for social media use and lessons in media literacy to children. Pairing moderate use with regular discussions about content and online interactions will ensure that teenagers are less likely to internalize negative emotions and learn healthy coping strategies.

Works Cited

- Alutaybi, Aarif, et al. "Combating fear of missing out (FOMO) on social media: The FOMO-R method." *International Journal of Environmental Research and Public Health*, vol. 17, no. 17, 2020, pp. 6128–6156, <https://doi.org/10.3390/ijerph17176128>.
- Anderson, Monica. "Teens, Social Media and Technology 2018." Pew Research Center, Pew Research Center, 31 May 2018, www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/.
- Barry, Ellen, and Cecilia Kang. "Surgeon General Calls for Warning Labels on Social Media Platforms." *The New York Times*, *The New York Times*, 17 June 2024, www.nytimes.com/2024/06/17/health/surgeon-general-social-media-warning-label.html.
- Boer, Maartje, et al. "Adolescents' intense and problematic social media use and their well-being in 29 countries." *Journal of Adolescent Health*, vol. 66, no. 6, 2020, pp. S89–S99, <https://doi.org/10.1016/j.jadohealth.2020.02.014>.
- "Children, Violence and Vulnerability 2023." Youth Endowment Fund, 13 Nov. 2023, youthendowmentfund.org.uk/reports/children-violence-and-vulnerability-2023/.
- "Comprehensive Mental Health Action Plan 2013–2030." World Health Organization, 2021.
- Craig, Shelley L., et al. "Media: A catalyst for resilience in lesbian, gay, bisexual, transgender, and Queer Youth." *Journal of LGBT Youth*, vol. 12, no. 3, 2015, pp. 254–275, <https://doi.org/10.1080/19361653.2015.1040193>.
- "Family Mourns Loss of Teenager Who Killed Herself Because of Bullying." *Boston 25 News*, *Boston 25 News*, 2017, www.boston25news.com/news/family-mourns-loss-of-teenager-who-killed-herself-because-of-bullying/538759491/.
- Garcia, Karen. "Graphic Imagery Online Can Lead to Psychological Harm, UC Irvine Experts Say." *Los Angeles Times*, *Los Angeles Times*, 19 July 2024, <https://www.latimes.com/california/story/2024-07-19/violent-images-social-media-public-health-threat>
- Haelle, Tara. "Online Risks Are Everyday Events for Teens -- but They Rarely Tell Their Parents." *Forbes*, *Forbes Magazine*, 26 Feb. 2024, www.forbes.com/sites/tarahaelle/2017/02/28/online-risks-are-everyday-events-for-teens-but-they-rarely-tell-their-parents/.
- Hickey, Megan. "A 15-Year-Old Boy Died by Suicide after Relentless Cyberbullying, and His Parents Say the Latin School Could Have Done More to Stop It." *CBS News*, *CBS Interactive*, 2022, www.cbsnews.com/chicago/news/15-year-old-boy-cyberbullying-suicide-latin-school-chicago-lawsuit/.
- Hinduja, Sameer, and Justin W. Patchin. "Bullying, cyberbullying, and suicide." *Archives of Suicide Research*, vol. 14, no. 3, 28 July 2010, pp. 206–221, <https://doi.org/10.1080/13811118.2010.494133>.
- Kao, Kathryn. "Social Media Addiction Linked to Cyberbullying." *UGA Today*, *University of Georgia*, 2021, news.uga.edu/social-media-addiction-linked-to-cyberbullying/.
- Khalaf, Abderrahman M, et al. "The impact of social media on the mental health of adolescents and young adults: A systematic review." *Cureus*, vol. 15, no. 8, 2023, <https://doi.org/10.7759/cureus.42990>.

- Marín Suelves, Diana, et al. "Cyberbullying: Education research." *Education Sciences*, vol. 13, no. 8, 2023, <https://doi.org/10.3390/educsci13080763>.
- Nawaz, Faisal A., et al. "Social media use among adolescents with eating disorders: A double-edged sword." *Frontiers in Psychiatry*, vol. 15, 2024, <https://doi.org/10.3389/fpsy.2024.1300182>.
- Nixon, Charisse. "Current perspectives: The impact of cyberbullying on adolescent health." *Adolescent Health, Medicine and Therapeutics*, vol. 5, 2014, pp. 143–158, <https://doi.org/10.2147/ahmt.s36456>.
- Nobre, Joana, et al. "Psychological vulnerability indices and the adolescent's good mental health factors: A correlational study in a sample of Portuguese adolescents." *Children*, vol. 9, no. 12, 14 Dec. 2022, pp. 1961–1971, <https://doi.org/10.3390/children9121961>.
- O'Keeffe, Gwenn Schurgin, and Kathleen Clarke-Pearson. "The impact of social media on children, adolescents, and families." *Pediatrics*, vol. 127, no. 4, 2011, pp. 800–804, <https://doi.org/10.1542/peds.2011-0054>.
- Patchin, Justin W, and Sameer Hinduja. "Cyberbullying Facts." Cyberbullying Research Center, 2024, <http://www.cyberbullying.org/facts>.
- Pellegrino, Alfonso, et al. "Research trends in social media addiction and problematic social media use: A Bibliometric analysis." *Frontiers in Psychiatry*, vol. 13, 2022, <https://doi.org/10.3389/fpsy.2022.1017506>.
- Racine, Nicole, et al. "Global prevalence of depressive and anxiety symptoms in children and adolescents during COVID-19." *JAMA Pediatrics*, vol. 175, no. 11, 2021, pp. 1142–1150, <https://doi.org/10.1001/jamapediatrics.2021.2482>.
- Reger, Rick. "Teens Tuning In." Local News Initiative, 6 Sept. 2023, <https://localnewsinitiative.northwestern.edu/posts/2023/09/06/medill-teen-news-engagement-survey/>
- Santarossa, Sara, and Sarah J. Woodruff. "#socialmedia: Exploring the relationship of social networking sites on body image, self-esteem, and eating disorders." *Social Media + Society*, vol. 3, no. 2, 2017, <https://doi.org/10.1177/2056305117704407>.
- Savoia, Elena, et al. "Adolescents' exposure to online risks: Gender disparities and vulnerabilities related to online behaviors." *International Journal of Environmental Research and Public Health*, vol. 18, no. 11, 2021, pp. 5786–5803, <https://doi.org/10.3390/ijerph18115786>.
- Shannon, Holly, et al. "Problematic social media use in adolescents and young adults: Systematic review and meta-analysis." *JMIR Mental Health*, vol. 9, no. 4, 2022, <https://doi.org/10.2196/33450>.
- "Social Media and the Internet." American Psychological Association, American Psychological Association, 2024, www.apa.org/topics/social-media-internet.
- Vidal, Carol, et al. "Social Media use and depression in adolescents: A scoping review." *International Review of Psychiatry*, vol. 32, no. 3, 2020, pp. 235–253, <https://doi.org/10.1080/09540261.2020.1720623>.
- Wang, Pengcheng, et al. "Social Networking Sites Addiction and Adolescent Depression: A

moderated mediation model of rumination and self-esteem.” *Personality and Individual Differences*, vol. 127, June 2018, pp. 162–167,
<https://doi.org/10.1016/j.paid.2018.02.008>.

Food Allergies and Evolutionary Mismatch By Naiya Sapra

Abstract

Food allergies affect people all over the world and the total number of allergies has been recently increasing. There have been various hypotheses regarding this increase in food allergies. These hypotheses include the hygiene hypothesis, vitamin D hypothesis, dual allergen exposure hypothesis, and the false alarm hypothesis. This paper will focus on the hygiene hypothesis, which suggests that modern, highly sanitized environments inhibit the immune system's typical development, leading to allergic reactions to harmless substances. Past environments had a large impact on the evolution of immune system development and thus allergies can be thought of as an evolutionary mismatch. There are current treatments for allergic reactions, such as oral immunotherapy and sublingual immunotherapy, but in addition to developing other therapies, existing treatments need to be made more accessible globally. This paper will address the question of whether the rise in food allergies has been caused by an evolutionary mismatch between adaptations of immune systems to past environments and hygienic environments in the present.

Introduction

Food allergies have grown more common in recent years, as the number of children with food allergies increased by 50% between 1997 and 2011 (Turke). An allergic reaction to food happens when a person is exposed to a certain protein or set of proteins in food that bind to a type of antibody called immunoglobulin E (IgE; National Institute of Allergy and Infectious Diseases). This event signals cells to release different chemicals, which produce the symptoms of an allergic reaction (National Institute of Allergy and Infectious Diseases). There has been debate among scientists and medical professionals about the exact causes of allergies, leading to the proposal of several hypotheses. For instance, the hygiene hypothesis explains how the human immune system, which evolved in less hygienic environments of the past, is prone to overreacting to non-harmful allergens.

The hygiene hypothesis originally emerged from a study done in 1989 that followed British children from birth until the age of 23 and recorded the total number of children in each household and the instances when each child had hay fever (Strachan). The results showed that for smaller families with fewer children, there were more instances of hay fever compared to larger families with more children (Strachan). Larger families were more likely to be exposed to bacteria and other microbes, since there were more people that could potentially expose each other to various microbes. Since larger families were more exposed to various microbes, the immune systems of the members of these families developed appropriately such that the risk of hay fever was mitigated. In more recent years, this study has been used to attribute the rising numbers of allergies to modern, hygienic environments. Another study suggested that the absence of immune-stimulating microbes in contemporary environments causes the immune system to behave abnormally and develop allergic diseases (Pfefferle et al.).

The hygiene hypothesis can be thought of as a form of evolutionary mismatch. Evolutionary mismatch takes place when a trait that was beneficial to humans in the past is no longer helpful and may even harm humans. The immune system was shaped by less hygienic conditions in past environments [across human history/since human evolution in Africa] and the immune system evolved to respond in order to fight diseases present at the time. However, with increased public health initiatives in modernity, environments have become more hygienic than in the past and the immune system has not been properly exposed to various microbes. As a result, the immune system overreacts and tries to fight infectious stimuli by reacting to harmless materials, even though they are not present in hygienic environments. This paper will address the question of whether the rise in food allergies has been caused by an evolutionary mismatch between adaptations of immune systems to past environments and hygienic environments in the present.

The Hygiene Hypothesis

One of the more popular hypotheses, the hygiene hypothesis, involves exposure to microbes or parasites, such as helminths (Haspeslagh et al.). Helminths can cause the immune system to react through the modified Th2 response (Haspeslagh et al.). Regulatory T cells help control and adjust how our immune system responds to antigens and prevents it from overreacting to harmless foreign matter (Haspeslagh et al.). When the body is exposed to certain microbes or parasites, T cells in the body should mount an appropriate immune response, as opposed to what occurs during an allergic reaction. The hygiene hypothesis suggests that without such exposure, the immune system is at risk of incorrectly responding to harmless substances. In addition, with the absence of helminths and other microbes, the total level of IgE antibodies decreases in the blood and the proportion of IgE antibodies that react to harmless substances increases (Allaerts and Chang). This phenomenon is because the IgE antibodies need to reach a certain density in order to cause an allergic reaction (Allaerts and Chang). This paper suggests that with decreased exposure to parasites and microbes, the body produces more IgE antibodies that are more likely to react to harmless substances, which could explain the presence of food allergies.

Various papers have provided additional evidence for this hypothesis since it was first proposed by Strachan. For example, one study found that children with parents who were farmers had a decreased risk of children becoming atopic and developing symptoms of allergic rhinitis, also known as hay fever (Braun-Fahrlander et al.). Both of these studies suggest that exposure to more microbiota can be beneficial and prevent various allergic diseases, including food allergies. A more recent study found that there was a very low prevalence of allergic sensitization in Amish children living in northern Indiana (Holbreich et al.). Amish families have a mean number of 5.9 children per family, and they mainly consume things that they directly produced (Holbreich et al.). This suggests that certain microbes present on their farms contribute to the low numbers of Amish children with allergic diseases.

Past Environments and the Immune System

Past environments have played an important role in human adaptation. Depending on the environment, the immune system has to respond and adapt so that humans are able to survive and reproduce. Since the human species first appeared in Africa, populations have faced a variety of selective pressures arising from changes in the environment (Hancock et al.). Some of these selective pressures include diverse diets, pathogens, altitudes, and ambient temperatures (Rees et al.). The immune system has adaptive mechanisms that allow vertebrates to respond to protect the body from potential infectious diseases (James et al.).

Around 10,000 years ago, groups of humans in the Near East began to farm and domesticate animals (Dow et al.). Previously, humans obtained their food by hunting animals or gathering plants in the wild (Dow et al.). The major shift to agriculture brought about many changes to human populations that adopted agriculture. For example, in China, India, and West Africa, populations increased exponentially and more villages and settlements began to form (Fuller et al.). Humans started settling down in villages and eventually larger cities and the population density increased. For instance, during the Industrial Revolution, people flooded into overcrowded cities and started to develop health problems from all of the smog and sewage, which led to increases in diseases such as cholera and dysentery and poor sanitation in many places (Davenport et al.). During this time, human immune systems had to respond to the conditions in the cities to be able to fight diseases and survive. The immune system adapted to specific environments in the past and this continues to impact our immune systems today.

In past, less hygienic environments, there were increased amounts of parasites, such as helminths (Rook). Scientists have hypothesized that these parasites traveled into our gut microbiomes and helped make our immune systems more tolerant. For instance, in one study, animal models indicated the existence of a window early in life during which exposure to microbes resulted in acquisition of tolerance (Stiemsma and Turvey). In modern, hygienic environments, the amount of parasites has decreased and scientists think it is possible that this is responsible for the overreaction of our immune systems. Without exposure to parasites, there is an imbalance of microbiota, breaking gut homeostasis and contributing to diseases (De Filippis et al.). This imbalance may also affect the integrity of the intestinal epithelial barrier, leading to the entry of antigens in the bloodstream and the abnormal stimulation of the immune system (De Filippis et al.). Our immune system is used to protecting our bodies from parasites and disease, so even with the lack of parasites and other microbiota in our guts in present environments, the immune system still responds. This causes the allergic reactions to different foods and environmental factors.

Food Allergy Hypotheses

Despite lots of evidence for the hygiene hypothesis, there are other hypotheses concerning the causes of food allergies. One of them is the vitamin D hypothesis, which argues that vitamin D deficiency increases the risk of developing food allergies (Allen et al.). One study found that infants with vitamin D deficiency were three times more likely to have an egg allergy

and eleven times more likely to have a peanut allergy (Allen et al.). Another hypothesis, known as the dual allergen exposure hypothesis, advances that the immune system reacts differently depending on the way that an individual is exposed to the allergen (Kulis et al.). For example, it has been shown that exposure through an impaired skin barrier can cause allergies, while oral exposure can lead to tolerance against the allergen (Kulis et al.). The false alarm hypothesis says that food allergies are caused by modern diets (Smith et al.). Modern diets in industrialized countries are enriched in advanced glycation endproducts (AGEs), and it has been proposed that the immune system mistakes these components for allergens (Smith et al.). AGEs are compounds that are formed when proteins and fats combine with sugars under high temperatures (Smith et al.). Modern cooking in fryers and microwaves and at fast food restaurants has contributed to these high temperatures and increased the amount of AGEs present in our diets (Smith et al.). These various hypotheses attempt to explain the causes of food allergies, so scientists can work on ways to help prevent them.

Current Treatments

There are currently several treatments for people with food allergies to utilize. One of these treatments is Oral Immunotherapy (OIT) and it involves ingesting the allergenic food in small doses and gradually increasing those doses until the individual is tolerant (Wood). However, one limitation of this treatment is that it must be carefully overseen by a professional to prevent the patient from ingesting too much at one time. Another treatment is Sublingual Immunotherapy (SLIT) and patients are treated by placing a small amount of the allergenic food under their tongue, which unlike OIT is not consumed (NIAID). One study performed different trials to evaluate if SLIT was safe and effective in peanut-allergic children (Kim et al.). The study found that the peanut SLIT was safe and the children continued to stay tolerant more than 17 weeks after the trials were discontinued (Kim et al.). While SLIT has fewer side effects than OIT, it is not as effective for higher doses of an allergen, so researchers are still working on ways to improve the treatment (NIAID).

Conclusion

In conclusion, food allergies remain a large problem around the world among people of all ages. The exact causes of these food allergies have been debated, leading to several hypotheses, including the popular hygiene hypothesis. Other hypotheses include the vitamin D deficiency, dual allergen exposure, and the false alarm hypothesis as potential causes of food allergies. Evolutionary factors, such as the shift from less hygienic past environments to modern hygienic ones, also play a role in how the immune system reacts today. Exposure to parasites and microbes in the past may have helped the immune system develop tolerance, while the lack of exposure in modern environments may have contributed to the rise in allergies.

Despite ongoing research on allergies, there will need to be more research for certain treatments to be used by the general public. Scientists need to perform research on possible ways to stimulate the immune system to mimic its typical development and prevent allergies.

Additionally, for treatments that already exist, they need to become more globally accessible. Many of the existing treatments mentioned above are extremely costly and can only be found in more affluent parts of the world with the correct medical supervision. Research needs to be done to find solutions to help those who do not have access to these treatments by working to make them more accessible. This should be paired with appropriate government policies to enact these solutions for increased equity in allergy treatments. Scientists also need to continue to develop medicines or treatments to act as preventative measures and eventually end childhood food allergies. Together, the advancement of evidence-based allergy treatment will allow treatments to be accessed by everyone and will eventually prevent food allergies altogether.

Works Cited

- Allaerts, Wim, and T. W. Chang. "Skewed Exposure to Environmental Antigens Complements Hygiene Hypothesis in Explaining the Rise of Allergy." *Acta Biotheoretica*, vol. 65, no. 2, 2017, pp. 117–134. <https://doi.org/10.1007/s10441-017-9306-7>.
- Allen, Katrina J., et al. "Vitamin D Insufficiency Is Associated with Challenge-Proven Food Allergy in Infants." *The Journal of Allergy and Clinical Immunology*, vol. 131, no. 4, 2013, pp. 1109–1116. e11166. <https://doi.org/10.1016/j.jaci.2013.01.017>.
- Braun-Fahrländer, Charlotte, et al. "Prevalence of Hay Fever and Allergic Sensitization in Farmer's Children and Their Peers Living in the Same Rural Community." *Clinical and Experimental Allergy*, vol. 29, no. 1, 1999, pp. 28–34. <https://doi.org/10.1046/j.1365-2222.1999.00479.x>.
- Davenport, Romola J., Mark Satchell, and Leigh M. W. Shaw-Taylor. "Cholera as a 'Sanitary Test' of British Cities, 1831-1866." *The History of the Family*, vol. 24, no. 2, 2018, pp. 404–438. <https://doi.org/10.1080/1081602X.2018.1525755>.
- Dow, Gregory K., Nancy Olewiler, and Chris G. Reed. *The Transition to Agriculture: Climate Reversals, Population Density, and Technical Change*. Simon Fraser University, 10 Sept. 2007. Retrieved 13 Aug. 2024, from https://www.sfu.ca/~reed/download/agriculture_08.pdf.
- De Filippis, Francesca, et al. "Specific Gut Microbiome Signatures and the Associated Pro-Inflammatory Functions Are Linked to Pediatric Allergy and Acquisition of Immune Tolerance." *Nature Communications*, vol. 12, no. 1, 2021, p. 5958. <https://doi.org/10.1038/s41467-021-26266-z>.
- Fuller, Dorian Q., et al. "Comparing Pathways to Agriculture." *Archaeology International*, vol. 18, 23 Nov. 2015. Retrieved 22 Aug. 2024, from <https://journals.uclpress.co.uk/ai/article/id/2051/>.
- Hancock, Angela M., et al. "Adaptations to New Environments in Humans: The Role of Subtle Allele Frequency Shifts." *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 365, no. 1552, 2010, pp. 2459–2468. <https://doi.org/10.1098/rstb.2010.0032>.
- Haspeslagh, Esther, et al. "The Hygiene Hypothesis: Immunological Mechanisms of Airway Tolerance." *Current Opinion in Immunology*, vol. 54, 2018, pp. 102–108. <https://doi.org/10.1016/j.coi.2018.06.007>.
- Holbreich, Mark, et al. "Amish Children Living in Northern Indiana Have a Very Low Prevalence of Allergic Sensitization." *The Journal of Allergy and Clinical Immunology*, vol. 129, no. 6, 2012, pp. 1671–1673. <https://doi.org/10.1016/j.jaci.2012.03.016>.
- James, W. P. T., et al. "Nutrition and Its Role in Human Evolution." *Journal of Internal Medicine*, vol. 285, no. 5, 2019, pp. 533–549. <https://doi.org/10.1111/joim.12878>.
- Kim, Emily H., et al. "Open-Label Study of the Efficacy, Safety, and Durability of Peanut]

- Sublingual Immunotherapy in Peanut-Allergic Children." *The Journal of Allergy and Clinical Immunology*, vol. 151, no. 6, 2023, pp. 1558–1565.e6.
<https://doi.org/10.1016/j.jaci.2023.01.036>.
- Kulis, Edwin D., et al. "The Airway as a Route of Sensitization to Peanut: An Update to the Dual Allergen Exposure Hypothesis." *The Journal of Allergy and Clinical Immunology*, vol. 148, no. 3, 2021, pp. 689–693. <https://doi.org/10.1016/j.jaci.2021.05.035>.
- Metcalf, Charlotte J. E., et al. "Hosts, Microbiomes, and the Evolution of Critical Windows." *Evolution Letters*, vol. 6, no. 6, 2022, pp. 412–425. <https://doi.org/10.1002/evl3.298>.
- National Institute of Allergy and Infectious Diseases (NIAID). Causes and Prevention of Food Allergy. 10 June 2024. Retrieved 13 Aug. 2024, from <https://www.niaid.nih.gov/diseases-conditions/food-allergy-causes-prevention>.
- National Institute of Allergy and Infectious Diseases (NIAID). Food Allergy Treatment Research. 23 May 2024. Retrieved 14 Aug. 2024, from <https://www.niaid.nih.gov/diseases-conditions/food-allergy-treatment-research>.
- Pfefferle, Pia I., et al. "The Hygiene Hypothesis—Learning from but Not Living in the Past." *Frontiers in Immunology*, vol. 12, 2021, article 635935.
<https://doi.org/10.3389/fimmu.2021.635935>.
- Rees, Jonathan S., et al. "The Genomics of Human Local Adaptation." *Trends in Genetics*, vol. 36, no. 6, 2020, pp. 415–428. <https://doi.org/10.1016/j.tig.2020.03.006>.
- Rook, Graham A. "Regulation of the Immune System by Biodiversity from the Natural Environment: An Ecosystem Service Essential to Health." *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 46, 2013, pp. 18360–18367. <https://doi.org/10.1073/pnas.1313731110>.
- Smith, Peter K., et al. "The False Alarm Hypothesis: Food Allergy Is Associated with High Dietary Advanced Glycation End-Products and Pro-Glycating Dietary Sugars That Mimic Alarmins." *The Journal of Allergy and Clinical Immunology*, vol. 139, no. 2, 2017, pp. 429–437. <https://doi.org/10.1016/j.jaci.2016.05.040>.
- Stiemsma, Leah T., and Stuart E. Turvey. "Asthma and the Microbiome: Defining the Critical Window in Early Life." *Allergy, Asthma, and Clinical Immunology*, vol. 13, 2017, article 3. <https://doi.org/10.1186/s13223-016-0173-6>.
- Strachan, David P. "Hay Fever, Hygiene, and Household Size." *BMJ*, vol. 299, no. 6710, 1989, pp. 1259–1260. <https://doi.org/10.1136/bmj.299.6710.1259>.
- Turke, Paul W. "Childhood Food Allergies: An Evolutionary Mismatch Hypothesis." *Evolution, Medicine, and Public Health*, vol. 2017, no. 1, 2017, pp. 154–160.
<https://doi.org/10.1093/emph/eox014>.
- Wood, Robert A. "Food Allergen Immunotherapy: Current Status and Prospects for the Future." *The Journal of Allergy and Clinical Immunology*, vol. 137, no. 4, 2016, pp. 973–982.
<https://doi.org/10.1016/j.jaci.2016.01.001>.

A Descent into the Dark Watery Mud By Jonah Miller

In "Tits-up in a Ditch," Annie Proulx's version of rural Wyoming is very narrow-minded and traditional. The people of the town maintain rigid gender expectations that combine with economic problems to make many of them unhappy. Although the population lives under powerful limitations, there are still a few people who pursue change. This change and innovation, however, is dismissed by a combination of a lack of resources and resistance from economically powerful members of the community. The majority of the population in Proulx's version of rural Wyoming is poor or resides in the middle class, and they resent the power of large corporations. At the same time, they pride themselves on working hard. Many of the characters exhaust their bodies performing intense ranch labor or similarly intense work. Although men and women both contribute to this work, the men receive more credit for it, while the women must simultaneously uphold family obligations. While both the male and female characters in Annie Proulx's "Tits-up in a Ditch" experience the difficulties of living in a deteriorating rural community defined by lack of opportunity, the women in the story have the added burden of having to mourn losses while trying to meet their obligations as wives and mothers.

The women born into the rural Wyoming community of Proulx's short story must deal with a male-dominated culture that undervalues their labor while also expecting them to make up for the faults of the men in their lives. One defining aspect of the community Proulx explores is gender hierarchy. Men are viewed as the head of the household, responsible for performing heavy ranch labor and ensuring their families' financial security. Women, on the other hand, are viewed primarily as mothers and wives. Despite this traditional perception of what gender roles should be, the reality is that women work as much or not more than men. Proulx writes that Wyoming ranch women hold "everything together...cooking for big crowds, nursing the sick and injured, cleaning, raising children...paying the bills...[and] riding with the men at branding and shipping time" (23). This list of responsibilities includes a mix of tasks that are associated with both masculinity and femininity. Women are still seen as second-class citizens, however. Despite their hard work both inside and outside the home, Proulx suggests that women are treated with "little more regard than the beef they helped produce" (23). This treatment is dehumanizing and demoralizing for women who are working hard and suffering alongside men but are not recognized for it. Bonita is an example of a woman who is forced to satisfy the responsibilities that Wyoming culture usually puts on men while also trying to meet the expectations to be a good wife and mother. She has to deal with an "impaired husband...endless labor...[and] a bad-girl daughter," all while displaying the "good humor" and pleasantness that is expected of a woman (3). In other words, she has to deal with the suffocating responsibilities of both men and women, and has to smile while carrying the burden. While her husband, who is supposed to display the masculine ability to "endure pain...and not bitch about it all day," deals with various injuries, Bonita has to be the one to "cowboy up" and handle hard labor even though she has "arthritis in her left knee" (3). Bonita seems similar to many women born into the rural Wyoming

community, who have to "suffer in silence" as they do more than men yet receive less recognition for it (3).

The young men raised in Proulx's rural Wyoming are expected to live up to an idea of toughness and resourcefulness even though the opportunities available to them often break their bodies down and strip them of their masculinity. Just as women in rural Wyoming have certain expectations to live up to, men also meet old-fashioned and potentially destructive expectations. Boys born into this rural Wyoming community are taught to embody the classic image of a cowboy, and be masculine, tough, and willing to engage in hard physical labor without complaining. These expectations don't leave a lot of room for individuality and imagination, something we see in Verl, whose "secret boyhood dream had been to become a charismatic radio man, meeting singing personalities, giving the news, announcing songs, describing the weather" (5). Proulx describes Verl's dream as "secret," emphasizing the embarrassment that he would feel if anyone knew he wanted to be something other than the stereotypical cowboy. Instead, he sets out on a career on the ranch, a life that is long, painful, and gruesome. While he probably enjoyed moments of this life, mostly before he suffered a string of injuries, he never got a chance to explore anything else. Even if Verl had the courage to attempt to follow his dreams of becoming a charismatic radio man, "He had no idea how to get into the radio game, and the plan faded as he grew into work on the home ranch" (5). The limitations placed on Verl by the culture of his town limit his ability to pursue the kind of life he wants. Verl and other men like him are born into a community with a very narrow sense of what life can be. When they are boys, they may have had interesting and unconventional dreams, but over time, Wyoming life puts them on a path of physical pain and shattered dreams.

While the world of "Tits-up in a Ditch" is not easy for either gender, women in the story are forced to bear the burden of loss and suffering while also living up to the social expectations of being wives and mothers. While everyone in Proulx's story experiences trauma, the culture of rural Wyoming is much harsher toward women. For example, when Dakotah is without a husband, out of a job, and extremely pregnant, she contemplates making the reasonable decision to "put the baby up for adoption" (20). Instead of acknowledging Dakotah's complicated and emotionally difficult situation, Bonita, a woman who has dealt with tragedy and loss in her own life, winds up harshly criticizing Dakotah, stating, "You can't bring anymore shame on this family" (21). Instead of focusing on Dakotah's personal suffering, Bonita focuses on how the community will perceive Dakotah's decision. Unlike men, who seem able to walk away from their responsibilities without shame or social stigma, women are expected to keep up with the obligations expected of them. Another example of this burden comes at the end of the story, when Dakotah loses Baby Verl. At this point, she feels loss mixed with guilt, even though Dakotah joined the army to provide her baby with a better life, the same decision required that the baby be left with an irresponsible man, Verl. As she is trying to process this loss, she is persistently asked by her mother-in-law for information on Sash. While the request for information itself is not so awful, when the mother-in-law tries to persuade Dakotah by stating, "You're his wife," she highlights how this society perceives women (37). Even though Sash has

not been a husband to Dakotah or a father to Baby Verl, the people within the community expect her to push down her understandable grief and anger to act like a "wife." The absurdity of this moment makes Dakotah feel "her hooves slip as she began her descent into the dark, watery mud," which alludes to her feeling that she is dying. Ultimately, this may be the way many of the women born into this community feel by the time they are adults.

Although there are examples of small victories or minor versions of success throughout "Tits-up in a Ditch," overall, no one thrives. The town embraces an old-fashioned ideology that resists change even as it acknowledges that traditional life is difficult and unpleasant. This commitment to living in the past is exacerbated by the poverty in the town, creating generational cycles of unhappiness. Like the cattle who end up killed or "tits-up in a ditch," the characters in this story are doomed to an unhappy end by the social and economical systems in their town. Whether it is particular moments of intense struggle or just an ongoing miserable life, the people of rural Wyoming cannot thrive where they are born. Their labor taxes their bodies, as well as the expectations and judgment of others. Although almost no one seems to be happy, the women have it worse, because the town's economic state requires them to work like men while upholding the family obligations of women, the same obligations that men are often excused from. Ultimately, this story emphasizes the power of expectations and how they can shape individual lives.

Analyzing the Impact of Climate Risk on Inflation: A Comprehensive Statistical Review

By Alice Zhang

Abstract

This study employs regression and scatter plot analysis to examine the relationship between inflation and climate risk, while accounting for various macroeconomic variables. Inflation, climate risk, GDP growth, and exchange rates have a robust positive correlation in the scatter plots, whereas wide money, unemployment, and GDP per capita reveal a negative correlation. Consistent and statistically significant findings across various models illustrate the considerable influence of climate risk on inflation in regression analysis.

Keywords inflation, climate risk, regression analysis, scatter plot, macroeconomic variable

Introduction

The determinant of inflation is important because it affects economic stability and individual and company behavior. The motivation behind the study lies in understanding what drives inflation, which helps policymakers design effective strategies to control price increases and ensure economic stability. It is clear that identifying the determinant of inflation provides insights for better economic planning, financial management, and policymaking. Existing literature typically categorizes the determinants of inflation into several key factors, including unemployment, GDP growth, exchange rate, oil rents, and broad money. This study examines the effect of climate risk on inflation, introducing a new determinant. Comprehending this connection is essential for developing effective policies. We utilize data from the World Bank to examine the correlation between inflation and economic indices, including GDP growth, exchange rates, money supply, unemployment, and GDP per capita. Our objective is to provide a comprehensive understanding of the principal factors affecting inflation by employing regression models and scatter graphs. The paper highlights the critical impact of climate risk on inflation, emphasizing its importance in the current economic landscape. It provides empirical evidence that there is a correlation between inflation and climate risk, thereby enabling policymakers to make more effective monetary and fiscal policy decisions. This research bridges the divide between environmental factors and economic indicators, contributing to the expanding literature on inflation and environmental factors. It is essential for businesses, policymakers, and individuals to comprehend the correlation between inflation and climate risk, as it is a critical factor in promoting economic stability.

This paper proceeds as follows: Section 2 provides background information on the existing literature. Section 3 highlights the empirical findings. Section 4 concludes the paper.

Literature Review

Recent studies have increasingly focused on the relationship between climate change and economic factors across various regions. Li et al. (2023) conducted an analysis of data from 26

nations worldwide and identified a positive correlation between inflation and temperature variation. Almajali (2023) investigated the diverse impacts of climate risk on inflation and GDP growth in 173 countries, underscoring a significant correlation between climate-related events and numerous macroeconomic variables. Extreme weather shocks have a substantial impact on inflation rates, notably in agricultural states, according to Liao et al. (2024). Luo et al. (2023) found that specific weather variables could improve inflation forecasts, underscoring the absence of a universally prevalent climate risk factor. In the European context, Safonov (2023) examines the correlation between escalating carbon prices and heightened electricity prices in the UN. While industrial power costs originally had no effect, Selmi (2023) indicates that investors are directing greater resources towards gold and copper rather than oil following the Paris Agreement.

As a result, the dependence between inflation and industrial metals continues to hold. Odongo et al. (2022) highlight the significant effects of weather patterns on food prices and overall inflation in Eastern and Southern Africa. Lyu (2019) analyzed the existing relationship between the green economy and the pollution present in China. Findings showed that speeding up both factors would accelerate reaching the climate goals for 2020 and lay the foundation for 2030's respective goals.

Konradt and di Mauro (2021) analyzed the impact of carbon taxes on inflation over a long period in Canada and Europe. Their findings suggested carbon taxes didn't increase the overall price level; instead, they had a slight negative impact on price.

Aguila and Wullweber (2024) emphasized that rising interest rates increase the cost of sustainable investments, delaying the green transition. Moreover, monetary policies should be greener and cheaper to address possible environmental crises and inflation, respectively.

Kunawotor et al. (2022) investigate the positive influence of climate change events on inflation in Africa from 1990 to 2017, emphasizing the importance of monetary policy. From 2005 to 2020, Yusifzada (2022) investigates the influence of climate change on inflation in Azerbaijan. They underscore the potential for climate change to contribute 1.3% to inflation in a normal scenario and 2.2% in a worst-case scenario by 2030. The results demonstrate the necessity for policymakers to establish a climate action plan.

Wahidah and Antriyandarti (2021) investigate the impact of climate change on inflation in Indonesia during the Covid-19 period. They demonstrate that COVID-19 cases adversely affected inflation, whereas recovered cases had a beneficial impact. Climate change also adversely affected inflation in the agricultural industry.

Zhang (2023) examines the influence of climate risk on inflation in BRICS nations. The results demonstrate that climate risk directly influences inflation in China, India, and Russia. In the largest Eurozone countries, Ciccarelli et al. (2024) examine the asymmetric impacts of temperature shocks, which serve as a proxy for climate change, on inflation. The results indicate that temperature fluctuations substantially influence inflation.

Bistline et al. (2023) evaluate the effect of the Inflation Reduction Act (IRA), the largest US climate response, emphasizing that the IRA reduces inflation but faces challenges from

higher interest rates and material costs. Boneva and Ferrucci (2022) emphasize the role of climate variables for forecasting inflation.

Moessner (2022) studies the relationship between inflation and climate policy between 55 countries. The results indicate that inflation is reduced by improved climate policies, while inflation is not increased by reduced carbon dioxide emissions. Dafermos et al. (2024) examine the influence of weather-related calamities on inflation in the Euro area. The results show that weather-related disasters positively influence inflation, albeit with a non-permanent effect. As climate crises become more severe, it is more difficult to control inflation.

Empirical Findings

Data

Table 1 in the appendix presents inflation (annual consumer prices), climate risk (CO2 emissions per 2015 USD of GDP), and various economic variables sourced from World Bank-World Development Indicators. The economic variables include unemployment (total % of the labor force), GDP growth (annual%), oil rents (% of GDP), exchange rate (real effective exchange rate), and broad money (% of GDP).

Variables	Observation	Mean	Std. Dev	Min	Max
Inflation	217	3.5469	7.4971	-2.5952	84.8643
Exchange Rate	209	785.9119	3670.465	0.3062	42000
Broad Money	185	79.4570	55.7692	0.0295	454.7032
Oil Rents	243	1.5651	4.0555	0	27.5816
Climate Risk	234	0.4565	0.3274	0.0477	1.6884
Unemployment	143	7.7408	5.0593	0.14	25.895
GDP Growth	257	-4.8575	8.3774	-58.3182	43.4796
GDP Per Capita	255	14667.95	21169.52	263.361	166822.6

Table 1. Descriptive Statistics

This table provides descriptive statistics for various macroeconomic and inflation variables used in the analysis. The inflation rate has 217 observations with a mean of 3.5469 and a standard deviation of 7.4971, ranging from a minimum of -2.5952 to a maximum of 84.8643. The exchange rate has 209 observations, with a mean of 785.9119 and a standard deviation of 3670.465, ranging from a minimum of 0.3062 to a maximum of 42000. Broad Money has 185 observations with a mean of 79.4570 and a standard deviation of 55.7692, ranging from a minimum of 0.0295 to a maximum of 454.7032. Oil rents have 243 observations with a mean of 1.5651 and a standard deviation of 4.0555, ranging from a minimum of 0 to a maximum of 27.5816. Climate Risk has 234 observations with a mean of 0.4569 and a standard deviation of 0.3274, ranging from a minimum of 0.0477 to a maximum of 1.6884. Unemployment has 143 observations with a mean of 7.7408 and a standard deviation of 5.0593, ranging from a minimum of 0.14 to a maximum of 25.895. GDP Growth has 257 observations with a mean of -4.8575 and a standard deviation of 8.3774, ranging from a minimum of -58.3182 to a maximum of 43.4796. GDP per capita has 255 observations with a mean of 14667.95 and a standard deviation of 21169.52, ranging from a minimum of 263.361 to a maximum of 166822.6.

Findings

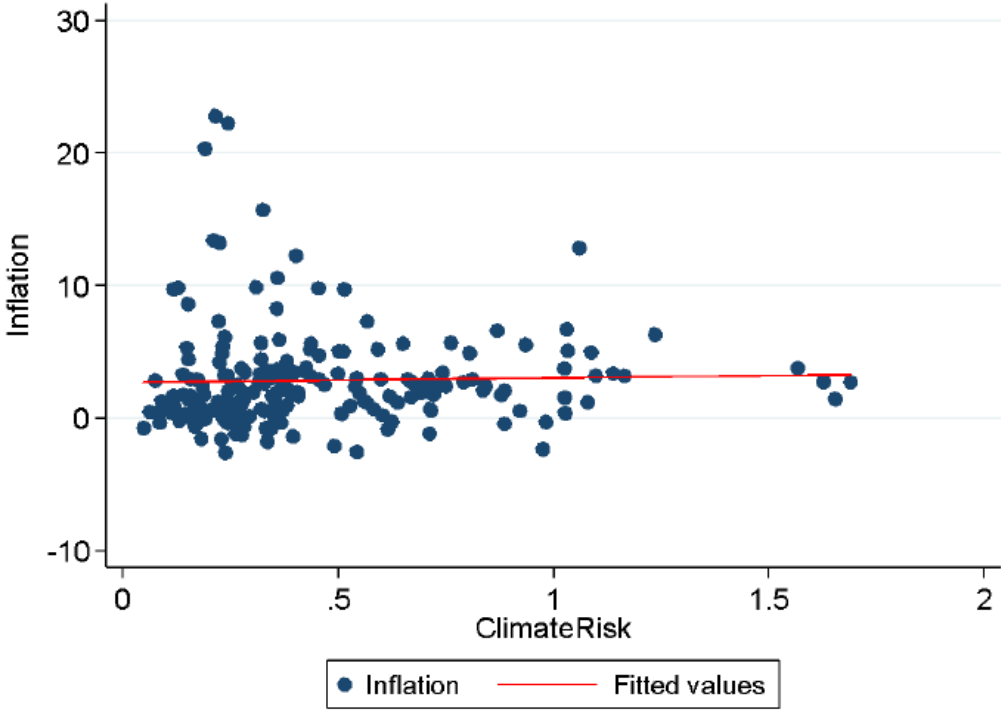
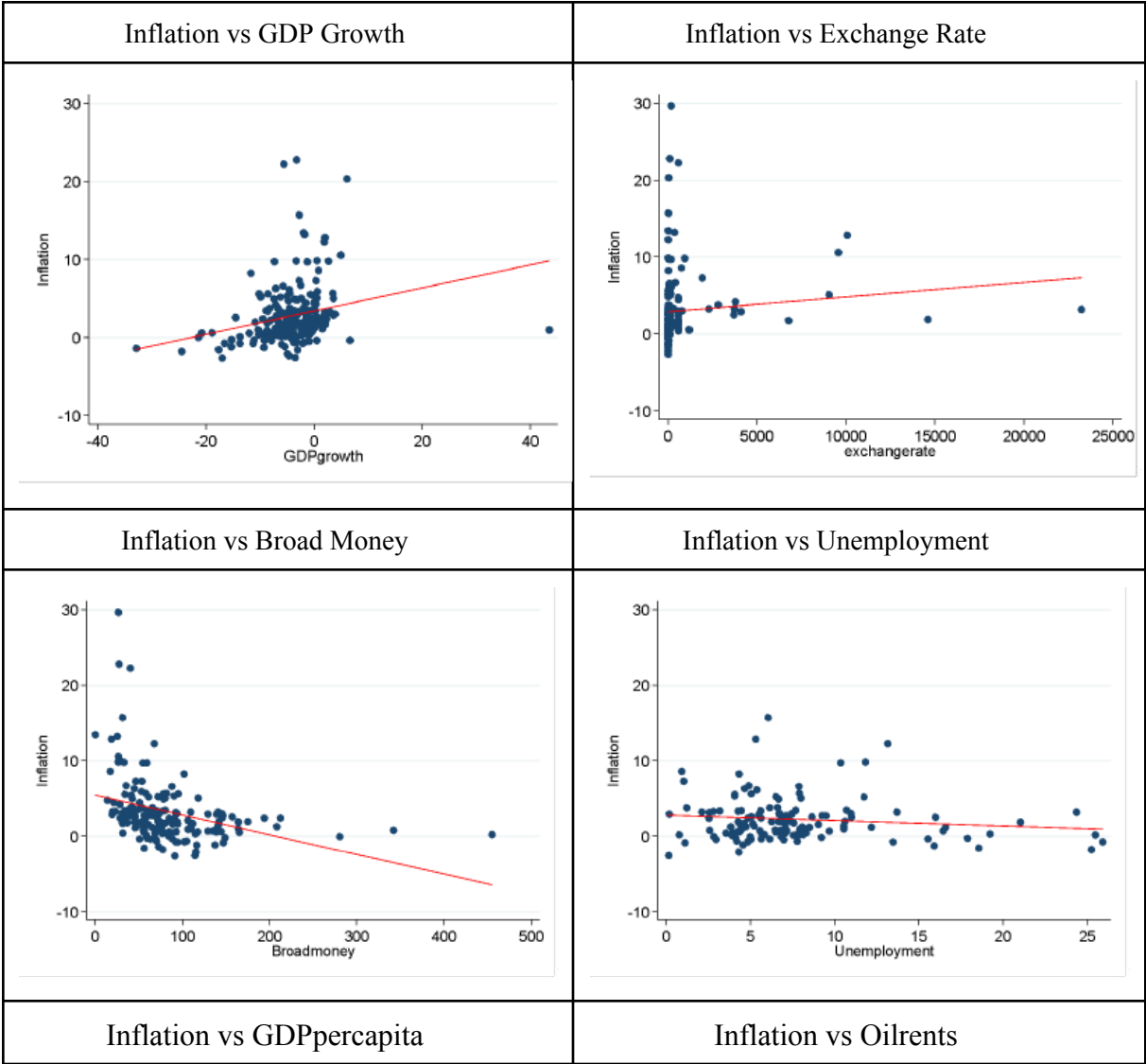


Figure 1a. Scatter Plot Figure of Inflation vs Climate Risk

The scatter figure shows a strong correlation between inflation, climate risk, and numerous economic indices. We expect increased climate risk to lead to higher inflation due to the significant connection between the two factors.



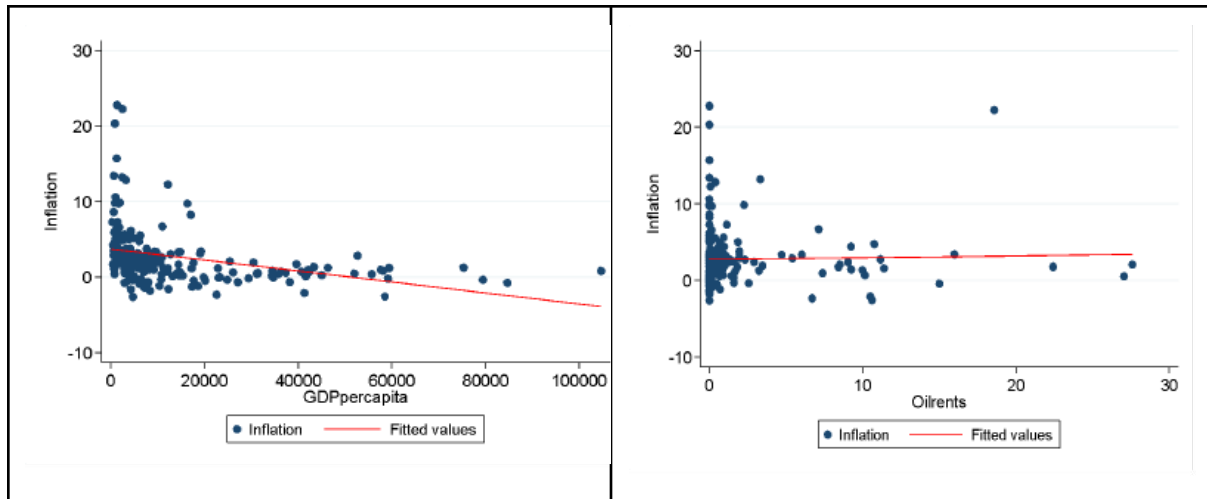


Figure 1b: Scatter Plots between Inflation and Economic Variables

Inflation positively correlates with GDP growth, suggesting that a higher level of GDP growth might contribute to higher inflation. Inflation has a positive correlation with exchange rates, suggesting that higher exchange rates may contribute to higher inflation.

Inflation negatively correlates with Broad Money, suggesting that higher levels of Broad Money might contribute to lower inflation. There is a slight negative correlation between inflation and unemployment, which implies that higher levels of unemployment may contribute to a slight decrease in inflation.

The negative correlation between inflation and GDP per capita suggests that a higher GDP per capita could potentially reduce inflation. Inflation slightly positively correlates with oil rents, suggesting that higher levels of oil rents might contribute to a slight increase in inflation.

Model

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_n X_{ni} + \epsilon_i$$

Where Y_i is the dependent variable for observation i , which refers to inflation. When all independent variables are zero, the constant term β_0 represents the expected value of the dependent variable. β_1 to β_n are the coefficients for the independent and control variables, which include climate risk, unemployment, GDP growth, oil rents, exchange rate, and broad money. While keeping all control variables constant, the coefficient of β_1 shows how much inflation change one climate risk variable changes by 1 unit. ϵ_i is the error term.

Inflation	Model-1	Model-2	Model-3	Model-4	Model-5
<i>ClimateRisk</i>	2.0797*** (0.6939)	1.8448 ** (0.7147)	2.2371*** (0.7505)	2.2917*** (0.8261)	0.7864 (0.9028)
<i>Unemployment</i>	-0.5250 (0.0545)	-0.00851 (0.0625)	-0.0315 (0.0626)	-0.0262 (0.0639)	-0.0647 (0.0619)

<i>GDPgrowth</i>		0.1155 (0.0723)	0.0106 (0.0724)	0.1331* (0.0796)	0.1548* (0.0889)
<i>Oilrents</i>			-0.1902* (0.0970)	-0.1980* (0.1002)	-0.3576*** 0.1062
<i>exchangerate</i>				0.000031 (0.000111)	-0.00001 0.00009
<i>Broadmoney</i>					-0.0327*** 0.0084
<i>Constant</i>	1.7538*** (0.5600)	2.0542*** (0.5675)	2.1970*** (0.5695)	2.3849*** (0.6513)	6.9032*** 1.3408
<i>R²</i>	0.0636	0.0875	0.1141	0.1193	0.2904
<i># of observations</i>	129	129	128	102	74

Table 2. Regression Analysis

The regression analysis in Table 2 shows that climate risk has a consistently positive and statistically significant impact on inflation, with p-values less than 0.05 in all models except Model 5, demonstrating its substantial contributions to inflation. Oil rents also have a negative influence on inflation, with several models indicating a 10-percent impact. GDP growth has a positive and significant impact on inflation, with p-values less than 0.1 in Models 4 and 5. In Model 5, broad money has a negative and statistically significant influence on inflation, with p-values less than 0.01. Table 1 indicates that neither unemployment nor exchange rates have a statistically significant effect on inflation.

Conclusion

This study highlights the significant relationship between climate risk and inflation, demonstrating that climate risk has a positive effect on inflation across various regression models. The results indicate that inflation increases in line with rising climate risk. Money supply and oil rents highlight the importance of considering climate risk in inflation dynamics. To minimize inflation risk, policymakers take environmental factors into account in economic decision-making. This extensive statistical analysis provides significant insight into how inflation is understood within the context of climate change.

This paper makes a significant contribution to the existing literature by offering a comprehensive statistical analysis, integrating climate risk into the study of inflation dynamics across various macroeconomic variables. Unlike previous studies, which often focus on specific regions or climate risk factors, this study implies a broader data set and multiple regression models to investigate the relationship between inflation and climate risk on a global scale. The paper is not only considering the positive correlation between climate risk and inflation but also identifies the effects of other macroeconomic variables such as GDP growth, broad money supply, and oil rents. This study examines the influence of climate risk on inflation, introducing a new viewpoint on inflation dynamics within the framework of climate change and providing

policymakers with enhanced insights into the ways environmental issues might affect economic results. This research bridges the gap between environmental and economic studies.

Works Cited

- Li, C., Zhang, X., & He, J. (2023). Impact of Climate Change on Inflation in 26 Selected Countries. *Sustainability*, 15(17), 13108
- Almajali, K. A. D. (2023). The Influence of Climate Disturbances on Growth and Inflation. *Migration Letters*, 20(S10), 148-157
- Liao, W., Sheng, X., Gupta, R., & Karmakar, S. (2024). Extreme weather shocks and state-level inflation of the United States. *Economics Letters*, 238, 111714.
- Luo, J., Fu, S., Cepni, O., & Gupta, R. (2024). Climate risks and forecastability of US inflation: evidence from dynamic quantile model averaging. In *Climate risks and forecastability of US inflation: evidence from dynamic quantile model averaging*: Luo, Jiawen| uFu, Shengjie| uCepni, Oguzhan| uGupta, Rangan. Pretoria, South Africa: Department of Economics, University of Pretoria.
- Safonov, M. (2023). Quantitative Analysis of The Impacts Of Carbon Pricing On Energy Inflation in The European Union.
- Selmi, R. (2023). Do investors care about carbon risk? The impact of the Paris agreement on the inflation hedging performance of commodities. *Economics Bulletin*.
- Odongo, M. T., Misati, R. N., Kamau, A. W., & Kisingu, K. N. (2022). Climate change and inflation in Eastern and Southern Africa. *Sustainability*, 14(22), 14764.
- Lyu, Y. F. (2019). An empirical analysis on the relationship between green economy and inflation in China. *Applied Ecology & Environmental Research*, 17(4).
- Bistline, J. E., Mehrotra, N. R., & Wolfram, C. (2023). Economic implications of the climate provisions of the inflation reduction act. *Brookings Papers on Economic Activity*, 2023(1), 77-182.
- Ciccarelli, M., Kuik, F., & Hernández, C. M. (2024). The asymmetric effects of temperature shocks on inflation in the largest euro area countries. *European Economic Review*, 168, 104805.
- Zhang, Z. (2023). Are climate risks helpful for understanding inflation in BRICS countries?. *Finance Research Letters*, 58, 104441.
- Yusifzada, T. (2022). Response of Inflation to the climate stress: Evidence from Azerbaijan.
- Wahidah, N. L., & Antriandarti, E. (2021, April). Impact of climate change and Coronavirus Disease (COVID-19) on inflation in Indonesia. In *IOP Conference Series: Earth and Environmental Science* (Vol. 724, No. 1, p. 012105). IOP Publishing.
- Kunawotor, M. E., Bokpin, G. A., Asuming, P. O., & Amoateng, K. A. (2022). The impacts of extreme weather events on inflation and the implications for monetary policy in Africa. *Progress in Development Studies*, 22(2), 130-148.
- Aguila, N., & Wullweber, J. (2024). Greener and cheaper: green monetary policy in the era of inflation and high interest rates. *Eurasian Economic Review*, 14(1), 39-60.
- Konradt, M., & di Mauro, B. W. (2021). Carbon taxation and inflation: Evidence from Europe and Canada1. *Combatting Climate Change: a CEPR Collection*.
- Boneva, L., & Ferrucci, G. (2022). Inflation and climate change: the role of climate variables in inflation forecasting and macro modelling.

Moessner, R. (2022). Evidence on climate policy, carbon dioxide emissions and inflation. *International Journal of Global Warming*, 28(2), 136-151.

Beirne, J., Dafermos, Y., Kriwoluzky, A., Renzhi, N., Volz, U., & Wittich, J. (2024). Weather-related disasters and inflation in the euro area. *Journal of Banking & Finance*, 107298.

Economic Mobility in Domestic Violence Experienced Families

By Max Wu and Andrew Li

Abstract

Economic factors play a large role in determining the likelihood of experiencing intimate partner violence (IPV). However, less studied is the impact of IPV on a victim's economic mobility. This paper offers a method to identify some of the issues that victims of IPV will face, particularly by examining metrics related to socioeconomic status. Using data from the Laurel House as a case study, this study conducts a thorough analysis of the struggles these survivors and their children will continue encountering even after separating from an abuser.

The results of this paper can be used as a framework to guide city agencies around the world that are looking to provide long-term solutions to support emergency interventions. This study provides various recommendations to help mitigate the chance of IPV, including providing educational programs at schools. The key conclusions are: (i) those with lower socioeconomic status are most likely to enter emergency IPV shelters, (ii) almost all victims of IPV were denied housing at least once during their search due to economic constraints, and (iii) children of victims living in IPV shelters were often relocated and needed to change schools, which could disrupt the learning process.

Introduction

Domestic violence, also commonly referred to as intimate partner violence (IPV), is defined as physical violence, sexual violence, stalking, and psychological aggression by a current or former intimate partner. IPV varies in both frequency and intensity, manifesting anywhere from one isolated episode of violence to chronic and severe episodes over many years. This violence is not limited to specific groups of people, but rather can occur in any relationship regardless of race, age, or sex. Potential lifelong or long-term consequences of IPV include emotional trauma, chronic health issues, or physical impairment. Additionally, incidents of IPV remain largely unreported, and it is estimated that under 40% of victims end up reporting these crimes or seek out help. Although it may seem counterintuitive for victims to choose to stay in abusive relationships or keep issues unreported, most victims of IPV do not seek help due to the fear of more harmful effects down the line. Thus, this violence tends to become normalized within many families and has the potential to generate a "cycle of abuse" that carries intergenerational repercussions.

Many of the challenges mentioned above have been exacerbated by the COVID-19 pandemic, which has highlighted several critical issues in combating IPV. Reports of IPV across the US during the pandemic jumped from 21% to 35% due to stay-at-home orders. Some of the main problems showcased by the pandemic include: limited operation of support services due to health lockdowns; isolation between victims and community support networks; and, perhaps most devastatingly, an increase in financial stress with the widespread closure of businesses and industries.

Consistent with pre-pandemic data, two of the largest factors influencing the likelihood and risk of IPV include a household's socioeconomic status and income imbalance between partners. In many cases, victims will report economic concerns as a major barrier if they were to leave their abusive partners. To gain economic leverage in a relationship, abusive partners can employ a variety of strategies including preventing resource acquisition through interfering with their partners' employability, monetarily controlling resource use, and exploitation through means such as stealing, gambling, or damaging possessions. Thus, these compounding economic factors in many cases of IPV makes it a leading cause of homelessness nationally.

There are about 133–275 million children who are exposed to or witness IPV. While some are primary victims, children who live in a household with IPV but are not involved in the violence directly are known as secondary victims. Understandably, the main focus of research has been on primary victims, with only a recent increase in concern for the effects that IPV has on secondary victims. Secondary victims, particularly children, often develop a range of emotional, psychological, and behavioral problems after encounters with IPV. These victims are at a higher risk of developing anxiety, depression, and post-traumatic stress disorder (PTSD). These problems are particularly resistant to change when occurring in a child's developmental stage, and become entrenched easily. The household environment plays a critical role in a child's development, and growing up in a home with IPV disrupts the sense of safety and stability that is essential for healthy growth. This toxic stress can impede cognitive development and hinder the acquisition of essential life skills. Moreover, the lack of positive role models and supportive adults can exacerbate feelings of isolation and helplessness.

Beyond emotional effects, childhood IPV exposure also causes substantial negative economic effects. The culmination of healthcare costs, increased crime costs, and productivity losses push the total costs from childhood IPV exposure to over \$50,000 per victim. There was also a significant gender disparity in these costs, with the estimated lifetime cost being \$103,767 for females and \$23,414 for males. For each birth year, these costs amount to over \$55 billion across the United States.

This study was undertaken to fulfill the following objectives: (i) analyze the correlation between socioeconomic factors and victims of IPV; (ii) assess housing transition challenges faced by residents; and (iii) explore the impact of violence on children of victims, particularly focusing on their educational and transportation needs.

Background

Domestic violence does not exist in a vacuum; it is influenced by various socioeconomic factors. The American Psychological Association (APA) defines socioeconomic status (SES) as an individual or group's social standing or class, typically assessed through a combination of income, educational attainment, financial security, and the perception of an individual through a society's lens. This includes mistreatment, privilege, and opportunity; in this case, it affects the chance of experiencing IPV by influencing how an individual is viewed and victimized.

Rates of IPV are generally considered to be most influenced by educational attainment. High IPV victimization rates have always been reported among women with low educational

backgrounds in the United States. This could be because lower educational attainment is associated with reduced life skills and empowerment, giving a sense of control to the abuser. However, while secondary education is generally considered more protective, it is equally important to recognize that there are specific instances in which this is not the case. For example, certain communities with traditional gender roles tend to have norms that tolerate abuse more, and the rate of IPV is not affected by the supposed protective nature of secondary education. The same study states that in these communities, women with higher education are more likely to be affected by IPV because they challenge male authority more, and are retaliated against more often than women without second education.

The relation between women’s employment and IPV is similarly double-sided. Women’s income can serve as a protective factor against IPV since it decreases financial dependence of their partner. Conversely, contributing more financially than a partner can also increase the risk of IPV, with abusers more likely to see them as financially exploitable. This is most apparent when their partners hold traditional beliefs regarding gender roles.

Financial dynamics within relationships, such as asking for or receiving money, can heighten the risk of IPV. Male control over income, along with low SES and insufficient family income, is a strong predictor of abuse. The presence of stress stemming from economic hardship is only one of many factors that can exacerbate these economic conflicts, potentially leading to IPV.

One organization looking to mitigate the effects of IPV is the Laurel House, a comprehensive domestic violence agency serving individuals and families in Montgomery County, Pennsylvania established in 1980. The agency offers an emergency shelter for those who have reached out to flee an abusive situation. The Laurel House also offers other services, including but not limited to: a 24-hour hotline, individual and group counseling, medical advocacy, legal representation, and community education efforts.

The data presented in the tables below are obtained from the Laurel House and are based on either (a) the 10 current households living at the Laurel House as of September 23, 2024 or (b) the 196 individual households that have accessed the Laurel House’s emergency shelter from June 2022 to September 2024; this data is not publicly available.

Table 1: Length of Stay of 10 Current Households at the Laurel House

Statistical Measure	Length of Stay (Days)
Mean	54
Median	42
Low	11
High	116

Table 2: Length of Stay of Past Households at the Laurel House

Statistical Measure	Length of Stay (Days)
Mean	29.92
Low	1
High	144

Table 1 and Table 2 display the length of stay of households that are either currently at or have previously accessed the Laurel House’s emergency shelter. These tables show the wide range of durations sought out by households, which is consistent with other analyses of length of stay in IPV emergency shelters. Factors that influence the length of stay for a household include: age, number of children, education, level of needs, and proportion of needs met by the shelter. Factors that do not seem to influence length of stay include: employment, type of abuse, difficulty adapting to the shelter, and relationships within the shelter. Additionally, services at IPV shelters like counseling decrease the chance of re-abuse if the stay is longer.

Results

Socioeconomic data from the Laurel House (June 2022 - September 2024) reveals the most common types of families affected by domestic violence. The following metrics were collected: public benefits, employment, and household composition.

Table 3: Public Benefits Analysis (June 2022 to September 2024)

Public Benefits	Number of Households
Received Public Benefits	150
Did Not Receive Public Benefits	0
Did Not Disclose	46

Table 3 describes the amount of households that received public benefits, including but not limited to: SNAP, Medicaid, CHIP, Medicare, SSI, SSDI. With the large majority of families receiving public benefits, this underscores the connection between economic vulnerability and domestic violence. The prevalence of public benefits among these households reflects a critical need for sustained economic support and social safety nets within this demographic. Additionally, the large number of families who did not disclose whether they were receiving such benefits suggests potential underreporting or gaps in accessing available resources due to either stigma or bureaucratic challenges. When compared to the general population of Pennsylvania, 76.4% (150 out of 196 families) of Laurel House households received public benefits compared

to 24.7% of Pennsylvanias. This does not account for Laurel House households who did not disclose, meaning the gap between percentages is likely larger.

Table 4: Employment Analysis (June 2022 to September 2024)

Employment Situation	Number of Households
Employed	87
Unemployed	109
Did Not Disclose	0

Table 4 describes the amount of households that were employed in any capacity, including part-time or full-time employment. The employment rate of 44.4% (87 out of 196 families) amount shelter residents is considerably lower than the employment rate in the general population of Montgomery County, where Laurel House is located, which stands at approximately 66.5%. This has two possible suggestions. One, IPV can limit a survivor’s ability to engage fully in the workforce, whether due to economic abuse, the demands of caregiving, or trauma. Two, unemployment causes victims to appear more vulnerable, increasing the chance of IPV.

Table 5: Head of Household Analysis (June 2022 to September 2024)

Head of Household	Number of Households
Female Headed	188
Male Headed	2
Did Not Disclose	6

Table 6: Types of Households (June 2022 to September 2024)

Type of Household	Number of Households
Single Adults	91
Families	105
Did Not Disclose	0

Table 5 and Table 6 describe the amount of different types of households. The overwhelming number of female-headed households not only aligns with national trends in IPV but also points to the intersection of gender, economic vulnerability, and housing insecurity. At

the same time, it is important to acknowledge that male-headed households that have experienced IPV may be less likely to report due to societal expectations and stigmatization. The amount of victims that are in a family is 53.6% (105 out of 196 families), which is lower than the Pennsylvania state average of 67.5%, indicating that families with children are correlated with reduced IPV reporting and a lower chance of seeking out emergency shelters. This is likely attributed to potential disruption of schooling and emotional strain on children.

A lack of financial and housing stability can be caused by the abuser both directly and indirectly. Examples of direct actions include ruining the victim’s credit, stealing from the victim, or finding ways to harass them at their work until they are fired. Indirect links would include the result of depression or post-traumatic stress disorder from some IPV cases, which could make it more difficult to go to work or pay bills, thus making it more likely survivors have trouble finding homes. Additionally, the process of acquiring housing for IPV victims has many considerations, including but not limited to: finding an area safe from the abuser, making sure there is a strategy to address potential trauma, and building support systems in the community.

Information provided from the Laurel House shows that almost all families were denied housing at least once during their search. Factors included: low credit score, poor rental history, lack of positive rental references, insufficient income, landlords unwilling to accept rental assistance, and landlords unwilling to have county-based safety inspections.

Table 7: Housing Outcomes (June 2022 to September 2024)

Housing Outcome	Number of Households
Found Housing Away from Their Abuser	147
Returned to Abuser	22
Did Not Disclose	27

Table 7 describes the housing outcome for the 196 households that have previously lived in the Laurel House’s emergency shelter. Ways people found housing away from their abuser includes but is not limited to: transitional housing, rapid rehousing, housing vouchers, independent rentals by client, rooms for rent, shared living situations, moving in with friends or families, and relocating out of the country or state for safety. While most found housing away from their abuser, approximately 11% (22 out of 196 families) returned to their abuser, most likely due to economic ties and difficulty finding independent housing.

On top of helping victims that are directly involved in IPV, it is also important to see the effects of IPV on secondary victims, especially children. Data from the Laurel House shows the effect of IPV on schooling outcomes, shown below.

Table 8: Schooling Outcomes (June 2022 to September 2024)

Schooling Outcome	Number of Households
Households Moving out of the Area that Remained in Original School	11
Households Moving out of the Area that Changed Schools	22
Households Moving within the Same District	105
Households without School-age Children	58

Out of the 33 households with children affected by changes in living conditions, two-thirds went with the change and one-third decided to stay in their school of origin. This option is allowed by the McKinney-Vento Act, which allows families experiencing homelessness to choose between whether to switch schools or not. Students that chose to attend the local school district were provided transportation by the district, and students remaining in their school of origin will have the Laurel House’s children’s advocate help network with the school of origin to provide transportation.

Discussion/Recommendation

The overall data underscores the effectiveness of programs at IPV shelters that focus on financial independence, such as job training and credit repair services. Moreover, the findings suggest that IPV survivors do not become more exploitable when they are more financially or educationally independent, and will actually provide a pathway towards sustainable economic stability for IPV-affected families.

Despite these findings, several limitations of the study should be acknowledged. The geographic limitations of our data set, with a majority of data points coming from urban areas, may have skewed the results. Rural IPV survivors may face entirely different economic barriers, such as access to resources, job opportunities, and legal support. Second, the study relied on self-reported data from survivors who sought help from a formal organization, Laurel House. This study likely overlooks the experience of those who do not have access to these services, either due to stigma, lack of awareness, or fear. As a result, the findings may not be fully representative of the broader population of families or individuals affected by IPV.

This leads to several areas of further exploration. Most clearly, the experiences of rural based survivors needs further research. However, beyond that, investigating the role of community-based solutions, such as local cooperatives or faith-based organizations, could offer insights into alternative ways to provide safety for IPV-experienced families. Moreover, future studies should aim to collect long-term data to track the economic trajectories of survivors, which would allow researchers to assess current solutions more accurately, and identify the

specific key factors these solutions demonstrated. Finally, the intersection of race and socioeconomic status needs more in-depth investigation. Marginalized communities are disproportionately affected by IPV and experience more barriers to economic mobility. An intersectional approach to this could capture the complexity of these issues and would help to design targeted intervention strategies.

The process of recovering from IPV is complex and different for everyone. As noted in this study, there are a wide range of challenges when helping victims find independence away from an abuser. In particular, this study analyzes difficulty related to obtaining housing, caused either by mental health issues following abuse or a lack of economic resources. This highlights why it is imperative that agencies centered around helping victims have the resources they need to best assist them. For example, the City of Philadelphia has set aside 20% of their Emergency Housing Vouchers for survivors of domestic violence, sexual assault, and trafficking. This allows survivors to leave abusive situations and a chance to become economically independent without needing to return.

Since domestic violence's long-reaching consequences can affect children and adolescents, schools have the power to play an important role in shaping healthy relationships. Comprehensive programs that occur in schools are crucial in decreasing rates of adolescents as primary victims of IPV and preventing further negative consequences in secondary victims. Furthermore, 33% of adolescents in America are victims of sexual, physical, verbal, or emotional dating abuse. This comes with 81% of parents admitting that they do not believe teen dating violence is an issue, which makes it important that intervention starts with young people directly. Students that are secondary victims of IPV-related issues also become less likely to do well in school. Factors in this could include decreased economic outcomes after encountering IPV which makes it difficult to access educational resources, as well as an overall sense of displacement and possible emotional trauma.

Educational programs in schools that focus on IPV are most effective when they are ingrained in curriculums and become mandatory, coordinated efforts. Schools should promote a whole-school approach to IPV education, where all staff members are trained to recognize the signs of abuse and understand the protocols for responding to disclosures. This includes everyone including teachers, administrators, counselors, and support staff. When everyone in the school community is equipped with the knowledge and skills to address IPV, students are more likely to receive the support they need in a timely and compassionate manner. It's also important to recognize that IPV education must be inclusive and culturally sensitive. Programs should be designed to address the diverse experiences and backgrounds of students, taking into account factors such as gender, race, sexual orientation, and socioeconomic status. This ensures that all students feel represented and understood, making the education more effective and accessible.

Additionally, schools should advocate for policies at the district and state levels that mandate IPV education as a standard part of the curriculum. Legislative support can help ensure that school IPV programs are adequately funded and prioritized, leading to more widespread and consistent implementation across schools. By advocating for such policies, schools can extend

their impact beyond their immediate community, contributing to a larger movement to combat IPV and protect future generations from its harmful effects.

To guarantee that secondary victims are also accounted for within schools, schools should prioritize holistic support. This includes enhanced counseling services to address emotional trauma and provide a safe space for students to discuss their experiences. Academic support programs can help students catch up on missed coursework and improve their performance. Community partnerships can connect students with essential resources and support services. Comprehensive awareness and education can empower students with knowledge about healthy relationships and the signs of abuse, but also help teachers understand the deep cutting effects that witnessing IPV as a secondary victim can have on a student's life. Finally, crisis intervention plans can ensure students' safety and well-being in emergency situations. By implementing these strategies, schools can create a more supportive environment for students affected by IPV and help them overcome the academic challenges they may face.

Conclusion

The findings highlight the relationship between economic mobility, IPV, and its impact on all members of the family. While a majority of IPV research is understandably placed on the immediate physical and emotional consequences of IPV, we look to emphasize the long-term economic repercussions, particularly on housing, employment, and the transmission of vulnerability. This study further demonstrates that (i) victims of IPV that access emergency shelters are primarily those with lower socioeconomic status, (ii) almost all households were denied housing at least once during their search, with around 11% returning to their abuser, and (iii) children of households living in IPV emergency shelters are often relocated and change schools.

Overall, we conclude that temporary interventions are helpful, but they are not sufficient in meeting the needs of families and children who experience ongoing instability. Support systems should be more comprehensive, addressing both academic and social challenges. Economic and social empowerment of victims will allow for the breaking of the cycle of domestic violence. Educational systems and public officials should play a more active role in preventing IPV.

The challenges faced by underserved communities, particularly children affected by IPV, are deeply intertwined with broader systemic issues in education and social support. Disrupted education, caused by frequent relocation and/or unstable environments, is a barrier for academic success. Addressing these barriers requires a more cohesive and sustained effort from a variety of groups – including educators, policymakers, social service providers – and the increased support of shelters like the Laurel House to ensure continuity in services provided directly to IPV victims.

Works Cited

- Abramsky, Tanya, et al. "Women's income and risk of intimate partner violence: secondary findings from the MAISHA cluster randomised trial in North-Western Tanzania." *NCBI*, 14 August 2019, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6694529/>. Accessed 25 September 2024.
- Adams, Adrienne E. "Development of the Scale of Economic Abuse." *Violence Against Women*, vol. 14, no. 5, 2008, pp. 563-588. *Sage Pub*, <https://journals.sagepub.com/doi/10.1177/1077801208315529>.
- Ahmadabadi, Zohre. "Income, Gender, and Forms of Intimate Partner Violence." *Journal of Interpersonal Violence*, vol. 35, no. 23-24, 2020, pp. 5500-5525. *Sage Journals*, <https://journals.sagepub.com/doi/abs/10.1177/0886260517719541>.
- American Psychology Association. "APA Deep Poverty Initiative." *Socioeconomic Status Portfolio*, <https://www.apa.org/pi/ses>. Accessed 25 September 2024.
- Anrows. "New report finds education is an important step in reducing rates of Gender-Based Violence." *ANROWS*, 15 November 2023, <https://www.anrows.org.au/media-releases/new-report-finds-education-is-an-important-step-in-reducing-rates-of-gender-based-violence/>. Accessed 22 August 2024.
- Centers for Disease Control and Prevention. "About Intimate Partner Violence | Intimate Partner Violence Prevention." *CDC*, 16 May 2024, <https://www.cdc.gov/intimate-partner-violence/about/index.html>. Accessed 5 August 2024.
- City of Philadelphia. "How Housing Barriers Affect Survivors of Domestic Violence." <https://www.phila.gov/2022-09-29-how-housing-barriers-affect-survivors-of-domestic-violence/>.
- Committee on Health Care for Underserved Women. "Intimate Partner Violence." *ACOG*, 2022, <https://www.acog.org/clinical/clinical-guidance/committee-opinion/articles/2012/02/intimate-partner-violence>. Accessed 5 August 2024.
- Commonwealth of Pennsylvania. "People Enrolled in DHS Benefits | Department of Human Services." *Commonwealth of Pennsylvania*, May 2023, <https://www.pa.gov/en/agencies/dhs/resources/data-reports/people-enrolled-in-dhs-benefits.html>. Accessed 29 September 2024.
- Costa, D., et al. "Male and female physical intimate partner violence and socio-economic position: a cross-sectional international multicentre study in Europe." *National Library of Medicine*, June 2016, <https://pubmed.ncbi.nlm.nih.gov/27262180/>. Accessed 25 September 2024.
- Cunradi, Carol B., et al. "Socioeconomic Predictors of Intimate Partner Violence Among White, Black, and Hispanic Couples in the United States." *Springer Link*, December 2002, <https://link.springer.com/article/10.1023/A:1020374617328>. Accessed 25 September 2024.

- Digby, Ali, and Emily Fu. "Impacts of homelessness on children -- research with teachers." *Kantar Public*, 2017. *ctfassets*, https://assets.ctfassets.net/6sxvmndn0s/1ExhoeDJaFxF89Sdcms4i0/43938d3bda5c7ff4665b136cb6aa3ead/2017_12_20_Homelessness_and_School_Children.pdf.
- DomesticShelters. "Are Domestic Violence Rates Dropping?" *DomesticShelters.org*, 20 April 2015, <https://www.domesticshelters.org/articles/statistics/are-domestic-violence-rates-dropping>. Accessed 5 August 2024.
- Gibson, Bobby. "Raising Awareness of Teen Dating Violence." *House Dems CT*, 2022, <http://www.housedems.ct.gov/node/20690#:~:text=One%20in%20three%20adolescents%20in,dating%20violence%20is%20an%20issue>. Accessed 22 August 2024.
- Gupta, Naman. "'I really just leaned on my community for support': Barriers, Challenges, and Coping Mechanisms Used by Survivors of Technology-Facilitated Abuse to Seek Social Support." *USENIX Security 2024*, 2024. *ResearchGate*, https://www.researchgate.net/publication/381918000_I_really_just_leaned_on_my_community_for_support_Barriers_Challenges_and_Coping_Mechanisms_Used_by_Survivors_of_Technology-Facilitated_Abuse_to_Seek_Social_Support.
- Harne, Lynne. *Tackling Domestic Violence*. McGraw Hill, 2008. *Google Books*.
- Holmes, Megan, et al. "Economic Burden of Child Exposure to Intimate Partner Violence in the United States." *ResearchGate*, May 2018, https://www.researchgate.net/publication/323259558_Economic_Burden_of_Child_Exposure_to_Intimate_Partner_Violence_in_the_United_States. Accessed 2 October 2024.
- Huecker, Martin. *Domestic Violence*. Treasure Island, 2023. *NCBI*, <https://www.ncbi.nlm.nih.gov/books/NBK499891/>.
- IPVHealth. "Frequently Asked Questions." *IPV Health*, <https://ipvhealth.org/faqs/>.
- Jewkes, Rachel. "Intimate partner violence: the end of routine screening." *National Library of Medicine*, 16 April 2013, <https://pubmed.ncbi.nlm.nih.gov/23598182/>. Accessed 25 September 2024.
- Jouriles, Ernest N., et al. "Children Exposed to Intimate Partner Violence: Conduct Problems, Interventions, and Partner Contact With the Child." *National Library of Medicine*, 2016, <https://pubmed.ncbi.nlm.nih.gov/27359091/>. Accessed 2 October 2024.
- Kourti, Anastasia. "Domestic Violence During the COVID-19 Pandemic: A Systematic Review." *Sage Journals*, vol. 24, no. 2, 2023, pp. 719-745. *SagePub*, <https://journals.sagepub.com/doi/full/10.1177/15248380211038690>.
- Laurel House. "About Us – laurel-house.org." *Laurel House*, <https://laurel-house.org/about-us/>. Accessed 25 September 2024.
- Lloyd, Michele. "Domestic Violence and Education: Examining the Impact of Domestic Violence on Young Children, Children, and Young People and the Potential Role of Schools." *Frontiers*, vol. 9, 2018. *Frontiersin*, <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2018.02094/full>.

- Mojahed, Amera, et al. "Intimate Partner Violence Against Women in the Arab Countries: A Systematic Review of Risk Factors." *PubMed*, 3 September 2020, <https://pubmed.ncbi.nlm.nih.gov/32878586/>. Accessed 25 September 2024.
- Niolon, Phyllis Holditch, and Shalom M. Irving. "Intimate Partner Violence Prevention Resource for Action." *CDC*, 2017, https://www.cdc.gov/violenceprevention/pdf/ipv-prevention-resource_508.pdf. Accessed 29 July 2024.
- NNEDV. *Domestic Violence, Housing, and Homelessness*. 2019. *NNEDV*, https://nnedv.org/wp-content/uploads/2019/07/Library_TH_2018_DV_Housing_Homelessness.pdf.
- Perez, S., et al. "The Role of PTSD and Length of Shelter Stay in Battered Women's Severity of Reabuse after Leaving Shelter." *Journal of Aggression, Maltreatment & Trauma*, vol. 21, no. 7, 2012, pp. 776-791. <https://doi.org/10.1080/10926771.2012.702712>.
- Peterson, Cora, et al. "Lifetime Economic Burden of Intimate Partner Violence Among U.S. Adults." *PubMed*, 2018, <https://pubmed.ncbi.nlm.nih.gov/30166082/>. Accessed 2 October 2024.
- Point2. "Montgomery County, PA Household Income, Population & Demographics | Point2." *Point2 Homes*, 2024, <https://www.point2homes.com/US/Neighborhood/PA/Montgomery-County-Demographics.html>. Accessed 29 September 2024.
- Reid, M. Jamila, et al. "Follow-up of children who received the incredible years intervention for oppositional-defiant disorder: Maintenance and prediction of 2-year outcome." *Science Direct*, 2003, <https://www.sciencedirect.com/science/article/abs/pii/S000578940380031X>. Accessed 2 October 2024.
- Sanders, Cynthia K. "Organizing for Economic Empowerment of Battered Women: Women's Savings Accounts." *ResearchGate*, vol. 14, no. 3, 2006, pp. 47-68. *ResearchGate*, https://www.researchgate.net/publication/233377663_Organizing_for_Economic_Empowerment_of_Battered_Women_Women's_Savings_Accounts.
- Schneider, Avie. "Nearly 10 Million Filed For Unemployment In Last 2 Weeks, A Dismal Record : Coronavirus Updates." *NPR*, 2 April 2020, <https://www.npr.org/sections/coronavirus-live-updates/2020/04/02/825383525/6-6-million-file-for-unemployment-another-dismal-record>. Accessed 5 August 2024.
- Smith-Clapham, Amber M., and Julia E. Childs. "Implications of the COVID-19 Pandemic on Interpersonal Violence Within Marginalized Communities: Toward a New Prevention Paradigm." *Am J Public Health*, vol. 113, no. S2, 2023, pp. 149-156. *National Library of Medicine*, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10282862/>.
- Sterne, Abigail. *Domestic violence and children: A handbook for schools and early years settings*. Routledge, 2009. *Google Scholar*.

- Sullivan, Cris M., et al. “‘There’s Just All These Moving Parts:’ Helping Domestic Violence Survivors Obtain Housing.” *Clinical Social Work Journal*, vol. 47, 2019, pp. 198-206. <https://www.proquest.com/docview/2016874022?fromopenview=true&pq-origsite=gscholar&sourcetype=Scholarly%20Journals>.
- Sullivan, Cris M., and Gabriela Lopez-Zeron. “Impact of the Domestic Violence Housing First Model on Survivors’ Safety and Housing Stability: Six Month Findings.” *Journal of Family Violence*, vol. 38, 2023, pp. 395-406.
- Swadley, Rachel L. “Returning to Abusive Relationships: Related and Predictive Factors.” *MSU Graduate Theses*, 2017. *Missouri State University*, <https://bearworks.missouristate.edu/cgi/viewcontent.cgi?article=4183&context=theses>.
- Taylor, Julie C., et al. “Barriers to Men's Help Seeking for Intimate Partner Violence.” *NCBI*, 25 August 2021, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9554285/>. Accessed 30 September 2024.
- Trujillo, Monica P., and Eric Quintane. “Why do they Stay? Examining Contributing Factors to Women’s Length of Stay in a Domestic Violence Shelter.” *Journal of Family Violence*, vol. 32, no. 1, 2017, pp. 89-100, <https://doi.org/10.1007/s10896-016-9904-1>.
- UNICEF. “Behind Closed Doors The Impact of Domestic Violence on Children.” *Belfast Women's Aid*, 2006, https://belfastwomensaid.org.uk/themainevent/wp-content/uploads/2021/09/UNICEF_BehindClosedDoors.pdf. Accessed 2 October 2024.
- United States Census Bureau. “Montgomery County, Pennsylvania - Census Bureau Profile.” *Census Data*, 2024, https://data.census.gov/profile/Montgomery_County,_Pennsylvania?g=050XX00US42091#employment. Accessed 29 September 2024.
- UN Women. “COVID-19 and Ending Violence Against Women and Girls.” *WEPs*, 2 April 2020, <https://www.unwomen.org/sites/default/files/Headquarters/Attachments/Sections/Library/Publications/2020/Issue-brief-COVID-19-and-ending-violence-against-women-and-girls-en.pdf>. Accessed 5 August 2024.
- Walters, Andrienne C. “The forgotten children: victims of domestic violence, victims of the system.” *Government Law Review*, 2018, <https://www.albanygovernmentlawreview.org/article/24015-the-forgotten-children-victims-of-domestic-violence-victims-of-the-system>. Accessed 2 October 2024.
- Watts, Charlotte. “How does economic empowerment affect women's risk of intimate partner violence in low and middle income countries? A systematic review of published evidence.” *Wiley Online Library*, 06 October 2008, <https://onlinelibrary.wiley.com/doi/10.1002/jid.1500>. Accessed 25 September 2024.
- A Woman's Place. “Teen Dating Violence - A WOMAN'S PLACE.” *A Woman's Place*, 2020, <https://www.awpdv.org/teen-dating-violence.html>. Accessed 22 August 2024.

Marketing for a Cause: An Exploration of How For-Profit and Nonprofit Brands Leverage Social Media to Engage Teenagers By Victoria Oh

Abstract

In today's digital age, social media platforms have become pivotal in shaping how organizations connect with their audiences, especially teenagers. This paper explores the most effective social media marketing tactics targeting teenagers and evaluated non-profit and for profit campaigns. A review of existing literature found that nonprofits focused on engaging a larger audience for their cause and for profits focused on brand loyalty and increasing sales. Despite having different goals, both non-profits and for-profits use similar tactics to leverage their brand, including authenticity, transparency, interactive content, and influencer collaboration. This paper also includes four case studies, evaluating the differences in marketing techniques.

Introduction

In today's hyper-connected world, effectively engaging teenagers has become both a necessity and a challenge for nonprofit and for-profit organizations alike. As digital natives, teens are one of the most active groups online, with 95% of teens reporting that they have access to a smartphone, and 45% stating that they are online "almost constantly" (Pew Research Center, 2018). This unprecedented level of connectivity makes them a critical demographic for organizations seeking to build long-term engagement and loyalty. However, despite the clear importance of this age group, there remains a surprising lack of research on how to tailor marketing strategies specifically to appeal to teens, particularly within the nonprofit sector. Research has shown that Gen Z is driven by authenticity and social responsibility. According to a report by the National Retail Federation, 67% of teens prefer brands that contribute to social causes, and they value companies that align with their personal beliefs (National Retail Federation, 2022). While these insights are vital, the gap in academic research leaves organizations without a robust framework to fully understand the most effective ways to harness these preferences through digital marketing strategies. In particular, nonprofits, which often rely on the support of socially conscious individuals, could benefit significantly from better understanding how to appeal to this key demographic.

The purpose of this study is to explore the marketing strategies used by both nonprofit and for-profit organizations to engage teens, highlighting the current state of marketing efforts and the apparent gaps in existing literature. By analyzing real-world case studies, this research aims to provide insights into successful campaigns and suggest areas where further research is critically needed. As digital platforms evolve and youth engagement becomes more complex, addressing this gap in knowledge will help organizations refine their approach to reaching younger audiences, ensuring more effective and lasting connections.

Methods

To explore effective marketing tactics for nonprofits and for-profits aimed at engaging teenagers, I conducted a multi-step research process that included initial exploration using online scholarly databases, followed by an independent analysis of case studies. My research began with Google Scholar, where I searched for studies on general marketing strategies, particularly focusing on nonprofit and for-profit organizations targeting teenagers. However, I found limited existing literature specific to youth engagement. As a result, I broadened my search to general marketing tactics across various demographics.

I supplemented this initial academic research by using Google to locate case studies on real-world campaigns. This helped me identify practical applications of marketing strategies, which informed the bulk of my analysis. I focused on case studies that highlighted successful social media campaigns, evaluating their methods, audience engagement techniques, and outcomes. Though I found relevant case studies online, much of the research was independently analyzed by synthesizing common patterns and strategies observed in both nonprofit and for-profit sectors.

To ensure a comprehensive approach, I categorized the campaigns by their conversion strategies, such as calls to action, engagement metrics, and user-generated content. This process allowed me to assess which tactics were most effective at appealing to younger audiences. Throughout this research, I aimed to answer how nonprofits and for-profits can optimize their marketing tactics to better engage teenagers, particularly through social media channels.

Results

Social media marketing refers to the use of social media platforms to connect with a target audience, build a brand, and ultimately drive sales or achieve other business goals. Through platforms like Instagram, TikTok, and Facebook, businesses and organizations create and share content designed to attract attention, engage users, and generate responses. The content can vary in form, including images, videos, written posts, and interactive elements such as polls or lives.

SMM is distinct from traditional marketing methods in its ability to reach broader audiences and allow for immediate interaction with followers. The two-way nature of social media enables businesses to build relationships with their audiences, receive real-time feedback, and adjust their strategies as needed. The primary goals of social media marketing often include increasing brand awareness, driving traffic to a website, boosting sales, and cultivating customer loyalty.

Teenagers represent a valuable demographic in the world of social media marketing for several reasons. Firstly, they are early adopters of new technologies and trends. Their ability to quickly engage with and share content makes them powerful influencers on platforms where virality can determine the success of a product or campaign. According to Forbes, parents spend approximately \$17,000 on a single teenager annually, meaning that teens also exert substantial influence on household spending (Name Forbes ____). Their active presence on social

media—85% of Gen Z uses these platforms to discover new products—makes them an ideal audience for campaigns aimed at driving consumer behavior.

For businesses and nonprofits alike, targeting teens allows for the potential to reach not only the teens themselves but also a broader audience through their social circles and families. Teenagers often act as trendsetters, capable of propelling products, causes, or ideas to viral status, making them a critical focus of social media marketing strategies.

Social Media Marketing Tactics

Authenticity refers to the quality of being genuine, real, and true to oneself or one's brand. In social media marketing, authenticity involves presenting content and messages that reflect the true values, mission, and identity of the organization or brand. For example, a nonprofit might share real stories from individuals who have benefited from its services, while a for-profit brand might highlight behind-the-scenes aspects of its production process.

Transparency means being open and honest about the practices, processes, and values of the organization. This includes clear communication about how funds are used (for nonprofits), or how products are sourced and made (for for-profits). Transparency helps build trust with the audience by demonstrating that there is nothing to hide and that the organization or brand operates with integrity.

- **Builds Trust:** Authenticity and transparency help in establishing credibility and trust with the audience, especially important in social media where users are quick to spot inconsistencies.
- **Fosters Loyalty:** When organizations are genuine and open, they create stronger emotional connections with their audience, leading to greater loyalty and support.
- **Enhances Engagement:** Audiences are more likely to engage with content that feels real and aligned with their own values and beliefs.

Interactive Content involves elements that require active participation from the audience rather than passive consumption. This can include a range of formats, such as:

- **Polls and Surveys:** Allowing users to express their opinions or preferences.
- **Quizzes:** Engaging users with fun or informative questions that lead to personalized results.
- **Challenges:** Encouraging users to complete specific tasks or share their own content based on a theme or prompt.
- **Live Streams:** Offering real-time interaction between the audience and the brand or organization.
- **Increases Engagement:** Interactive content encourages users to actively participate, leading to higher levels of engagement compared to static content.

Visual Storytelling involves using visual elements—such as images, videos, infographics, and graphics—to tell a story and convey messages in a compelling and engaging way. It leverages the power of visuals to capture attention, communicate information quickly, and evoke emotional responses.

- Captures Attention: Visual content is more likely to stand out in crowded social media feeds compared to text-based content.
- Enhances Understanding: Visuals can simplify complex information and make it easier for the audience to grasp key messages.
- Evokes Emotion: Well-crafted visuals can create emotional connections and drive stronger responses from the audience.

Influencer Collaboration involves partnering with individuals who have significant followings on social media to promote a brand or cause. Influencers can range from celebrities to micro-influencers with smaller, niche audiences. Collaborations can include sponsored posts, product reviews, shoutouts, or co-created content.

- Expands Reach: Influencers can introduce a brand or cause to their established audience, extending the reach of the campaign.
- Builds Credibility: Influencers' endorsements can lend credibility to the brand or cause, as their followers often trust their opinions and recommendations.
- Increases Engagement: Influencer collaborations can drive higher engagement rates by leveraging the influencer's connection with their audience.

Case studies

Nonprofit Case Studies

The Dresember campaign serves as a prominent example of how a nonprofit can successfully engage teenagers through a participatory challenge. The campaign encourages individuals, particularly teens, to wear a dress or tie every day throughout December to raise awareness and funds for combating human trafficking. This study examines how Dresember uses social media platforms like Instagram and TikTok to mobilize teens through user-generated content, personal storytelling, and peer-to-peer fundraising. The methods of analysis focus on how Dresember utilizes emotional storytelling, encourages long-term commitment, and builds a sense of community among participants to achieve its conversion goals of increased awareness, donations, and sustained engagement.

The Red Cross disaster relief efforts provide an insightful look into how nonprofits use real-time social media updates during crises to engage and inform teenage audiences. This case study analyzes the Red Cross's use of platforms like Twitter and Facebook to disseminate crucial information, build community support, and drive immediate action, such as donations or volunteer sign-ups. The methods also explore the organization's use of influencer partnerships during natural disasters to amplify their message and how this collaboration with prominent

social media figures enhances engagement among teens. The analysis focuses on how real-time updates and crisis communication help convert engagement into tangible outcomes such as volunteer participation and disaster relief donations.

The Make-A-Wish Foundation's #ShareYourEars campaign illustrates the power of brand partnerships and visual storytelling in nonprofit social media efforts. This case study examines how Make-A-Wish partnered with Disney to encourage teens and families to post pictures wearing Mickey Mouse ears on social media, with each post resulting in a donation from Disney to the foundation. The analysis investigates how this campaign successfully utilized a clear visual call to action, user-generated content, and a brand partnership to engage teens and create widespread participation. The methods include an evaluation of the emotional appeal of the campaign and how Make-A-Wish harnessed the collective social media power of teens to drive donations and awareness for their cause.

For-Profit Case Studies

Apple's ShotOniPhone campaign highlights how for-profit organizations use social media to promote their products and engage teenagers through user-generated content. This case study focuses on how Apple encourages teens to showcase photos taken with their iPhones by sharing them on platforms like Instagram and Twitter. The methods of analysis examine how Apple combines product demonstration with user-generated content to foster brand loyalty and demonstrate the quality of the iPhone camera. Special attention is given to how Apple strategically uses social media influencers to amplify the campaign's reach and how this peer-to-peer engagement strengthens the connection between teens and the Apple brand, ultimately driving sales.

The Dove Real Beauty campaign serves as a for-profit case study in leveraging emotional appeal and social commentary to engage teenage audiences. This analysis explores how Dove used a combination of social media platforms and traditional advertising to challenge beauty standards by featuring real women of diverse shapes, sizes, and ethnicities. The methods focus on Dove's strategy of building emotional connections with teens by promoting body positivity and challenging societal norms. The study also examines how Dove's integration of influencer collaborations and user-generated content helps foster long-term brand loyalty and how the campaign's resonance with teens translates into increased sales and customer retention. Always' #LikeAGirl campaign offers an example of how a for-profit company can use social media to redefine stereotypes and empower teenage audiences. This case study examines how the campaign used a powerful video and social media channels to transform the phrase "like a girl" from an insult into a message of empowerment. The methods evaluate how to effectively engaged teens by addressing social issues, building a strong emotional connection, and encouraging teens to share their own stories of empowerment on social media. The analysis also considers how the campaign's focus on social change contributed to increased brand loyalty and product sales among teenage consumers.

Data for these case studies were collected through various channels, including social media metrics (e.g., likes, shares, and comments), campaign reports, and online articles detailing the success of the respective campaigns. Each campaign was analyzed based on its engagement rates, conversion outcomes, and the overall impact on the target demographic. Metrics such as the number of posts, reach, user interactions, and conversion rates (e.g., donations, sales, participation) were used to assess the effectiveness of each campaign in achieving its goals. The analysis also included the role of influencer partnerships and peer-to-peer engagement as drivers of campaign success. Ultimately, these methods allowed for a comprehensive understanding of the tactics nonprofits and for-profits use to leverage social media to engage teenagers effectively.

Discussion

Nonprofits tend to focus on long-term relationship building, aiming to convert engagement into sustained support through donations, participation in events, and volunteer sign-ups. Campaigns like Dressemer and Make-A-Wish's #ShareYourEars relied heavily on emotional storytelling and community building to foster ongoing commitment from teens. These campaigns demonstrate that nonprofits often prioritize emotional connections and shared values to drive conversions, with success measured by long-term impact rather than immediate financial gain.

For-profits, by contrast, emphasize more immediate conversions, such as sales and customer retention. Campaigns like Apple's ShotOniPhone and Dove's Real Beauty highlight how for-profit companies leverage social media to showcase product quality, build emotional connections, and foster brand loyalty, ultimately driving purchases. For-profits often use multi-channel marketing strategies and influencer partnerships to maximize their reach and prompt immediate consumer action. The emphasis is on aligning the brand with the values of teenage consumers while ensuring that the engagement leads to direct financial returns. Emotional Appeal plays a crucial role in both sectors, but nonprofits typically use it to drive long-term engagement and advocacy, while for-profits use it to foster brand loyalty and prompt immediate purchases. Both sectors rely on visual storytelling and influencer partnerships to convert engagement into tangible outcomes, but the types of conversions differ significantly—nonprofits focus on participation and awareness, while for-profits prioritize sales and customer retention.

Conclusion

This research brought up several interesting points regarding marketing strategies for nonprofits and for-profits, particularly when engaging teenagers. While my analysis highlighted effective tactics used in case studies, I did not delve deeply into some of the broader marketing techniques that surfaced during the process. As we can see, there remains a significant lack of academic research in this area, particularly regarding targeted strategies for engaging teens in nonprofit and for-profit marketing. This gap underscores the need for further studies that specifically address the unique challenges and opportunities in reaching younger audiences through social media and other digital platforms.

Works Cited

- Forbes. (2022). "Teen Spending Power: How Much Parents Spend on Their Teens Each Year." [Forbes Website].
- Dressemer Campaign. (2023). "Raising Awareness for Human Trafficking Through Fashion Activism." [Dressemer Website].
- Red Cross. (2023). "Social Media Response During Natural Disasters." [Red Cross Website].
- Apple. (2022). "Shot on iPhone Campaign: The Power of User-Generated Content." [Apple Website].
- Dove. (2022). "Real Beauty Campaign: Challenging Beauty Standards." [Dove Website].
- Make-A-Wish America. (2023). "#ShareYourEars Campaign: Disney and Make-A-Wish Partnership." [Make-A-Wish Website].

The Legacy of Chinese Theater: Beijing Opera's Trail Through Macbeth By Mei Teng Shao

Abstract

This paper examines Beijing Opera (*jingju*) adaptations of Shakespeare's *Macbeth* through the years 1986-2022. It first discusses Wu's *Kingdom of Desire*, a late 20th-century adaptation of the play into traditional Chinese operatic form, exploring the power of the Chinese people against their rulers. This paper argues that *Kingdom of Desire* reminds the PRC government that they can be overthrown. It then discusses *Who Is Macbeth*, a more recent adaptation in mixed theatrical forms that draws parallels between Macbeth and figures from Chinese history and myth. The progression between the two adaptations reveals a resurgence of pride in Chinese history. At the same time, it reveals an abandonment of traditional Chinese theater elements. Together, these progressions are emblematic of the current Chinese standpoint on its position in the world relative to the West: despite regaining pride in its culture, China is increasingly turning to Western forms in order to represent that culture, a trend that is likely to continue deeper into the 21st century.

Paper

According to Ruru Li, Shakespeare "was neither forced onto China by colonial education as in India, nor was he introduced by English literary scholars as in Japan. Instead, as they always had with foreign cultures since ancient times, Chinese people, having found Shakespeare useful, either appropriated and assimilated him into Chinese or used him to attack Chinese culture" (*Shashibiya*, 15). Discussion of Shakespeare in China, therefore, yields a nuanced and complex understanding of China's relationship with western ideas and art parallel to China's relationship with its own ideas and art. This essay will examine this intersection through the evolution of Chinese efforts to adapt *Macbeth*, from the late 1980s to today, whether it be into traditional forms of Chinese theater such as Beijing Opera (*jingju*), more westernized styles of spoken theater (*huaju*), or a combination of both. I argue that these efforts over the past thirty years reflect a shift away from idolizing Western culture and towards the advancement of Chinese culture on a world stage.

After the end of the Cultural Revolution in 1976, the Chinese appetite for Shakespeare increased dramatically. The Cultural Revolution set out to lift China up from the humiliation it has been suffering since the Opium Wars, but from a cultural and artistic perspective, the Cultural Revolution was a step backwards: only a few "model plays" (which adhered tightly to Communist Party ideology) were allowed to even be performed, and countless intellectuals and artists were persecuted for expressing nuanced and critical opinions (Durdin). Furthermore, countless works of existing art were targeted for being emblematic of bourgeois ideas. The dusk of the Cultural Revolution heralded a widespread understanding amongst citizens that Mao's China was largely detrimental. This, combined with a lingering fear of expressing original

critical ideas, as well as an admiration for the better-off countries west of China, created an extremely high demand for adaptations of Shakespeare (*Shashibiya*, 57).

So, why must we examine *Macbeth* specifically? Not only is *Macbeth* one of Shakespeare's more mainstream works, *Macbeth*, as a piece with prominent themes of power, tyranny, and ambition, echoed many of the struggles the Chinese people had to contend with. Additionally, the original *Macbeth* is theorized to have been written as a response to a plot to kill England's then King James I. The play contains jokes made at the expense of the conspirators as well as legitimations of King James I's paranoia and right to rule (Neves). Thus, *Macbeth* is a work which was conceived to both appease the ruling power but also communicate themes against tyranny. As such, *Macbeth* perfectly abetted the cultural and artistic need to speak to the power of governing bodies without fear.

These were the circumstances under which Wu Hsing-kuo created his 1986 *jingju* adaptation of *Macbeth* in Taipei, titled *Kingdom of Desire*. Though there are non-negligible differences in the opinions of Taiwanese and Chinese citizens, *Kingdom of Desire* quickly became incredibly well known, accepted, and discussed in both the mainland as well as Taiwan. In fact, it resonated in the mainland so strongly, a joint-production of the play was carried out in Beijing in 2001, and a Shanghai production of the play was carried out in 2010 (*The Soul*, 247). Thus, despite being originally a Taiwanese production, *Kingdom of Desire* can be used to parse opinions from the mainland as well.

There are many substantial differences between Shakespeare's original *Macbeth* and Wu's *Kingdom of Desire*, as will be explored in the next close readings. Some of these differences, such as the setting of the piece being changed to China's Warring States period (circa 403 BCE - 221 BCE), and the transformation of the three witches into a more Chinese concept of a mountain ghoulish, were necessitated by the switch to the *jingju* form. Other narrative changes, however, compel further examination. Of most interest, Macbeth and Lady Macbeth (Aoshu and Lady Aoshu) are no longer driven primarily by a hunger for power. Instead, while trying to convince Aoshu to kill Duncan (Jihou), Lady Aoshu brings up the concern that if Banquo (Mengting) tells Jihou of the prophecies, Jihou would immediately order Aoshu's death. This is exacerbated by Jihou's request for Aoshu to ride with him and Mengting into battle the next day, which Lady Aoshu points out, will "certainly make [Aoshu] victim to a poisonous plot, in which thousands of arrows shall pierce [his] heart and mutilate [his] corpse" (*Cast*, 52). It is only then that Aoshu relents and begins to plot Jihou's assassination with his wife. This actually makes the couple more sympathetic; they believe themselves to be in a cutthroat world, in which any inaction might as well be self-annihilation. This, perhaps, humanizes the couple, replacing what was only raw greed and ambition in Shakespeare's original with something more vulnerable, something more nakedly human.

Running contrary to their newly bequeathed humanity, Lady Aoshu and Aoshu also display their cruelty in new ways. After the assassin sent by the Aoshus to kill Mengting and his son returns with word that he failed, Lady Aoshu flies into such a rage that she kills the assassin with Aoshu's sword. The assassin, having been stabbed out of his shaking and kowtowing, can

only reach desperately for Lady Aoshu. Lady Aoshu herself, dressed in red robes as opposed to her pre-Empress purple, recoils at the sight of real blood on her hands and leaps to cower behind her husband. Red is the color of happiness and fortune in China. It is splashed over the lanterns of the Chinese New Year, the gifts of *hongbao*, and it shares its place on the modern Chinese flag with only gold, meaning royalty and wealth. Red is the color of the people, and wearing that same red, Lady Aoshu murders a servant who was only trying to fulfill his missive. With the addition of chaotic and climatic music as she recoils, it is clear that Lady Aoshu, in this moment, fundamentally betrayed the people of China with her actions, despite the Aoshus' earlier humanity. The blood that the original Lady Macbeth sees on her hands is not real, only a psychological figment of her guilt. However, the blood Lady Aoshu sees on her hands is real, spilt fresh from a servant she has really murdered. Though both Lady Macbeth and Lady Aoshu collapse into self loathing and guilt, it is only when Lady Aoshu is met with the sight of actual blood on her hands that she even begins to regret her actions. *Kingdom of Desire* tells us here that despite the couple's sympathetic origins, they're still capable of being just as, if not more cruel, than any other ruler of the people once their fear turns into anger.

Aoshu does not behave remarkably differently than in the original play, yet his death, influenced by Akira Kurosawa's 1957 film adaptation of *Macbeth, Throne of Blood*, is changed significantly. After announcing to his army from atop a ledge that they will not lose unless the very forest moves towards them, the army works itself into a celebratory frenzy that seems almost forced in its abandon and chaos until a messenger arrives and fearfully announces that the forest *has* started to move towards them. With rage, Aoshu physically comes apart as he tears off pieces of his costume, symbolically losing the evidence of his accomplishments and his qualification to be emperor. There is stillness as the army considers this, before they all silently turn on Aoshu and start to shoot at him. Instead of death at the hands of Macduff, Aoshu dies at the hands of his own army, thousands of arrows piercing his heart and mutilating him. This is extra ironic, as when he and his wife were debating whether they should kill the Emperor, this was his wife's warning to him, that "thousands of arrows shall pierce [his] heart and mutilate [his] corpse." Aoshu's ultimate fate, as we are warned in this mirroring, is the result of his fear.

As a culmination of the above close readings, we will examine metaphorical meanings of the Aoshus' deaths taken together. Again, the original *Macbeth* was not only written to appease power, but also to speak to that power. Although Shakespeare flattered King James I with legitimations of his lineage and claim to the English throne, *Macbeth*, at its core, is still a play about a king who went mad with power and was thusly killed. Through this lens, we may be able to examine the Aoshus as an allegory for the ruling power, the People's Republic of China (PRC). In this case, the fact that the Aoshus were humanized and allowed character motivations, such as fear of betrayal, reveal a respect for the origins of the Party, whether that be genuine or a form of appeasement. The PRC, as opposed to the Nationalist Party, was indeed the faction organizing an armed resistance to the Japanese in World War II, so a genuine respect makes sense. With time, however, Mao's various programs and initiatives, as well as their effect on the people of China, revealed themselves to have horribly wide-reaching consequences. The

deadliest famine in the history of China happened during Mao's Great Leap Forward. Given these just beginnings, combined with a horrible and deadly end, the Aoshu mirror the PRC, and that is why Aoshu's death is so important. China's 5000-year long history is rife with civil war and the transmutation of old empires into new ones under different leadership. Contrary to the divine monarchies of Europe, the ideas behind China's divine monarchy evolved to accommodate rebellion. The Mandate of Heaven is an ancient Chinese political doctrine which states that any time an empire is successfully overthrown, that means that the overthrown empire was judged unworthy by heaven and no longer holds divinity. This doctrine, and the incredible amount of power it places in the hands of the people, have driven Chinese politics for millenia. This is why Aoshu's death by his own soldiers' arrows matters so much; the arrows speak to the ruling party, warning that ultimately, if the soldiers, the metaphorical regular citizens of China, are upset, then there will be retribution.

In more recent times, China has continued to adapt Shakespeare. One such adaptation, *Who Is Macbeth*, produced by Le Bojuan, was first unveiled in Beijing in 2013, and has continued to evolve since then, most recently being performed in 2022. *Who Is Macbeth* tells the story of Macbeth alongside two famous persons from the Chinese historical canon, Zhao Guangyi and Gongsun Zidu. Macbeth's story is told unchanged up until his decision to kill Duncan. However, his story is then continued through the similar circumstances of the 10th-century CE scion Zhao Guangyi, who plotted with the then-emperor's concubine to kill the emperor. The narrator describes Macbeth as an "incarnation" of Zhao Guangyi (Song). After the emperor's death, the metaphorical torch is finally passed to Gongsun Zidu, a circa 8th- to 5th-century BCE general who betrayed another general on his own side and then went mad with guilt. The particular arrangement of these segments as parallel to Macbeth's story is almost haughty; it implies that these ancient Chinese historical figures are equivalent to, and in their precedence and quantity, perhaps even surpassing the importance of the original 11th-century CE Macbeth. Le's play takes pride in the history of China over that of the west, to such an extent that it almost reads as revisionist. Whereas *Kingdom of Desire* found commonalities and points of commiseration with Shakespeare's work, *Who Is Macbeth* seemingly intends to overwrite Shakespeare's same work. This progression reveals that since the Cultural Revolution, China has steadily regained confidence in itself and now, moreso than ever, holds itself as a major power capable of rivaling the west.

However, the progression of the format of these adaptations actually runs contrary to the progression of the ideas displayed in these two adaptations. *Kingdom of Desire* is adapted directly into *jingju* form, which is the official traditional operatic form of Chinese theater, as sanctioned by the PRC government, even through the Cultural Revolution (*The Soul*, 8). Traditional Chinese operatic forms, especially *jingju*, serve as a powerful representation of Chinese history, passed down through 5000 years of Chinese civilization, along with countless strict rules for how to create meaning on the stage. Costumes, arias, and *mise-en-scenes* of *Kingdom of Desire* are passingly faithful to the traditions of *jingju*, though it is important to note that even in 1986, many critics and citizens alike decried *Kingdom of Desire* for doing away with

too much of traditional *jingju* convention, from the complexity of Macbeth and Lady Macbeth's characters—which do not align with the traditional role types of *jingju*—to the uncharacteristic movements being utilized on stage (*The Soul*, 246). These elements had to bend to accommodate the story of Macbeth, which was not in any way like the traditional folktales and dramatized history usually enacted on the *jingju* stage.

However, the more recent *Who Is Macbeth* reveals that attitudes towards the west are not simply shifting towards hostility; there is still a perhaps unconscious respect or acknowledgement towards the traditions of the west. The format of *Who is Macbeth* has deviated away from traditional Chinese practice. Though *Who Is Macbeth* preserves a sparse and economical *mise-en-scene*, traditional Chinese costuming, and an older style of dialogue commonly used in traditional Chinese theater, practically all other aspects of *Who Is Macbeth* have been entirely westernized. Information is communicated largely through spoken word, meaning no arias, no regimented gestures, and incredibly limited martial arts and dance. Though spoken word theater has existed in China since the beginning of the 20th century, citizens living immediately post-Cultural Revolution still would've scoffed at such a production even being categorized as theater; if singing, dancing, and acrobatics are not all showcased, there would be no point in watching. Additionally, though *jingju* has been historically directed by the actors themselves, whose performances and choices (like adding culturally relevant references into the performance) have been the main driving force behind *jingju* performance, modern *jingju* performance has been shifting towards the director-actor model employed in the west (Fu, 3). So, contrary to the shift towards China exhibited in the content of China's Shakespearian adaptations, a diametrically opposite shift towards western styles can be observed in the format of said adaptations. In modern Chinese society, *jingju* is falling out of favor with younger generations while western styles have been blossoming. For example, although it is facing struggles with censorship despite its focus on apolitical, quotidian topics, stand-up comedy enjoyed a huge surge in popularity during the pandemic years that is still ongoing (Wang).

Though China has started to hold its culture as an equal to the west, it has moved away from doing so in a Chinese method, so to speak; today, it is adopting western traditions and methodology, attempting to compete with the west on the west's terms, if not in theater, then also in many other aspects of its society. Not only can this trend be seen in the most recent adaptation of Macbeth as examined in this essay, it can be seen in the likes of the movie *The Wandering Earth* (2019). Though the movie prioritized more Chinese values in the emphasis that a group of heroes as opposed to a singular hero saved the day in the end, it still sought after the style of the American Hollywood blockbuster in all its special effects. Another example is the recent video game *Black Myth: Wukong* (2024), which, despite having strong traditionally Chinese story elements from *Journey to the West*, is still categorized by both westerners and Chinese themselves as incorporating and borrowing the “souls-like” style, which was a style established by the popular western game *Dark Souls*. It seems that despite rediscovering pride in its own culture, Chinese tradition continues to be left behind. Judging by the trends in Chinese media of all types, not just theater, this development will continue on a wide scale into the 21st century.

Works Cited

- Cast, Chuanqi. 慾望城國 = The Kingdom of Desire. 法蘭克福國際工作室出版 ; 成信文化總經銷, 2000.
- Chaffee, John W. Branches of Heaven: A History of the Imperial Clan of Sung China. Harvard University Asia Center : Distributed by Harvard University Press, 1999.
- Durdin, Tillman. "China Transformed by Elimination of 'Four Olds.'" The New York Times, The New York Times, 19 May 1971, www.nytimes.com/1971/05/19/archives/china-transformed-by-elimination-of-four-olds.html.
- Fu, Jin. Chinese Theatre. Cambridge University Press, 2012.
- Gushikuang, Guoguo, director. Baidu, 2022, <http://baike.baidu.com/l/RIJWgDle>. Accessed 2024.
- Li, Ruru, and Eugenio Barba. The Soul of Beijing Opera: Theatrical Creativity and Continuity in the Changing World. Hong Kong University Press, 2010.
- Li, Ruru. Shashibiya: Staging Shakespeare in China. Hong Kong University Press, 2012.
- Neves, Kathryn. "King James I and Macbeth." Utah Shakespeare Festival, 2019, www.bard.org/news/king-james-i-and-macbeth/.
- Song, Jigang. "《谁是麦克白》云南艺术学院演出_哔哩哔哩_bilibili." _哔哩哔哩_bilibili, 21 Dec. 2022, www.bilibili.com/video/BV1884y1s7go/?spm_id_from=333.337.search-card.all.click&vd_source=324fcd4011ca4b72f84b25647ab704b3.
- Wang, Wei. "China's Growing Comedy Scene Feels Censorship Chill." BBC News, BBC, 4 June 2023, www.bbc.com/news/world-asia-china-65776797.

Unraveling Japanese Encephalitis: Insights into Pathogenesis, Immune Response, and Biomarker Development By Miwa Davis

Abstract

Japanese encephalitis, or JE, is a form of encephalitis that leads to drastic inflammation of the brain. It is a viral infection that spreads through the JE virus, a single stranded RNA virus belonging to the flavivirus family (Misra and Kalita). The transmission of this virus occurs through a zoonotic cycle including culex mosquitoes, pigs, and ardeid water birds. Humans serve as a dead end host for the virus (Unni et al.). Although a vaccine has been developed for JE, the exact mechanisms of the infection are still being studied. The blood brain barrier of the central nervous system, otherwise known as BBB, is a selectively permeable membrane that is responsible for regulating the chemical microenvironment of neurons in the brain (“Blood–brain Barrier”). In patients with JE, the neuroinvasive virus invades the BBB, gaining access to the central nervous system. Numerous intracellular mechanisms such as the “trojan horse” method, paracellular entry, and translocation of cerebral spinal fluid [CSF] are used by the Japanese encephalitis virus (Hsieh and St John). JE is endemic in various regions in Eastern and Southeastern Asia, resulting in approximately 50,000 cases and 15,000 deaths annually (Solomon et al.). Additionally, numerous studies have given potential candidates for JE biomarkers; however new long-term experiments can enhance the accuracy of JE diagnosis and neurodegenerative treatment.

Epidemiology and Public Health Aspects

Japanese Encephalitis was first identified in horses and humans as early as 1871, yet scientists are still conducting research on the virus for further understanding (Solomon et al.). While the exact origins of the virus are unknown, laboratory studies have revealed phylogenetic similarities to African viruses, such as West Nile Virus. An epidemic was first reported in Japan in 1924, and subsequent decades saw major outbreaks affecting over 6,000 civilians approximately every 10 years. Notably, the JE virus was finally isolated from a human brain sample in Tokyo in 1934 (Misra and Kalita).

While JE is traditionally considered endemic, epidemic patterns have been traced in regions such as Taiwan, North Korea, Northern Vietnam, and Northern India (Wang and Liang). Although adults still may be infected, JE is known to primarily affect children (World Health Organization: WHO). Most individuals living in rural areas possess immunity due to childhood infections of encephalitis. To mitigate the spread of JE, the World Health Organization, WHO, has recommended several public health measures such as: 1) Increasing immunization in endemic regions 2) Modifying agricultural practices to reduce risks of transmission 3) Implementing preventative measures such as mosquito nets and repellants (World Health Organization: WHO).

Disease Mechanism

Japanese Encephalitis (JE) infection begins with a bite from a mosquito who carries the pathogen. The JE virus enters the body, primarily infecting the dendritic cells, fibroblasts, endothelial cells, and lymph nodes (Ashraf et al.). This is followed by primary asymptomatic viremia, in which the virus circulates the bloodstream.

To initiate the neuroinvasion, the virus crosses the endothelial cell membrane through passive transport, utilizing a concentration gradient to move without the input of energy. Endothelial cells, which form sheets that separate bodily compartments, enable the virus to pass through capillary walls: a process known as diapedesis or transendothelial migration (Ashraf et al.). This process is a part of the innate immune response, acting as the body's first line of defense against pathogens (Alberts et al.).

Diapedesis triggers the inflammatory response, an automatic reaction to tissue damage that consists of soluble blood elements and leukocytes relocating to the site of the injury. Transendothelial migration (TEM) allows leukocytes to move through the subendothelial basement membrane, as they cannot self-propel (Muller). While TEM typically occurs via paracellular invasion, other mechanisms such as transcellular transport, and retrograde neuronal transport are being explored in relation to JE (Ashraf et al.).

Paracellular transport occurs between cells through intercellular shunt pathways that are driven by passive diffusion—the net movement from higher to lower concentrations (Yu). More specifically, paracellular transport involves passing through the epithelium, the thin layer of tissue above the organs and glands of the body (“NCI Dictionary of Cancer Terms”). By traveling inside intercellular spaces, this form of transport is under the control of the tight junctions that constitute the major rate limiting barrier towards ion permeation (Maiti). Although paracellular transport is often viewed as restrictive, variations in electrical resistance and ion selectivity impact the extent of protein delivery. Shared biophysical properties within the junctions of adjacent cells such as size, charge selectivity, permeability, and pH sensitivity are corroborated by the function of Claudin proteins (Maiti). This family consists of at least 18 proteins that define apical and basal membrane polarity by limiting exchange within the membrane (Heiskala et al.). Subsequently, when the tight junctions become compromised, the BBB weakens, allowing JEV to invade the brain.

In contrast, transcellular transport involves active processes driven by energy from adenosine triphosphate (ATP), where molecules are transmitted through membrane proteins (Chen and Lui). The energy originates from hydrolysis, where ATP is broken down into adenosine diphosphate (ADP) and an inorganic phosphate (Yu). This process can occur via macropinocytosis, micropinocytosis, or exocytosis. Micropinocytosis and macropinocytosis are forms of endocytosis, where cells internalize molecules through vesicles—a small liquid filled sac (“NCI Dictionary of Cancer Terms”). On the contrary, exocytosis is where substances are expelled through membrane fusion with vesicles. In this way, JEV gains the ability to directly infect endothelial cells in order to access the brain parenchyma (Hsieh et al.).

The “Trojan Horse” mechanism is another method of neuroinvasion, where JEV disguises itself as a benign entity in order to enter the central nervous system (CNS). After

infecting a phagocyte, JEV manipulates the cell to migrate towards the brain (Alberts et al.). Conversely, the JEV may suppress phagocyte activation, enabling infected cells to reach the brain through normal circulation. Once in the brain, infected phagocytes cross the BBB by attaching to brain capillaries through the previously mentioned pathways of paracellular or transcellular transport (Muller).

The inflammatory response in the brain stems from the replication of Ly6C^{hi} monocytes, which migrate into the CNS via CCR2 signaling to induce inflammation. Following entry to the CNS, JEV activates microglia and astrocytes, triggering the neuroinflammatory responses through signaling cascades that produce inflammatory mediators. In studies involving cultured mice injected with JEV, activation of TLR-3 and RIG-I-mediated ERK/MAP-Kp38/AP-1/NF- κ B signaling cascades has been observed (Ashraf et al.). This resulted in the formation of inflammatory cytokines, associated with symptoms such as the confusion, motor dysfunction, and seizures.

Although JEV does not directly damage the BBB, its presence simulates cellular responses that compromise the barrier. As an arbovirus, JEV causes viremia, often leading to asymptomatic cases or flu-like symptoms. Fever is a hallmark symptom of numerous infectious diseases including JE, and is the result of an immune response to a virus. When the JEV infects the body, it is detected by both pattern recognition receptors and Toll-like receptors (PRRs and TLRs) on immune cells. PRRs detect specific evolutionarily conserved structures on pathogens called PAMPs: pathogen-associated molecular patterns. After detecting PAMP engagement, PRRs trigger intracellular signaling cascades that lead to the expression of proinflammatory molecules (Mogensen). TLRs belong to the family of PRRs, and in addition to PAMP recognition, these molecules activate the immune response and induce co-stimulatory molecules (Sameer and Nissar). After the JEV pathogen is recognized by receptors, immune cells such as macrophages and dendritic cells release pro-inflammatory cytokines such as interleukin-1 beta (IL-1 β), interleukin-6 (IL-6), and tumor necrosis factor alpha (TNF- α) (Faraj and Jalal). IL-1 β , IL-6 , and TNF- α are all involved in the process of pathological pain (Zhang and An). IL-1 β has been found to increase the production of substance P and prostaglandin E₂ (PGE₂) in neuronal cells (Faraj and Jalal). After the cytokines travel to the hypothalamus, the hormone center of the brain, they bind to receptors on hypothalamic endothelial cells, leading to additional production of PGE₂ from the enzyme cyclooxygenase-2 (COX-2) (Jin et al.). The PGE₂ proceeds to interact with receptors on the hypothalamus where preoptic neurons are silenced, causing a disinhibition of presympathetic neurons (Maiti). The pre optic area (POA) of the hypothalamus controls thermoregulatory responses (Osaka). This interaction leads to an elevation in body temperature, resulting in fever accompanied by shivering, vasoconstriction, and behavioral changes (Dall and Stanford).

Seizures, particularly in children, are another frequent symptom associated with JE. In addition to the pro-inflammatory response induced by the cytokines, damaged neurons release excessive levels of glutamate (Banerjee and Tripathi). Glutamate serves as the major excitatory neurotransmitter in the nervous system (National Academies Press (US)). N-methyl-D-aspartate

(NMDA) receptors are glutamate-gated channels that represent a critical component of the CNS. Abnormal NMDA receptor expression can lead to cognitive defects and neurological disorders, including seizures (Blanke and VanDongen). When neurons are infected, excitotoxicity arises from overstimulation of NMDA receptors. Excitotoxicity is defined as the potential of cell death from excessive neuronal activism (Blanke and VanDongen). Hyperactivation of NMDA channels result in an influx of Ca^{2+} ions, damaging neuronal cell structures such as proteases, lipases, and endonucleases (Blanke and VanDongen). Seizures originate from high frequency action potential bursts and neuronal hypersynchronization (Bromfield et al.). This is reflected as “spikes” on electroencephalography (EEG). The combined effects of neuronal cell death and excitotoxicity lead to abnormal electrical activity, triggering seizures. This activity can spread to adjacent neurons through cortical connections and long association pathways such as the corpus callosum (Bromfield et al.).

Treatments

As research into the exact mechanisms of JE continues, there remains no definitive cure. However, several vaccines have been developed as preventative measures. These vaccines stimulate the immune system to produce antibodies against JEV. The first vaccine was derived from mouse brain tissue in the 1930s (*Inactivated Japanese Encephalitis Virus Vaccine Recommendations of the Advisory Committee on Immunization Practices (ACIP)*). Currently, there are four variations of the JE vaccine available: 1) Mouse brain-derived killed-inactivated vaccine: This is the most common form of the vaccine given in Asia. It was first isolated from the brain of a patient infected with JE. By injecting the virus into suckling mice, the vaccine was made by inactivating the supernatant of the brain homogenate with formalin. Then, the virus would be suspended by ultracentrifugation (Yun and Lee). 2) Cell culture-derived live-attenuated vaccine: This version of the vaccine was created by a SA₁₄-14-2 strain from a group of *Culex Pipiens* mosquito larvae in East Asia. This vaccine consists of a combination of infected hamster kidney cells and plaque purification from chick embryos (Erra et al.). 3) Cell culture-derived killed-inactivated vaccine: This version ensures the inactivation of JEV while allowing the immune response to continue without the risk of infection (Yun et al.). 4) Genetically engineered live-attenuated chimeric vaccine: A combination of different viral strains for maximum efficacy and safety (Yun and Lee).

While the mouse brain-derived killed-inactivated vaccine provides the most benefits, its drawbacks include high production costs, the risk of post-vaccination infections such as disseminated encephalomyelitis (another brain/spinal cord inflammatory condition, similar to encephalitis and meningitis), and the necessity for multiple doses and boosters. These difficulties led to the creation of the live-attenuated vaccines that have been licensed to multiple countries in Asia, including China, South Korea, Nepal, and India (Yun and Lee).

Research Directions

Many efforts are currently being made to acquire more knowledge about the cellular components of JEV infection. Several clinical trials are currently underway exploring broad spectrum (non-specific) antiviral molecules. Among these, rosmarinic acid has been investigated for its ability to lower the viral replication of GP78 in mouse brains, alongside the phytochemical curcumin. Additionally, minocycline has shown promise by deregulating the ubiquitin protease system, thereby mitigating neuronal apoptosis and offering protective effects on the BBB. Another potential form of treatment involves nucleic-acid based anti-JEV molecules. These particles target the viral genome of JEV, preventing the propagation of transcription and translation. Although this approach has proved to lack sensitivity compared to that of broad spectrum antiviral molecules, it has shown promise in preventing infection in both in vitro and in vivo studies. Small interfering RNA (siRNA) and microRNA (miRNA) have been utilized to silence the viral genes, linking methylene morphine ring backbones with phosphorodiamidate groups to achieve this effect (Yun and Lee).

Biomarkers

Biomarkers (i.e biological markers) are an integral resource used by scientists as a form of diagnosis and characterization. Otherwise known as molecular markers or signature molecules, these traits can be used as a prediction or sign of an underlying condition (Benisek). By connecting a biomarker to a pathophysiological effect in the body, clinical endpoints and intermediate outcomes can be observed in a clinical setting (Aronson and Ferner). In the context of JE, biomarkers have gained increasing importance due to clinical similarities between JE and other neurological conditions. Accurate diagnosis with the help of clinical biomarkers can be a vital tool for differentiation between JE and other disorders, enabling prompt and targeted treatment.

This is particularly critical for pediatric patients who are at the highest risk for JE. Early intervention is essential in order to prevent life-long neurological deficits. The following analysis goes through multiple candidates in modern medicine that have the potential to serve as biomarkers for JE.

Molecular Markers

The first molecular marker under consideration is a neurological mutation. In a National Institutes of Health (NIH)-funded study investigating the molecular basis of JE, a specific strain of the JE vaccine (SA14-14-2) was cloned, leading to the creation of a virus derived from its complementary DNA (cDNA) termed SA14-14-2 MCV. To identify the specific variants of the SA14-14-2 MCV strain that contribute to neurovirulence, cell cultures of both the parent and cloned virus were analyzed within the BHK-21 cells of hamster kidneys, a common medium used in JEV propagation. In Vivo traits were compared in a study involving 3-week-old ICR mice (n=20) who were infected with both SA14-14-2 and SA14-14-2 MCV as detailed in Table 1 (Yun et al.)

Table 1: Comparison of In Vitro Growth Properties and In Vivo Phenotypic Traits of JE Virus Strains

Virus	In vitro growth properties	In vivo phenotypic traits
SA14-14-2	<ul style="list-style-type: none"> - Uncloned parent virus - Foci/plaques ~30% smaller at 4-days post infection than CNU/LP2 (reference strain) 	<ul style="list-style-type: none"> - No clinical manifestations of JEV intramuscular (IM), and intraperitoneal (IP) inoculation - 5-10% of infected mice developed clinical symptoms/death after high dosage intracerebral (IC) inoculation
SA-14-142MCV	<ul style="list-style-type: none"> - Consistent replication rate with parent virus - Similarities in the propagation of genomic RNA in the 24-hour period after infection (1 PFU/cell) - Growth kinetics at steady range for 4 days following infection in comparison to parental virus at a multiplicity of infection (MOI) of 0.1, 1, and 10 PFU/cell - foci/plaques ~30% smaller at 4-days post infection than CNU/LP2 (reference strain) 	<ul style="list-style-type: none"> - No clinical manifestations of JEV intramuscular (IM), and intraperitoneal (IP) inoculation - 5-20% of infected mice developed clinical symptoms/death after high dosage intracerebral (IC) inoculation

The abundant similarities between the parent and cloned virus allowed for an investigation in the genetic virulence factor of JEV using SA14-14-MCV. A key genetic mutation was found in this study through the creation of 8 derivatives of the cloned virus, with each containing 1 of 8 different point mutations found through genetic analysis of the SA14-14-2 MCV using synthetic RNA transfection. Following a methodology similar to previous studies comparing in vivo traits between SA14-14-2 and SA14-14-2 MCV, experiments were conducted with n=10 ICR mice per grouping. In the group containing the G¹⁷⁰⁸A substitution, an IC LD₅₀ of <1.5 PFU (plaque-forming units) was found, compared to the IC LD₅₀ values of >1.5x10⁵ PFU in the other 7 mice groupings (Yun et al.) IC LD₅₀, otherwise known as the median lethal dose, is a dosage that will lead to incapacitation in a given subject (Erhirhie et al.). This meant that the G¹⁷⁰⁸A substitution had the capability of killing mice within ~7 dpi, or days post infection, when exposed to a dose of 1.5 PFU/mouse. This highly lethal mutation consisted of a single amino acid substitution from Glycine to Glutamic acid (Gly->Glu) at position E-244 (residue 244 of the E glycoprotein). The E-244 mutation exhibits promise as a biomarker for JE due to the implications it has towards neurovirulence in encephalitis (Yun et al.).

CSF Biomarkers

CSF, or cerebrospinal fluid biomarkers, is already showing significant promise in conditions such as Alzheimer's disease; However, there is a significant gap in research development and molecular knowledge for CSF biomarkers in neurological diseases such as JE (Teunissen et al.). CSF has a close proximity to the brain, as its homeostasis is regulated by the BBB and blood-cerebrospinal fluid barrier, BCB (Kadry et al.). A study conducted by King

George's Medical University (KGMU) hospital investigated CSF samples from both patients suffering from JE and those of a healthy neurological condition. By utilizing a proteomic comparison, potential CSF biomarkers for JE were identified as summarized in Table 2.

Table 2: CSF Proteins Associated with Japanese Encephalitis: A Compilation of Diagnostic and Proteomic Biomarkers

Protein Type	Expression: JE	Expression: AES (Acute encephalitis syndrome)	Conclusion	Citations
DBP (known as Vitamin D Binding Protein or Group Specific Component) Multi-functional protein. ~5µM in the human body.	Increased (Indicating higher levels of cytoplasmic actin in CSF)	Normal	Elevated DBP has shown to negatively affect recovery. Cytoplasmic action is related to severe axonal degeneration.	(Chun et al.) (Sengupta et al.) (Yang et. al.)
Complement C3 Protein in the immune system that protects against infection.	Increased (Implying increased activation of the complement system)	Normal	Increased levels of C3 protein have been linked to synaptic loss and neurological impairment.	(Chen et al.) (Sengupta et al.)
Complement C4b Protein in the immune system that protects against infection.	Increased (While C4b can clear harmful debris and pathogens, overactivation can cause exacerbate neuronal inflammation)	Normal	The complement proteins that injured neurons produce after acute neuronal injury can lead to BBB damage and the release of inflammatory anaphylatoxins.	(Orsini et al.) (Sengupta et al.)
Cytoplasmic Actin Protein element that is a major component of our cytoskeleton. Can be found in neurons.	Increased (In studies on Multiple Sclerosis, increases in extracellular gelsolin, GSN, and DBP promote an increase in actin that alters protein concentrations in CSF)	Normal	Increased cytoplasmic actin is an indicator of axonal degeneration in MS, so similar neurodegenerative effects may be present in JE.	(Kulakowska et al.) (Sengupta et al.) (Venkatesh et al.)
Fibrinogen Gamma Chain A glycoprotein that takes part in the inflammatory response in the	Increased (Fibrinogen is normally not present in the brain due to blockage by the BBB.	Normal	Fibrinogen has been seen to exacerbate neuropathological symptoms including	(Petersen et al.) (Sengupta et al.) (Wen et al.)

brain.	Increased levels indicate disruption of the BBB in patients with JE)		inflammation, PNS/CNS damage, and oxidative stress. Fibrinogen is also a promoter of cognitive decline.	
Serum Albumin Protein that plays a critical role in brain health. Is present in relatively low concentrations except in CSF. Is the main protein of blood plasma.	Increased (Increased serum albumin has been seen in patients with dementia, and is associated with neuronal apoptosis and tauopathies)	Normal	Enhanced levels of serum albumin is present in the brain after BBB disruption, leading to neurodegeneration.	(Cui, Yiyuan, et al.) (Hou, Sheng-Jie, et al.) (Moser et al.) (Shojai et al.)

Conclusion

As Japanese Encephalitis continues to become an increasingly prevalent neuroinflammatory disease, accurate diagnosis is necessary in order to treat patients before the virus leads to long-term neurological deficits. Throughout this review, specific mutations such as the G¹⁷⁰⁸A substitution showed potential in serving as a candidate for neurovirulence detection. Numerous CSF proteins have shown promise in current JE research, including: DBP, complement proteins C3 and C4b, cytoplasmic actin, fibrinogen, and serum albumin. However, most significant findings to date have emerged from studies related to other diseases, such as multiple sclerosis (MS). Consequently, there remains a significant gap in research specifically addressing the roles these proteins play in the mechanisms underlying JE.

Current research limitations further restrict our understanding of JE. For example, while in vitro studies provide valuable insights into viral mechanisms, they do not accurately represent the complexity of neuronal tissues in living organisms. Moreover, the lack of long-term studies on JE narrows the understanding of the efficacy of potential biomarkers. These challenges highlight the pressing need for further research to validate biomarker candidates, ultimately promoting early intervention and improved patient recovery outcomes in JE patients.

Works Cited

- Alberts, Bruce, et al. "Innate Immunity." *Molecular Biology of the Cell* - NCBI Bookshelf, 2002, www.ncbi.nlm.nih.gov/books/NBK26846/#:~:text=The%20innate%20immune%20responses%20are,initiate%20specific%20adaptive%20immune%20responses.
- Aronson, Jeffrey K., and Robin E. Ferner. "Biomarkers—A General Review." *Current Protocols in Pharmacology*, vol. 76, no. 1, Mar. 2017, <https://doi.org/10.1002/cpph.19>.
- Ashraf, Usama, et al. "Pathogenicity and Virulence of Japanese Encephalitis Virus: Neuroinflammation and Neuronal Cell Damage." *Virulence*, vol. 12, no. 1, Mar. 2021, pp. 968–80. <https://doi.org/10.1080/21505594.2021.1899674>.
- Banerjee, Arup, and Aarti Tripathi. "Recent Advances in Understanding Japanese Encephalitis." *F1000Research*, vol. 8, Nov. 2019, p. 1915. <https://doi.org/10.12688/f1000research.19693.1>.
- Benisek, Alexandra. "What Are Biomarkers?" WebMD, 13 Jan. 2023, www.webmd.com/a-to-z-guides/biomarkers-overview.
- Blanke, Marie L., and Antonius M. J. VanDongen. "Activation Mechanisms of the NMDA Receptor." *Biology of the NMDA Receptor* - NCBI Bookshelf, 2009, www.ncbi.nlm.nih.gov/books/NBK5274.
- Bromfield, Edward B., et al. "Basic Mechanisms Underlying Seizures and Epilepsy." *An Introduction to Epilepsy* - NCBI Bookshelf, 2006, www.ncbi.nlm.nih.gov/books/NBK2510.
- Cancer.gov. "NCI Dictionary of Cancer Terms." www.cancer.gov/publications/dictionaries/cancer-terms/def/epithelium#.
- Cui, Yiyuan, et al. "Protective Role of Serum Albumin in Dementia: A Prospective Study From United Kingdom Biobank." *Frontiers in Neurology*, vol. 15, Aug. 2024, <https://doi.org/10.3389/fneur.2024.1458184>.
- Dall, Lawrence, and James F. Stanford. "Fever, Chills, and Night Sweats." *Clinical Methods* - NCBI Bookshelf, 1990, www.ncbi.nlm.nih.gov/books/NBK324.
- Erra EO, Kantele A. The Vero cell-derived, inactivated, SA14-14-2 strain-based vaccine (Ixiaro) for prevention of Japanese encephalitis. *Expert Rev Vaccines*. 2015;14(9):1167-79. doi: 10.1586/14760584.2015.1061939. Epub 2015 Jul 10. PMID: 26162529.
- Erhirhie, Earnest Oghenesuvwe, et al. "Advances in Acute Toxicity Testing: Strengths, Weaknesses and Regulatory Acceptance." *Interdisciplinary Toxicology*, vol. 11, no. 1, May 2018, pp. 5–12. <https://doi.org/10.2478/intox-2018-0001>.
- Faraj, Shalaw Sardar, and Paywast Jamal Jalal. "IL1 β , IL-6, and TNF- α Cytokines Cooperate to Modulate a Complicated Medical Condition Among COVID-19 Patients: Case-control Study." *Annals of Medicine and Surgery*, vol. 85, no. 6, Apr. 2023, pp. 2291–97. <https://doi.org/10.1097/ms9.0000000000000679>.
- Heiskala, Marja, et al. "The Roles of Claudin Superfamily Proteins in Paracellular Transport." *Traffic*, vol. 2, no. 2, Jan. 2001, pp. 92–98. <https://doi.org/10.1034/j.1600-0854.2001.020203.x>.
- Hsieh, Justin T., and Ashley L. St John. "Japanese Encephalitis Virus and Its Mechanisms of Neuroinvasion." *PLoS Pathogens*, vol. 16, no. 4, Apr. 2020, p. e1008260. <https://doi.org/10.1371/journal.ppat.1008260>.

- Inactivated Japanese Encephalitis Virus Vaccine Recommendations of the Advisory Committee on Immunization Practices (ACIP). 8 Jan. 1993, www.cdc.gov/mmwr/preview/mmwrhtml/00020599.htm#:~:text=An%20inactivated%20JE%20vaccine%20derived,distributed%20by%20Connaught%20Laboratories%20Inc.
- Japanese Encephalitis. 6 Aug. 2024, www.who.int/news-room/fact-sheets/detail/japanese-encephalitis.
- Jin, Kaipeng, et al. "Cyclooxygenase-2-Prostaglandin E2 Pathway: A Key Player in Tumor-associated Immune Cells." *Frontiers in Oncology*, vol. 13, Jan. 2023, <https://doi.org/10.3389/fonc.2023.1099811>.
- Kadry, Hossam, et al. "A Blood–brain Barrier Overview on Structure, Function, Impairment, and Biomarkers of Integrity." *Fluids and Barriers of the CNS*, vol. 17, no. 1, Nov. 2020, <https://doi.org/10.1186/s12987-020-00230-3>.
- Maiti, Sabyasachi. "Nanometric Biopolymer Devices for Oral Delivery of Macromolecules With Clinical Significance." Elsevier eBooks, 2017, pp. 109–38. <https://doi.org/10.1016/b978-0-323-52725-5.00006-x>.
- Mogensen, Trine H. "Pathogen Recognition and Inflammatory Signaling in Innate Immune Defenses." *Clinical Microbiology Reviews*, vol. 22, no. 2, Apr. 2009, pp. 240–73. <https://doi.org/10.1128/cmr.00046-08>.
- Muller, William A. "Mechanisms of Leukocyte Transendothelial Migration." *Annual Review of Pathology Mechanisms of Disease*, vol. 6, no. 1, Jan. 2011, pp. 323–44. <https://doi.org/10.1146/annurev-pathol-011110-130224>.
- National Academies Press (US). "Overview of the Glutamatergic System." *Glutamate-Related Biomarkers in Drug Development for Disorders of the Nervous System - NCBI Bookshelf*, 2011, www.ncbi.nlm.nih.gov/books/NBK62187.
- Orsini, Franca, et al. "Versatility of the Complement System in Neuroinflammation, Neurodegeneration and Brain Homeostasis." *Frontiers in Cellular Neuroscience*, vol. 8, 2014, doi:10.3389/fncel.2014.00380.
- Petersen, Mark A., et al. "Fibrinogen in Neurological Diseases: Mechanisms, Imaging and Therapeutics." *Nature Reviews Neuroscience*, vol. 19, no. 5, 2018, pp. 283–301. doi:10.1038/nrn.2018.13.
- Sameer, Aga Syed, and Saniya Nissar. "Toll-Like Receptors (TLRs): Structure, Functions, Signaling, and Role of Their Polymorphisms in Colorectal Cancer Susceptibility." *BioMed Research International*, vol. 2021, Sept. 2021, pp. 1–14. <https://doi.org/10.1155/2021/1157023>.
- Sengupta, Nabonita, et al. "Cerebrospinal Fluid Biomarkers of Japanese Encephalitis." *F1000Research*, vol. 4, 23 July 2015, doi:10.12688/f1000research.6801.1.
- Shojai, S., et al. "Human Serum Albumin in Neurodegeneration." *Reviews in Neuroscience*, vol. 33, no. 7, 2022, pp. 803–17. doi:10.1515/revneuro-2021-0165.
- Solomon, Tom, et al. "Origin and Evolution of Japanese Encephalitis Virus in Southeast Asia." *Journal of Virology*, vol. 77, no. 5, Feb. 2003, pp. 3091–98. <https://doi.org/10.1128/jvi.77.5.3091-3098.2003>.

- Unni, Salini Krishnan, et al. "Japanese Encephalitis Virus: From Genome to Infectome." *Microbes and Infection*, vol. 13, no. 4, Jan. 2011, pp. 312–21. <https://doi.org/10.1016/j.micinf.2011.01.002>.
- Wang, Huanyu, and Guodong Liang. "Epidemiology of Japanese Encephalitis: Past, Present, and Future Prospects." *Therapeutics and Clinical Risk Management*, Mar. 2015, p. 435. <https://doi.org/10.2147/tcrm.s51168>.
- Yun, Ki Wook et al. "Safety and immunogenicity of a freeze-dried, Vero cell culture-derived, inactivated Japanese encephalitis vaccine (KD-287, ENCEVAC®) versus a mouse brain-derived inactivated Japanese encephalitis vaccine in children: a phase III, multicenter, double-blinded, randomized trial." *BMC infectious diseases* vol. 15 7. 8 Jan. 2015, doi:10.1186/s12879-014-0744-4

Investments and Maintenance: Their Impact on Perceived Closeness in Adolescent Friendships By Zhen Xu

Abstract

Friendships are on the third level of Maslow's hierarchy of needs, meaning it is a fundamental aspect of human lives. While much is known about the basic qualities that define friendship, research on how specific factors contribute to the deep connections between friends is limited. This study explores how investments and maintenance levels affect the perceived closeness of a friendship. It is hypothesized that friendships characterized by significant investments and high maintenance efforts will correlate with an increased perception of closeness within that friendship, and there will be statistically significant differences between the results of male and female participants. A total of 34 participants (7 male, 25 female, and 1 preferred not to say), aged 15-18, completed a survey that measured perceived closeness, investment levels, and maintenance levels. Results showed a moderate positive correlation between investment behaviors and perceived closeness¹ (measured by the Intimate Friendship Scale), and that maintenance behaviors were significantly correlated with perceived closeness² (measured by the Inclusion of Other in Self Scale). Additionally, investment and maintenance were significantly associated with each other. Investment levels were measured to be marginally significantly higher for women, while no significant gender differences occurred in levels of maintenance or perceived closeness. This research highlights the importance of active investment and maintenance in fostering the sense of closeness in friendships. This provides insights for the real world in bettering personal relationships and community interventions.

Keywords: investments, friendship maintenance, perceived closeness, friendship dynamics

Introduction

“Friendship is ‘necessary’ for the pursuit of wisdom, necessary in the same way that a horse is for a long journey” (Cuddeback 2004, p. 33). Friends may provide emotional and moral support to encourage oneself to embrace challenges; friends from diverse backgrounds can introduce a variety of perspectives that enriches oneself's understanding of the world; and friends can offer honest feedback constructively to help oneself refine their thoughts or actions. These aspects of friendships all make friendship itself an essential component of individual social and intellectual fulfillment.

Most people have friends, but their standards for whom may be considered a “friend” may differ slightly. There is no objective definition of what a friendship is. The term's meaning may vary to different people, so “friendship may be best conceptualized as a prototype concept” (Fehr 1996, p. 6). Based on the vague prototype, each individual develops a personal understanding of what friends are. For the purposes of this study, friendships are defined as a “voluntary interdependence between two persons over time, that is intended to facilitate socio-emotional goals of the participants, and may involve varying types and degrees of

companionship, intimacy, affection, and mutual assistance” (Hays 1988, p. 395). Friendships are at the “Love and Belonging” level in Maslow’s hierarchy of needs. (Maslow 2023). This means that it is a fundamental need for us to connect with others. Friendships are also different from other relationships. Relationships with family members, neighbors, or co-workers are oftentimes results of biological networks or public obligation, whereas friendships can be formed on a voluntary basis, as we can freely choose with whom we want to befriend. This voluntary nature of friendships makes it a relatively accurate reflection of an individual’s identity and values.

Friendships also contribute to how we value ourselves. “The extent to which we value something can be influenced by our awareness of the value others attribute to it, [so] recognizing that others value our own lives may impact on our sense of self-worth” (Leibowitz 2018, p. 101). In other words, a sense of validation can be gained from friendships. Because of this, friendships are beneficial to each individual participant.

So, how do friendships form? The process of friendship formation spans over a long period of time. In the initial stages of friendship formation, proximity plays a key role—people are more likely to become friends with those they encounter frequently (e.g., classmates); the recurring physical closeness provides many opportunities for interactions (Rubin & Shenker 1978). Then through conversations and interactions, shared values may be discovered and shared experiences may be developed, and those similarities foster an immediate sense of connection between the two friends. After some level of familiarity has been established, the individuals become acquaintances. As they grow increasingly comfortable with each other, they may disclose more personal information. Disclosure deepens the bond between the two friends, because they demonstrate they are willing to be vulnerable and invest emotionally in the friendship; it is during this stage that they truly become friends (Field-Springer & Stephens 2016). Individual personality traits, mutual interest, shared experiences, and sociocultural contexts are all factors that can possibly alter the mechanisms through which friendships form and evolve. This paper seeks to focus specifically on the investments individuals make in their friendships, the maintenance efforts they employ, and examine how these efforts collectively influence the perceived closeness within these friendships.

Investments

Investments are the building blocks of a sturdy foundation in a friendship. Investment in a friendship refers to the magnitude and significance of the resources dedicated to it (Rusbult, 1980). These resources can include time and emotional effort, which indicate the commitment each person attributes to the relationship. Initial investments (e.g., offering support) help establish trust and build the foundation for the friendship’s development. Over time, sustained investments create a deeper connection and commitment.

In friendships, investment is the tangible and intangible resources that an individual puts into it, including time, emotional energy, financial resources, et cetera (Goodfriend & Agnew 2008). Commitment is not to be confused with investment: commitment is the intent to continue a relationship and the long-term orientation towards maintaining that friendship (Rusbult, Martz,

& Agnew 2005). In this study, commitment is being considered as a part of investment because (1) it is an emotional investment; (2) it motivates other types of investments; and (3) it sustains investments over time. Previous studies have investigated investment in friendships from a lens of cost-benefit analysis, which evaluates a friendship based on the amount of resources dedicated to it and the advantages gained from the friendship (Gompers 2016). There have also been several studies that examined the role of investment on the stability of friendships, but investment was seldom evaluated in relation to perceived closeness and maintenance (Branje et al. 2007; Fehr 1999; Young 1990).

Maintenance

Friendships develop over time and are regularly maintained. Individuals deliberately engage in friendship maintenance to sustain desired levels of friendship commitment and satisfaction (Oswald et al. 2004). Effective maintenance strategies include self-disclosure, providing support and assurance, maintaining levels of rewards, and shared time (Fehr 1996). The level of maintenance can be consciously manipulated by individuals. For example, one can choose to show “more affection, less antagonism, and to interact and share more with one’s friend” to increase the quality of their friendship (Huić & Smolčić 2015). Conflict resolution is also an important aspect of friendship maintenance—two unique individuals will inevitably disagree on some matters (Tehrani & Yamini 2020). Regular maintenance stabilizes friendships, and that sense of security is foundational for deeper emotional connections, and thus perceived closeness.

Perceived Closeness

Perceived closeness is a subjective sense of intimacy, trust, and mutual understanding between friends (Ben-Ari & Lavee 2007). This closeness reflects how an individual views another individual is involved with his/her/their life. A closeness in friendship is characterized by a strong, frequent, varied, and enduring pattern of interdependence (Kelley et al. 1983). It is often associated with the quality of friendships. In this study, perceived closeness will be quantified as a composite measure of these feelings and explore its correlation with investments and maintenance efforts within the friendship.

Gender Differences in Friendships

Research on gender differences in friendships indicates that men and women seek different qualities based on their functional needs and lean toward different investment or maintenance methods due to societal expectations. For example, a study (Williams et al. 2022) finds that men prioritize status and mate-seeking qualities, while women focus more on emotional support. Because men and women value different aspects of friendships, the ways in which they interpret perceived closeness and approach investment or maintenance methods will also likely be different. Previous research (Aukett et al. 1988) also suggests that the “socialization process of males” teaches them to “repress their emotions and form less intimate

and beneficial friendships than women.” Societal expectations play a fundamental role in shaping friendships, ultimately influencing the men and women to adopt different mechanisms in forming and maintaining friendships.

Overview of the Current Study

Past studies separately address investment, maintenance, and perceived closeness. This study seeks to integrate these three variables to assess their impacts on friendship dynamics and each other. Through empirical research, this paper aims to elaborate on how varying levels of investment and maintenance affect perceived closeness and, by extension, the overall quality of the friendship.

An improved understanding of these dynamics will not only be beneficial for academic purposes but also for ameliorated practical applications. Understanding how friendships work will make maintaining personal relationships much easier. Better designed community and therapeutic interventions will also eventually emerge from a more clear concept of friendships. Strong friendships are related to improved mental health, greater resilience to stress, and better physical health (Ozbay et al. 2007). By deepening our understanding of what strengthens friendships, this study contributes to broader efforts targeted at making social networks more supportive and enduring.

The relationships between the amount of investment, the quality of maintenance, and the level of perceived closeness is the core subject this paper aims to assess. This study hypothesizes that increased investment and maintenance in friendships will be strongly associated with higher levels of perceived closeness because previous research shows that investment requires contribution of personal resources and maintenance requires a significant amount of involvement in the friend’s life (Branje et al. 2007; Fehr 1996; Oswald et al. 2004; Young 1990). Thus, both investment and maintenance would lead to improved quality of a friendship. Additionally, the study predicts that women, on average, will show higher levels of investment, maintenance, and perceived closeness in their friendships compared to men because they tend to be more expressive and desire a more intimate friendship than men (Aukett et al. 1988).

Methods

Participants

34 participants (7 male, 25 female, and 1 preferred not to say) were recruited on a voluntary basis; the confidentiality of their responses was also clearly communicated at the beginning of the survey. This study was conducted in accordance with the ethical guidelines set forth by the American Psychological Association. All participants provided informed consent before participating in the survey, were assured of their anonymity, and were informed of their right to withdraw from the study at any point without penalty. The majority (24 out of 34) of the participants identified to be Asian. All but one participant were within the age range of 15-18. The demographic information collected included age, gender, and country of residence to explore

potential variations in friendship dynamics across different demographic groups or to note confounding variables. One participant was excluded from the data analysis due to a clear pattern of random response behavior, as evidenced by inconsistent answers and the unusually short duration taken to complete the survey.

Materials

This study used a survey to investigate the underlying relationship between perceived closeness, investments, and maintenance activities in friendships by quantifying each of the three variables. The survey was designed to be complete within 10 minutes. The survey was created using Microsoft Forms. Participants were informed beforehand to answer a series of questions related to how they feel toward the friendship they felt were the closest. The survey link was distributed online through various channels, including direct messages, social media platforms, and online forums dedicated to discussions about interpersonal relationships.

The survey consisted of three main sections, each designed to measure a specific aspect of friendships: perceived closeness, investments, and maintenance activities. The participant was asked to think of one specific friend (their closest friend) for the entire survey to ensure consistency.

Perceived Closeness. Perceived closeness was measured using two distinct scales (Appendix A and Appendix B). The first perceived closeness scale (closeness1) is the Intimate Friendship Scale (IFS), which quantifies closeness with a 5-part likert scale ranging from “Strongly Disagree” to “Strongly Agree” (Allgood 2009). 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree. The second perceived closeness scale (closeness2) is the Inclusion of Other in Self (IOS) Scale, which assesses the intimacy and closeness of relationships through graphical representations of overlapping circles (Gächter et al. 2015). Participants were asked to select the set of circles that best represented the closeness of their relationship with their friend.

Investments. Investments were assessed using the Investment Model Scale (appendix C) through a 5-part Likert scale ranging from "Strongly Disagree" to "Strongly Agree," 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree (Rusbult, Martz, & Agnew 2005). The “Satisfaction Level–Global items” and “Quality of Alternatives–Global items” portions of the scale were excluded for the purposes of this study: the investment defined in this research does not include satisfaction levels because satisfaction is not a part of the definition used for investments in this study, and the quality of alternatives applies more to romantic partners (due to the exclusive nature of romantic relationships) and not to friendships.

Maintenance. Maintenance levels were measured using the Factor Structure of the Friendship Maintenance Scale (appendix D) through a 5-part Likert scale related to the regularity of various friendship-sustaining behaviors, including communication frequency, participation in shared activities, and efforts to resolve conflicts (Oswald et al. 2004). The scale is designed to gauge the frequency and depth of maintenance activities, ranging from “Never” to “Almost always.” 1 = Never, 2 = Happened once or twice, 3 = Sometimes, 4 = Often, 5 = Almost always.

Results

It was hypothesized that friendships that are greater in investment and maintenance levels will have higher perceived closeness. It was also predicted that women would have higher perceived closeness, investment, and maintenance levels in general. A Spearman's rho was conducted to evaluate the relationship between investment levels and perceived closeness. There was a moderate positive correlation between investment and closeness1, which was statistically significant, $r_s(30) = .533, p = .002$. However, the correlation between investment and closeness2 was not statistically significant, $r_s(30) = .170, p = .351$, indicating no substantial association between these variables for this measure (Table 1).

Another Spearman's rho was conducted to test the relationship between maintenance behaviors and closeness. For the relationship between maintenance behaviors and closeness, the correlation with closeness1 (measured by the IFS) was marginally significant, $r_s(30) = .325, p = .069$. This suggests a possible positive association between maintenance behaviors and perceived closeness, but it does not meet the conventional threshold for significance. In contrast, maintenance was positively correlated with closeness2 (measured by IOS scale), $r_s(30) = .483, p = .005$, indicating a moderate and statistically significant relationship between maintenance behaviors and this closeness measure (Table 1).

Spearman's rho was also conducted to assess the relationship between investment and maintenance behaviors. A statistically significant positive correlation was found between investment and maintenance behaviors, $r_s(30) = .379, p = .032$, demonstrating that higher levels of investment are associated with greater levels of maintenance behaviors (Table 1).

The two closeness measures, interestingly, did not have a statistically significant correlation, $r_s(30) = .316, p = .078$, because this correlation did not reach statistical significance at the 0.05 level (Table 1).

The findings indicate that investment behaviors are significantly associated with closeness1, while maintenance behaviors show a significant association with closeness2. Although the two variables each correlate with a different variable of perceived closeness, investment and maintenance behaviors are statistically significantly correlated.

It was also hypothesized that women would in general have higher perceived closeness in their friendships. Independent samples t-tests were performed to evaluate whether women in general had higher levels of investment, maintenance, and perceived closeness than men (Table 2).

For investments, the results indicated that women ($M = 4.13, SD = 0.37$) had marginally significant higher investment than men ($M = 3.87, SD = 0.58$), $t(29) = 1.61, p = .059$.

For maintenance, the results indicated there was no significant difference between women ($M = 3.80, SD = 0.45$) and men ($M = 3.54, SD = 0.59$), $t(29) = 1.25, p = .221$. The results also showed there was no significant difference between friendship closeness of women ($M = 4.09, SD = 0.26$) and men ($M = 4.00, SD = 0.19$), $t(29) = 0.88, p = .386$.

Discussion

In the original hypotheses, this study proposed that greater levels of investment and maintenance in friendships will lead to a higher perceived closeness. It was also hypothesized that women will report greater investment, maintenance, and perceived closeness in their friendships compared to men.

The results somewhat support the first hypothesis, which stated that greater levels of maintenance and investments in friendships contributes to a higher sense of perceived closeness. The findings show that investment has a significant positive correlation with closeness1, while maintenance has a significant positive correlation with closeness2. A significant correlation with one of the closeness variables should mean a significant correlation with the other closeness variable, but that was not the case; this discrepancy will be discussed later in this section.

One of the key findings of this study was the significant positive correlation between investment and closeness1. This implies that individuals who invest more in their friendships tend to perceive greater levels of closeness, as supported by the investment model of relationships (Rusbult 1980). Higher levels of investment means the individual is likely highly committed in the friendship. Thus, high investment levels creates a sense of attachment that increases the overall perception of closeness. This finding reinforces the notion that friendships, like other close relationships, require investment to attain closeness.

The relationship between maintenance behaviors and closeness was more nuanced. While the correlation between maintenance and closeness1 was not statistically significant, a significant positive correlation was found between maintenance and closeness2. It is possible that closeness2 (perceived closeness measured by the IOS scale) correlated with friendship maintenance because higher levels of maintenance entails greater involvement of the friend in one's life, thereby leading to more overlap between the "self" and "other" circles. In addition, the two closeness tests had no significant correlation with one another either. The lack of correlation between the two closeness tests indicates the subjectivity of definitions of the "closeness" variable in psychology. The Intimate Friendship Scale (closeness1) defines closeness using three sources—the definition of friendship as defined by Webster's Dictionary of Synonyms, sociological studies on social distance, and psychoanalytical literature—while the Inclusion of Other in Self Scale (closeness2) views closeness as a subjective judgment, thus defining closeness by emphasizing how connected individuals feel in a relationship. It also may have affected the exact sense of closeness measured by the two scales, as the Intimate Friendship Scale (closeness1) uses verbal descriptions while the Inclusion of Other in Self Scale (closeness2) uses visual representations. Many variables, in the field of psychology, do not have uniform definitions, which somewhat hinders the progress of research.

Investment and maintenance were found to have a statistically significant positive correlation in the context of friendships. This correlation does not indicate a causal relationship, however. These two events are mutually inclusive and are likely to take place simultaneously. Investment requires emotional and time commitment to the friendship, and so

does maintenance. Both are effective methods that strengthen the friendship, therefore can often be seen together at similar relative levels within a friendship.

The second hypothesis was not supported by the data. The results revealed no significant gender differences in all three variables. There are no significant gender differences in either perceived closeness or maintenance, and there is merely a marginally significant difference between women and men's investments in friendships, with women's investment levels being higher overall. This finding is inconsistent with previous research that suggests that men and women may experience and express closeness differently (Feeney 1999; Sprecher 1992; Wood & Inman 1993). However, the validity of this finding is limited. In this study, the sample size for males is only 7, meaning the findings do not come from a representative sample that can be further generalized to a broader population. It is possible, though, that there are actually no significant gender differences in investment, maintenance, and perceived closeness. This would suggest a potential universality of friendship experiences across genders, but further research with a larger and more gender-representative sample is required to support this conclusion.

One variable that was overlooked in this study is the current phase of the friendship. Friendships are dynamic: they evolve and go through various stages, including periods of conflict where there exists a temporary estrangement. The neglect of this factor might explain some of the inconsistencies in the findings. For instance, a friendship in the midst of conflict may result in lower levels of perceived closeness at the time the survey is taken, regardless of high-level investment or maintenance efforts. This temporary period of low perceived closeness may skew the data. Future research should account for the current phase of the friendship (e.g., fighting, growing, stable) as a potential moderator in the relationship between investment, maintenance, and perceived closeness to provide more nuanced insights.

Another critical aspect of friendships that this study should have taken into account is individual desired-intimacy in a friendship. Not all individuals expect the same level of closeness in their friendships. Thus, the question of "how close do you want the friendship to be?" should have been included in this research, as it is a possible factor that affects perceived closeness correlates with investment and maintenance levels.

This study also failed to account for individual differences in personality, which vastly affects how people invest in and maintain their friendships. For example, some individuals may prefer shared experiences that require more time, while others may prioritize intimate conversations that require more emotional investment. These individual differences in how people prefer to engage in their friendships could impact the levels of investment and maintenance quantified by the scales used in this study. Future studies should consider including personality measures, such as extroversion, agreeableness, or openness to experience, to potentially examine how these traits influence investment, maintenance, and perceived closeness in friendships.

In an age where superficial connections often dominate, this research reinforces and emphasizes the value of deeper, more intentional friendships. This study provides valuable insights into the dynamics of investment, maintenance, and perceived closeness in friendships.

Furthermore, the strong relationship between investment and closeness suggests that fostering deeper connections in friendships requires effort and active engagement. Future research should continue to explore the complex nature of perceived closeness and the different ways in which individuals maintain and invest in their friendships.

Works Cited

- Allen, Lisa Farinelli, et al. "Emotional Investment: An Exploration of Young Adult Friends' Emotional Experience and Expression Using an Investment Model Framework." *Journal of Social and Personal Relationships*, vol. 29, no. 2, 2, Mar. 2012, pp. 206–27. *DOI.org (Crossref)*, <https://doi.org/10.1177/0265407511431054>.
- Allgood, Sarah F. "The intimate friendship scale: factors and association with drinking patterns among college aged friends." (2009).
- Aukett, Richard, Jane Ritchie, and Kathryn Mill. "Gender differences in friendship patterns." *Sex roles* 19 (1988): 57-66.
- Ben-Ari, Adital, and Yoav Lavee. "Dyadic Closeness in Marriage: From the inside Story to a Conceptual Model." *Journal of Social and Personal Relationships*, vol. 24, no. 5, Oct. 2007, pp. 627–44. *DOI.org (Crossref)*, <https://doi.org/10.1177/0265407507081451>.
- Branje, S., et al. "You're my best friend: Friendship commitment and stability in adolescence." *Personal Relationships* 14 (2007): 443-459.
- Cuddeback, John A., and John G. Trapani. "Truth and Friendship: The Importance of the Conversation of Friends." *Truth Matters: Essays in Honor of Jacques Maritain* (2004): 26-33.
- Feeney, Judith A. "Issues of Closeness and Distance in Dating Relationships: Effects of Sex and Attachment Style." *Journal of Social and Personal Relationships*, vol. 16, no. 5, Oct. 1999, pp. 571–90. *DOI.org (Crossref)*, <https://doi.org/10.1177/0265407599165002>.
- Fehr, Beverley. *Friendship Processes*. SAGE Publications, Inc., 1996. *DOI.org (Crossref)*, <https://doi.org/10.4135/9781483327440>.
- Fehr, Beverley. "Stability and commitment in friendships." *Handbook of interpersonal commitment and relationship stability*. Boston, MA: Springer US, 1999. 259-280.
- Field-Springer, Kimberly, and Kathryn Stephens. "'She Gets Me': Forming a Friendship from a Place of Vulnerability." *Health Communication*, vol. 32, no. 3, Mar. 2017, pp. 386–88. *DOI.org (Crossref)*, <https://doi.org/10.1080/10410236.2016.1138387>.
- Gächter, Simon, et al. "Measuring the Closeness of Relationships: A Comprehensive Evaluation of the 'Inclusion of the Other in the Self' Scale." *PLOS ONE*, edited by MariaPaz Espinosa, vol. 10, no. 6, 6, June 2015, p. e0129478. *DOI.org (Crossref)*, <https://doi.org/10.1371/journal.pone.0129478>.
- Gompers, Paul A., et al. "The Cost of Friendship." *Journal of Financial Economics*, vol. 119, no. 3, 3, Mar. 2016, pp. 626–44. *DOI.org (Crossref)*, <https://doi.org/10.1016/j.jfineco.2016.01.013>.
- Goodfriend, Wind, and Christopher R. Agnew. "Sunken Costs and Desired Plans: Examining Different Types of Investments in Close Relationships." *Personality and Social Psychology Bulletin*, vol. 34, no. 12, 12, Dec. 2008, pp. 1639–52. *DOI.org (Crossref)*, <https://doi.org/10.1177/0146167208323743>.

- Hackenbracht, Joy, and Karen Gasper. "Feeling More and Feeling Close: Affect Intensity Influences Judgments of Interpersonal Closeness." *Social Cognition*, vol. 31, no. 1, 1, Feb. 2013, pp. 94–105. *DOI.org (Crossref)*, <https://doi.org/10.1521/soco.2013.31.1.94>.
- Hays, Robert B. "Friendship." (1988).
- Hudson, Nathan W., and R. Chris Fraley. "Adult Attachment and Perceptions of Closeness." *Personal Relationships*, vol. 24, no. 1, 1, Mar. 2017, pp. 17–26. *DOI.org (Crossref)*, <https://doi.org/10.1111/pere.12166>.
- Huić, Aleksandra, and Ivana Smolčić. "Friendship maintenance strategies and quality of friendship." *IARR 2015 Mini-conference "Self-regulation and Close Relationships"*. 2015.
- Johnson, Amy Janan. "Examining the Maintenance of Friendships: Are There Differences between Geographically Close and Long-distance Friends?" *Communication Quarterly*, vol. 49, no. 4, 4, Sept. 2001, pp. 424–35. *DOI.org (Crossref)*, <https://doi.org/10.1080/01463370109385639>.
- Kelley, Harold H., et al. "Analyzing close relationships." *Close relationships* 20 (1983): 67.
- Leibowitz, Uri D. "What Is Friendship?" *Disputatio*, vol. 10, no. 49, 49, Nov. 2018, pp. 97–117. *DOI.org (Crossref)*, <https://doi.org/10.2478/disp-2018-0008>.
- Maslow, Abraham Harold. *Motivation and Personality: Motivation And Personality: Unlocking Your Inner Drive and Understanding Human Behavior by AH Maslow*. Prabhat Prakashan, 2023.
- Oswald, Debra L., et al. "Friendship Maintenance: An Analysis of Individual and Dyad Behaviors." *Journal of Social and Clinical Psychology*, vol. 23, no. 3, 3, June 2004, pp. 413–41. *DOI.org (Crossref)*, <https://doi.org/10.1521/jscp.23.3.413.35460>.
- Ozbay, Fatih, et al. "Social Support and Resilience to Stress: From Neurobiology to Clinical Practice." *Psychiatry (Edgmont (Pa.: Township))*, vol. 4, no. 5, May 2007, pp. 35–40.
- Rubin, Zick, and Stephen Shenker. "Friendship, Proximity, and Self-Disclosure." *Journal of Personality*, vol. 46, no. 1, Mar. 1978, pp. 1–22. *DOI.org (Crossref)*, <https://doi.org/10.1111/j.1467-6494.1978.tb00599.x>.
- Rusbult, Caryl E. "Commitment and Satisfaction in Romantic Associations: A Test of the Investment Model." *Journal of Experimental Social Psychology*, vol. 16, no. 2, Mar. 1980, pp. 172–86. *DOI.org (Crossref)*, [https://doi.org/10.1016/0022-1031\(80\)90007-4](https://doi.org/10.1016/0022-1031(80)90007-4).
- Rusbult, Caryl E., et al. "The Investment Model Scale: Measuring Commitment Level, Satisfaction Level, Quality of Alternatives, and Investment Size." *Personal Relationships*, vol. 5, no. 4, 4, Dec. 1998, pp. 357–87. *DOI.org (Crossref)*, <https://doi.org/10.1111/j.1475-6811.1998.tb00177.x>.
- Sprecher, Susan. "How Men and Women Expect to Feel and Behave in Response to Inequity in Close Relationships." *Social Psychology Quarterly*, vol. 55, no. 1, Mar. 1992, p. 57. *DOI.org (Crossref)*, <https://doi.org/10.2307/2786686>.
- Tehrani, Hossein Dabiriyani, and Sara Yamini. "Personality Traits and Conflict Resolution Styles: A Meta-Analysis." *Personality and Individual Differences*, vol. 157, Apr. 2020, p. 109794. *DOI.org (Crossref)*, <https://doi.org/10.1016/j.paid.2019.109794>.

Williams, Keelah E. G., et al. "Sex Differences in Friendship Preferences." *Evolution and Human Behavior*, vol. 43, no. 1, Jan. 2022, pp. 44–52. DOI.org (Crossref), <https://doi.org/10.1016/j.evolhumbehav.2021.09.003>.

Wood, Julia T., and Christopher C. Inman. "In a Different Mode: Masculine Styles of Communicating Closeness." *Journal of Applied Communication Research*, vol. 21, no. 3, Aug. 1993, pp. 279–95. DOI.org (Crossref), <https://doi.org/10.1080/00909889309365372>.

Young, Timothy L. "Differentiating stable and unstable friendships: commitment, investment, and reward." (1990).

TABLE 1
Correlation Table

Table 1. Correlations for closeness1, closeness2, investment, and maintenance

		closeness1	closeness2	investment	maintenance
closeness 1	Spearman's Rho		0.316	.533**	0.325
	p-value		0.078	0.002	0.069
closeness 2	Spearman's Rho	0.316		0.170	.483**
	p-value	0.078		0.351	0.005
investment	Spearman's Rho	.533**	0.170		.379*
	p-value	0.002	0.351		0.032
maintenance	Spearman's Rho	0.325	.483**	.379*	
	p-value	0.069	0.005	0.032	

** . Correlation is significant at the 0.01 level (2-tailed).
* . Correlation is significant at the 0.05 level (2-tailed).

TABLE 2
Gender Difference in Investment

Table 2. t-test results for investment comparing men and women

Group Statistics							
	sex	N	Mean	Std. Deviation	t	df	p-value
investment	Male	7	3.87	0.58	1.61	29	0.059
	Female	24	4.16	0.37			

Narratives of Control: How Greek Mythology Reinforces Patriarchal Ideals through Female Archetypes By Suhana Malli

Introduction

Mythology, likely written by men, served as a way to explain the world and establish the “truths” of its time, with every detail reflecting the values and beliefs of the society that created these stories. It is because of this tendency to weave values into storytelling that we can now read mythology and tie those themes to the beliefs that made up the lives of the ancient Greeks. In ancient Greece, where the patriarchy was deeply ingrained, fear played a significant role in shaping daily life and decisions, particularly the fear of the gods. This fear was woven into their mythology, where female characters were often depicted as either manipulative and powerful or helpless and burdensome. Females were portrayed as manipulative through characters such as monsters and witches, as powerful but emotional through goddesses and other immortal beings, and burdensome through mortals. In Greek mythology, fear is represented in many ways, notably through evil or powerless female characters who serve as cautionary symbols, reflecting the ancient Greek patriarchy and need to maintain male dominance.

Fear in the Greek language

In Ancient Greece, fear was present through their language; understanding parts of their language gives background on what fear meant and how it was represented. In the Greek language there were different ways to describe fear. *Dēos* was total terror, or fear, and similarly *deinōs* was dread or awe-inspiring. *Phōbos* was fear that resulted in running away and instilled panic and battle rout. These words are almost identical to the names of the Greek gods of fear, Deimos and Phobos. Another word was *ekplektos*, which was shock, almost like striking someone dumb. Fear can also be *krūoeis*, chilling, freezing, numbing; or *smerdalēos*, which is an adjective that means huge and terrifying sound such as the crushing of bones, the gnashing of teeth, or the thunder of Zeus. Fear could make even the mightiest warriors “jabber their teeth” and “tremble their limbs” (Alexander 1). These many ways to describe fear all show how fear was such an important tool during ancient Greece.

Similarly fear was intricately linked to the social and political structures of the time, including the patriarchal system that dominated Greek society. The powerful influence of fear in mythology and language reflects how it also diffused through the hierarchical order, shaping societal norms and personal behavior. The gods and creatures that embodied fear were not just abstract concepts but active forces that affect decision-making and social control, paralleling how fear continues in various forms today.

During these ancient times, Greece was a male dominated society, at levels much more extreme than today. There were very few powerful, intelligent, or influential women in ancient Greece. However in Greek mythology there were many supernatural and mortal female beings that played a key role in these myths. These mythological females either played the role of the weak burden, damsel in distress, or cruel/seductive characters. These character types all brought

trouble and represented the need to portray women in any way other than powerful and good. The fear of loss of power is felt in every person, but during this time that fear was primarily felt by men. The fear that women could be intelligent, powerful, good, and would bring a loss of dominance for men was represented through female characters in Greek mythology.

The purpose of Greek mythology was to reflect the values held by the ancient Greeks. Mythology was a way to try to explain scientific phenomena and the origins of the Earth. These stories were supposed to give people in this society somebody to aspire to be. Men were supposed to try to live up to the heroes in the stories, and women were supposed to be good wives, not supposed to cause problems, and were supposed to be beautiful. (Mark). In mythology the women that caused problems were portrayed as monsters, witches, and outcasts often to serve to reinforce societal norms and control behavior. This created a widespread fear which was used as a tool to keep women subject to patriarchal control.

The Greek Gods of fear: Deimos and Phobos

Deimos and Phobos were the Greek gods of fear. Deimos represented terror and dread, while Phobos represented panic, flight, and fear. The two gods were brothers and their parents were Ares and Aphrodite. Ares was the god of war, and Deimos and Phobos commonly rode with their father in his chariot into battle. They did this to spread fear into the enemy's ranks. Because Aphrodite, goddess of love, was their mother, the twins also represented the fear of loss. In art, they were usually represented as “unremarkable youths” (Atsma 1), but Phobos was sometimes seen with a lion head or leonine features.

The two gods were usually seen with warriors as a symbol to strike fear into their enemy. For example, Agamemnon had the image of a Gorgon (the name of the species of Medusa) on his shield and is described this way “There were ten circles of bronze upon it, and set about it were twenty knobs of tin, pale-shining and in the very center another knob of dark cobalt. And circled in the midst of all was the blank-eyed farce of the Gorgo with her state of horror, and Deimos was inscribed upon it, and Phobos” (Homer, *Iliad* 11. 36 ff). The fact that one of the most important characters in Greek mythology, Agamemnomom, had the head of Medusa imprinted on it emphasizes the importance of Medusa as a character. The intended effect of this was to make the general a symbol of fear to his enemies and his own soldiers. This is just one example of how Deimos and Phobos were presented in Greek mythology, and the use of the image of the Gorgon. Note that the gorgon is in fact another name for Medusa, a female. This means that the image of a female is used to put fear into the hearts of their enemies. These two gods were primarily mentioned during battle for the purpose of striking fear into the enemy in order to achieve victory. Medusa’s face was commonly associated with these gods. If the two gods of fear, not gods that were feared, but fear itself were brought up at the mention of Medusa, then clearly she too embodies fear.

Mythical creatures

There were many mythical creatures that embodied fear, two of the most important being monsters and witches. One thing to realize however is, “that all these unpleasant creatures of childhood were female” (Feldman 9). These characters were always cruel, manipulative, and in some cases seductive. There were some of the most powerful females in Greek myth, and every one was portrayed as bad with evil intentions. They were meant to warn men of the trap that women are, and are only good as long as they are controlled. In *The Odyssey* by Homer, the dead spirit of Agamemnon states when speaking to Odysseus in the underworld, “A warning even to you. Indulge a woman never, and never tell her all you know” (199 Fitzgerald). This is a clear representation of the fear of giving women any power, in this example, the power of knowledge. This reflects a broader anxiety that intelligent women could challenge male authority and disrupt the established order. Their intellect could shake the structures that keep men in control, making them a formidable force in a world that seeks to silence them.

One famous creature in Homer’s *The Odyssey* is Scylla. On Odysseus’ departure from the island Aeaea he and his crew must escape past the monster Scylla. She is described as having twelve tentacle legs, six heads upon serpent necks, triple rowed fangs, and gullets of black death. Her heads are described as “nightmares of ferocity” (Fitzgerald 212). This is obviously a fearsome creature that’s full of strength, however she is described in a disgusted tone. This may be the result of the belief that women were gross and inferior compared to that of the male body. In addition, when Hercules fights Cerberus he is described with, “three mighty heads with lion-like manes” (182 Green). This is a clear difference in the way male versus female monsters were described. Both creatures were fearsome monsters with multiple heads, yet Scylla’s appearance is from nightmares versus Cerberus’ who appears as mighty.

In *The Odyssey* she snatches six men from Odysseus’ ship, one man per head. This monster represents a choice between two evils. The whirlpool Charybdis is also described as female, and lies on the other end of the straight, which is the reason Odysseus chooses to risk the jaws of Scylla; better to lose six men than the whole crew.

Scylla is one female monster out of many on Odysseus’ journey. Before encountering Scylla he and his crew passed the Sirens. The Sirens are described as beauties that bewitch men with their beautiful music. They had wings and the tails of birds, and lured sailors to their deaths. The Sirens make an appearance in *The Odyssey* and the tale of the Argonauts. They are all female, and are a clear representation of the fear of women controlling men. In a patriarchal society, where women’s voices were drowned out, women needed to be resourceful and clever about how they would accomplish their goals. There was no fear of any physical might of women, more so what they say of what they do and how they can manipulate people.

The theme of the Sphinx’s story is similar to that of the Sirens: the Sphinx was a being whose upper half was female, her lower half was leonine, and she had the wings of an eagle. The story of the Sphinx dates back to the 8th century BCE in the Greek poet Hesiod’s work, “*Theogony*”. However, the sphinx is most commonly connected to the play written by Sophocles “*Oedipus Rex*”. She is said to have guarded the entrance to the city of Thebes, and told riddles to travelers in order to allow them to pass with their lives. Finally, a man named Oedipus

answered her riddle and being distraught she threw herself from the mountain. This is a representation of the fear of women who hold knowledge over men. The Sphinx was female and outsmarted every male to cross her path, until finally a man came and beat her. However, this ending wasn't enough to destroy the full entity of the Sphinx. The story concludes with not only her death but suicide. This is not an honorable way to die, leaving the Sphinx with a legacy of dishonor.

Another creature very similar to Sirens are Harpies. They are fierce female creatures that had the head and torso of a woman and body of vultures or eagles. They are known for reaping chaos and filth wherever they go. Part of their stench was said, “to emanate from their ever-filthy oozing under parts” (O’Malley 24). The Harpies appear in Homer’s *Odyssey* and his other poems, Hesiod’s *Theogony*, and other playwrights at the time. This is a representation of views on women’s menstrual cycle and reproductive parts of their body. It is meant to reflect the opinion that the bodily functions of women are gross and impure. All this negative representation of women was to lower the opinions of females and make them non-threatening.

Possibly the most feared creature was Medusa, the Gorgon. In Greek mythology there were three Gorgon sisters; Stheno, Euryale, and Medusa. Medusa was the only sister capable of turning men into stone with a look. In some versions of the myth, Medusa was mortal, but her two sisters were immortal. It is said that the Gorgon’s wore snakes wrapped around their waists, had large wings, vipers for hair, a beard, and sometimes fur and tusks. Medusa is pictured with a snake-girdled chiton, “snaky locks, huge distended mouth, pendant tongue and the ‘Gorgon-glaring’ eyes” (Feldman 2). What really was special about Medusa however, wasn’t her body, but her head. This concept of not being feared not for physical prowess, but for other characteristics is similar to the powers of the Sirens. It has been etched onto numerous shields, and objects to freeze the lookers with fear. The Gorgon head is famously on Agamemnon and Pallas Athena’s shields.

The story of Athena’s shield is that she and Hermes guided the Greek hero, Perseus, on his quest to bring back the Gorgon head. After Perseus succeeded, Athena placed the head on either her breastplate, or on the Aegis (the name for her shield), which is a symbol of terror and power. It’s far more powerful than Agamemnon’s shield because it was crafted by immortals, and is even owned by Zeus, but loaned to other gods. No matter whose shield her face is plastered to, the terrible eyes of Medusa are described by Homer when he states, ‘And thereon was set as a crown, Gorgo, grim of aspect, glaring terribly, and about her were Terror and Fear’ (Feldman from *Iliad* 11. 36). This quote from the *Iliad* describes Hektor gazing at Agamemnon’s shield and mentions the fear brought by Deimos and Phobos. It is important to realize what makes the Gorgon truly different from other creatures is that it’s not the body (meaning: teeth, venom, tail, or strength) of the creature that strikes fear, but instead the “gorgoneion” (Feldman 4), the head and specifically her gaze.

Witches in particular are a very important representation of the views on powerful women because unlike monsters, their appearance is identical to that of human women. Additionally their movements and ability to speak as a human does, allows for the reader to fully

see how women with knowledge and strength were perceived as evil. The witch Circe was the daughter of the sun Titan Helios, and was banished to the island Aea for turning Scylla into the monster she is forever remembered as. After being banished, sailors would come to her island and she would turn them into pigs with her magic; she did the same to most of Odysseus' crew. However, Hermes gave Odysseus the herb moly to protect him from her magic, and she is forced to turn all the pigs back into men. When the crew first see Circe her voice is described as "beguiling" and her magic "[a] drug of evil" (Fitzgerald 172). She is surrounded by wolves and lions, and works at her loom.

Although turning men into pigs is not a kind act, the fact that a woman is all alone on an island, and might be using whatever skill she has to protect herself, is not at all mentioned. The assumption that she is evil is immediately made and intentionally disregards the fact that perhaps a crew full of grown men might be threatening. She is also described as a seductress as she makes persistent attempts to lure Odysseus into her bed, but a fact that is often forgotten is that he agrees. It's as much his choice as it is hers, and Odysseus is married while Circe is not. He first says, "now it is I myself you hold, enticing into your chamber, to your dangerous bed" to take my manhood" (175 Fitzgerald). He then accepts her offer and goes on to calling her bed "flawless" and calls her "loveliest goddess" (175-180 Fitzgerald) after sleeping with her. All of a sudden she's not so evil anymore! Even after freeing the crew and being held at swordpoint by Odysseus, she hosts them for a year and guides them so that he might make it back to Ithaca. She is a woman of knowledge, power, and beauty. She actually defeats the crew at first, and is described as evil throughout this whole scene. However, only after being defeated and pleasing Odysseus in bed is she described as good. Author of the novel *Circe*, Madeline Miller states that her work was an "incarnation of male anxiety about female power". This is a clear explanation of how men feared powerful women, and only saw women as good after controlling them.

Goddesses and how they were perceived

Beyond witches and monsters, in the mythical world were the gods. There were 12 Olympian Gods who had thrones on Mt. Olympus and are the 'main Gods,' However, there are approximately 400 minor gods. The twelve Olympian gods include: Zeus, Poseidon, Hera, Demeter, Aphrodite, Athena, Artemis, Apollo, Ares, Hephaestus, Hermes and Dionysus (sometimes replaced with Hestia). Zeus is the King of the Gods, and his wife was Hera. Hera is the goddess of marriage and childbirth, Demeter is the Goddess of agriculture, Aphrodite is the goddess of love, Athena is the goddess of wisdom, Artemis is the goddess of the hunt, and Hestia is the goddess of the hearth. Although these titles convey power, their actions in mythology are often less grand (Mills 17-18).

Hera, wife of Zeus, was respected by all women and men because she was Queen of Olympus, and yet Zeus was known for his many affairs with many creatures. One would think that this would lead to many fights between the couple and result in Hera confronting her husband in some way. Unfortunately, this was not the case. Hera instead took out her anger on the lovers of Zeus, even though most of the time Zeus either raped them, or disguised himself

putting his lovers not at fault. In *Mythology: Myth and Legends of the World*, Dr. Alice Mills states, “Zeus’s wandering eye meant Hera spent much of her time plotting revenge against his lovers” (Mills 16). This jealousy and lack of reason is the opposite of what one would expect of a queen, and not just that but the queen of queens. Her role implies that she should be reasonable, confident, and respectful of everyone, but that clearly is not demonstrated. An example of this would be when Zeus had an affair with the Titaness Leto, and Hera “did not allow Leto to stay anywhere on solid earth when it was time for her to give birth” (Mills 25). As the goddess of childbirth, such behavior is considered unbecoming. These portrayals of the queen of Olympus are degrading to all women. Because she symbolizes marriage, and yet her husband is notoriously unfaithful, it appears as if marriage meant that the wife should fight to keep her husband at all costs and that he is never in the wrong. This supports the concept of a male-dominated society.

Another Olympian goddess, one of the most famous, is Aphrodite. She was the goddess of love and beauty. She was married to Hephaestus, god of blacksmith, but was known for her affairs with Ares, the god of war. As the goddess of love she was, “unrestrained, beyond reason and common sense” (Mills 56). This description accurately describes her power because love is a complicated thing; she is the mother of fear, and lover of violence and Ares god of war. An abstract concept that is important to think about is that in Greek mythology, a man could give birth, although not in the same way as women. An example would be that Zeus gave birth to Athena through his head. In other words, fear is born from a woman, but wisdom is born from man.

On the topic of Athena, she was the goddess of wisdom, strategy and warfare, and the arts. She was Zeus’s favorite daughter and was a very powerful figure. There were instances where her jealousy got the better of her, but that has happened to each of the Olympians on a number of occasions. She is known for her role as mentor in many hero’s stories; she aided Odysseus, Hercules, and Persues, but yet she was never a hero herself. This represents the idea that during this time, women were not permitted to attain full glory for their actions. Athena was immortal and powerful, and yet the actions of mortal men were the extent of her achievements. She was also one of the three virgin goddesses: Athena, Artemis, and Hestia. These three goddesses are in fact powerful figures and embody the strength of womanhood, they are wise, kind, strong, wild, and many other beautiful qualities. However, as maidens, they lack the ability to bear children, which is not the essence of womanhood, but rather a gift that it provides—the ability to create life. Why can’t powerful and beautiful women also be mothers, why are these ideas separated?

Sister of Zeus, Demeter was the goddess of agriculture and wasn’t a very popular character. However, she was a very powerful goddess, she was respected, and she was a mother. She was one of the kindest goddesses and “played little part in the battles of mortal men or the sexual entanglements of the gods of Olympus” (Mills 68). She is also “one of the goddesses of ancient Greece of whom little is told” (Mills 68). This is unusual given that agriculture and harvest are vital for the sustenance of all mortals. She is one of the few goddesses who do not

follow the main categories of evil and manipulative or weak and petty. However, there is one case when she is described as being vengeful, but really her actions are justified and labeled negatively because they oppose the will of Zeus. Her daughter Persophne was the goddess of spring, and was kidnapped and raped by Hades, god of the underworld. Demeter searched the world for her daughter, but could not find a trace anywhere on the surface. However, Zeus did see the crime, and fearful of causing Hades' anger, withheld the information. Demeter lets all the crops and creatures slowly die in her vengeance, until Zeus sends Hermes to inform Demeter of the location of her daughter. Additionally, there is no sympathy and instead Demeter is described as a problem-causer, although it was in fact Hades who spontaneously decided to kidnap and marry her daughter. This is clearly seen when Dr. Alice Mills writes, "At last Zeus decided that Demeter's grief must be healed and her daughter given back, or else all that lived on Earth would perish and the gods would receive no more sweet-smelling sacrifice to enjoy" (Mills 69). Not only was Zeus's decision not prompted by altruism, but it was also motivated by the desire to maintain worship. By ensuring the fertility of the land and the survival of crops, Zeus aimed to sustain human populations. This was crucial because human survival and prosperity were directly linked to their ability to offer sacrifices and maintain rituals in his honor. In essence, by securing the agricultural bounty that supported human life, Zeus was also securing the annoying devotion and worship that cemented his divine authority and influence.

Ultimately, an agreement was reached, allowing Persephone to spend part of the year with her mother and the other part with her husband. This is where the Greeks believed the seasons came from, and why half of the year is cold and infertile, because of Demeter's grief. Although Demeter's actions are harsh, they are justified in the fact that her daughter is stolen, and the information is withheld from her. Demeter's power is immense and is against Zeus's will, which is exactly the opposite of what a male-dominated society would want to hear. As a result her actions are called vengeful and irrational.

A lesser known goddess, but very influential nonetheless was Eris. She was the goddess of strife and the ultimate cause of the Trojan War. She is also said to have birthed Hunger, Pain, and Oblivion. During a wedding between Peleus and Thetis she dropped a golden apple with the inscription 'For the fairest' causing Hera, Athena, and Aphrodite to fight amongst it, each claiming it was meant for her. This presents women as highly envious and lacking in self-control. This portrayal also uses three of the most important goddesses, which puts all goddesses in a bad light. Moreover, the fact that the Trojan War, described as 'the greatest war among men in Greek myth' (Mills 170), resulted from the actions of multiple females suggests an irrationality and emotional impulsiveness associated with the female sex.

In examining the portrayal of goddesses in Greek mythology, it becomes evident that their representations highlight the theme of male superiority. While goddesses like Hera and Aphrodite display power and influence, their actions often emphasize the patriarchal biases of their mythological context. Hera's vengeful responses and Aphrodite's irrational love highlight how female deities are frequently depicted through a lens of jealousy and emotional instability. Athena, despite her wisdom and strength, remains a supporting figure rather than a hero,

reflecting the limitations placed on women's roles. Demeter's story, while showing her immense power and maternal devotion, is framed by a male-dominated perspective that undermines her justified actions. Eris's role in sparking the Trojan War further exemplifies how mythological narratives can preserve stereotypes about female irrationality and envy. Collectively, these portrayals offer a view of how ancient Greek society viewed women, revealing a complex relationship between power, gender, and societal expectations.

Mortal women

Mortal women were primarily seen as burdens and ultimately brought the downfall of men. More mystical beings such as the gods, monsters, and witches held power. Mortal women had a much smaller influence than these beings and usually their stories ended in tragedy. This might symbolize the belief that women usually caused problems and were hazards to themselves and society if not controlled.

One of the most iconic characters was Pandora, today a popular eponym. The story of Pandora is that she was created by the gods as the first woman for the purpose of punishing the titan Prometheus. He created humans, and gave them fire, which was against Zeus' order. As a result the beautiful Pandora was made from the gods together, and sent as a gift to be Prometheus' wife. In addition they gave the couple a jar (a common misconception is that it's a box) and ordered them never to open it. However, Hermes had made Pandora with a heart of curiosity, and so she opened the lid and allowed all the evils of man to escape. The only thing that was left partially caught on the lid was hope, or better translated from Greek to English as anticipation. This way humans would bear the hardships of life in hope of better in the future. This representation of a woman being curious and unable to control herself is common to that of Eve, the first woman in the Bible. She too is tempted and eats the forbidden fruit in the garden of evil, dooming mankind. In both scenarios the woman is created for the purpose of male companionship as well. This implies that during ancient Greece, the belief was that women were weaker than men, and that their purpose was solely to aid men. An example of this is how the woman character rarely fights and toils for her own purpose. Penelope denies the suitors, but because she was loyal to her husband; this is obviously a good reason but there is no instance when a woman stands for what she wants for herself, and then attains it.

Penelope was the wife of Odysseus and mother of Telemachus. After her husband fought in the Trojan War, on his way back he encountered many obstacles that extended his journey by ten years. Through it all many suitors storm Penelope's castle demanding that she pick a new husband to be king. Although she refuses all of them and cleverly drives them away, she is usually depicted as frail and weepy in wait for her husband's return; "When she was upstairs with the servants she wept for Odysseus" (Rouse 19). While it is natural to grieve, she had a son who needed her and a kingdom to maintain. She could have driven the suitors away herself, she and her son had control of the guards, but instead she is depicted as only the waiting wife.

While Penelope represents the ideal of patience and fidelity, Circe embodies a different, yet equally compelling facet of female power in Greek mythology. Circe was a powerful witch,

renowned for her magical abilities, and her story connects with that of her siblings, who were also prominent figures in the mythological landscape. Her sister's name was Pasiphae who was the Queen of Crete, wife of King Minos, and mother of the minotaur. Their brother's name was Aeetes, King of Colchis and keeper of the Golden Fleece. Aeetes had a daughter skilled in witchcraft, Medea. Jason and his famous crew of heroes, The Argonauts, journey to Medea's home in hopes of acquiring the Golden Fleece. Medea performs magic to aid Jason in his quest, but only after he promises to marry her. She does perform evil acts that allow her to escape with her husband, but he later abandoned her. Although she does cruelly acquire her revenge, both Jason and Medea made a deal that he broke. This is similar to Medea's cousin Adriane, first daughter of Pasiphae.

Pasiphae also practiced witchcraft, although her daughters were entirely mortal. Her son however was the Minotaur, half bull-half man. This was the result of being cursed by Posideon because her husband, King Minos of Crete, broke his promise to sacrifice the Creten bull to the god. His wife was punished by being overridden with desire for the bull, which resulted in her hybrid son. Although it was meant to punish Minos' honor, it society and psychologically hurt his wife more than it hurt him. Once again leaving a woman at the mercy and will of men.

This unfortunate trend runs in the story of the women in the bloodline, for Adriane is also left by a man who promises to have her. Theseus, prince of Athens, goes to Crete as a sacrifice to the Minotaur, but with the help of Adriane, slays the creature and leaves for Athens. On the way they stop at an island, and abandon Adriane there. Luckily Dionysis finds her and makes her his bride, but the action of leaving and then just picking up women just adds fuel to the idea that the patriarchy was unjust. Even worse than this, Thesus later marries one of Adriane's sisters, Phaedra.

Helen of Troy is one of the most famous women from Greek myth, and is portrayed as the cause of the Trojan War, even though it was in fact the goddess Eris who started the events that would cascade into the "kidnapping" of Helen. Helen was the queen of Sparta and wife of Menelaus. She was kidnapped as a child and forced to marry Theseus, but was returned to Sparta and grew up without the title of married. She is known for being the most beautiful mortal women in Greek mythology, so much so that a deal was forced to be made so that her suitors would not quarrel amongst themselves. The deal was that Helen would choose who she wished to marry, and the remaining suitors would come to Helen's aid whenever she was in need of it. This is why almost all of Greece went to war with Troy when Paris, prince of Troy, took Helen to his kingdom. Some versions claim that Helen was kidnapped, others say that she ran away to Egypt and her phantom was in Troy, but the most common story was that Helen fell in love with Paris and went with him willingly. She was the cause of one of the most famous wars in Greek history and this theme can be interpreted as women being the cause of conflict because of unstable character. The war that Helen supposedly brought on can be interpreted as women, specifically beautiful women, being dangers to men and unloyal.

The previous mortal women mentioned were the cause for the downfall of the men around them, and were portrayed as emotionally unstable. Two examples of women that were

powerful, but not considered weak were Atalanta and the hunters of Artemis. Atalanta was cast out by her father as a baby because he wished for a boy. However, she was raised by hunters and became an incredible athlete and hunter. She is known for being the only female to be a part of the Argo and participated in the Calydonian boar hunt. However, similar to the hunters of Artemis she despised the prospect of marriage. Later her father accepted her and requested that she marry. Atalanta agreed, so long as her suitor beat her in a foot race. She was known for her speed, and would kill anyone who lost the race. A man named Hippomenes was aided by Aphrodite and raced against Atalanta. He dropped a total of three golden apples throughout the race so that she would stop to pick them up, giving him enough time to beat her. A repeating theme in Greek mythology was that strong and capable women refused the love of a man. This is because during ancient Greek times there was a very strict interpretation of masculinity and femininity. However, Atlanta and the hunters of Artemis participate in many practices that were considered to be rites of passage to young men in ancient Greece. Atalanta also expresses themes of femininity through marriage and the footrace. Foot races and dances were common practices that young ladies would perform before womanhood. The contrast between being desirable and powerful was not common for females during this time, making Atalanta a very rare type of character.

Goddesses, witches, monsters and women

To summarize, the most common character types that women were perceived as during ancient Greece were manipulating and evil or emotionally unstable and burdensome. Goddesses were usually a mix of these two because they were powerful because of their immortality which often puts them in the perception of being manipulative and vengeful, but when compared to males (immortal and in some cases mortal) they are over emotional and irrational which was considered a female associated tendency. Monsters were evil, savage, and dirty without any sort of awe associated with their might. Witches were power hungry, seductive threats. Mortal women were preyed upon constantly and were put into the position of being a burden or the cause of conflict.

Amazons

Although the story of Atalanta is a rare one, one group of women embodied the impossible theme of her story. This group of women were known as the Amazons. Ironically they were, “the mythic archenemies of the ancient Greeks” (Mayor). The Amazons were warrior women that lived northeast of Ancient Greece. The Amazons were first mentioned by Homer and the most popular account of them in mythology is tied to the ninth labor of Hercules. Hercules was the demigod son of Zeus, and Hera charged Hercules with ten nearly impossible tasks to complete. The ninth one was to steal the girdle from the Amazon queen, Hippolyta. Hercules succeeds and ultimately dishonors the Amazons. The fact that this warrior group was so skilled increases the success of Hercules’ theft, but discredits the Amazons.

Another amazonian warrior, Penthesilia, went to Troy to aid the Trojans during the Trojan War and fought Achilles to the death. Right before she died, Achilles lifted her helmet off her head to gaze upon her face and fell in love with her. This simple act is what made the Amazons so different from other warrior women in ancient Greece, the Amazons accepted men, to an extent. The Amazons were phenomenal warriors in all aspects, but also interacted with men which was a very uncommon combination. Additionally, although they are deeply rooted in myth, there is also physical evidence that suggests that a group of nomadic women warriors existed in Central Asia and other areas of Asia (Brooklyn Museum: Amazon) The Amazons go against the archetypes set for females in Greek mythology. They accept the company of men, but are completely independent of their society. Although this group of female warriors goes against the popular portrayals of women, the fact that the Amazon group was greatly disliked by the Greeks, reinforces the idea that women were silenced to empower men.

Modern literature about women from myth

In the early 2000s a movement of retelling Greek mythology from the perspective of female characters sparked to life. This movement reflects a cultural shift towards recognizing underrepresented voices and perspectives. This shift comes from ongoing discussions about general equality; this allows for a sense of progress in understanding history and identity. These novels gave these female characters a voice and the ability to tell their own stories without being a secondary, subservient character on the side. In many primary sources regarding Greek mythology, such as Homer's works and ancient Greek historians and poets such as Hesiod and Herodotus, the female characters' full story isn't told, unlike the male characters. However, some female characters are vital to the story, yet their past and future aren't elaborated. Madeline Miller has written two outstanding novels, *Circe* and *The Song of Achilles*. When asked what drew her to write about Circe she responded, "We learn nothing of why she's transforming men, or how she came to power. I wanted to imagine her whole story, and I also wanted to push back against the one-dimensional way she's been treated in popular culture" (Miller). She clearly wanted to take the *Odyssey* and give as much effort that went into telling Odysseus' story, into Circe's story.

The novel *Circe* tells the story of the witch from Homer's *Odyssey* without excluding any of the events that take place in the traditional epic. *The Song of Achilles* tells the story of the events leading up to the Trojan War through the perspective of Achilles' best friend and possible lover, Patroclus. This point of view allows the reader to see the depth of emotion these characters have, who were portrayed as only warriors in the origin tellings. Other novels include *Electra* by Jennifer Saint which focuses on the perspectives of Clytemnestra, Electra, and Cassandra. Clytemnestra was the wife of Agamemnon and was famous because she murdered her husband after he returned from the war because he sacrificed their eldest daughter Iphigenia to Artemis so that the ships could sail to Troy. This novel makes emotional connections between Clymenstra and her youngest daughter Electra. These modern retellings are different from traditional stories because they offer many perspectives which allow for more themes to be shared. Rather than the

traditional themes that may have been seen only through the perspective of the main characters. Cassandra was a princess from Troy who was cursed by Apollo to have the gift of prophecy, but no one would believe her. She was cursed because she refused to sleep with the god, and so this novel gives the opinion of a Trojan citizen rather than a Greek. Saint also wrote *Adriane*, about a Cretan princess who is tricked by the hero Theseus. There are many other novels that give the perspectives of men and women, monsters and witches, princesses and peasants. There are also novels about other types of mythology; an example of this would be *The Witches Heart* by Genevieve Gornichec which is about a witch from Norse Mythology.

Before this new trend, mythology and mythological characters were a part of popular culture in subtle ways. Athena, goddess of wisdom set an archetype which many characters in other franchises follow. One such example is Hermione Granger from *Harry Potter*, who is an intelligent and strategic girl. The British girl group, Bananarama, in the 1980s released a song called "Venus" which includes themes of love, similar to the Roman goddess Venus, who is the counterpart of the Greek goddess Aphrodite (Mills 18).

In addition to archetypes and novels, there are many films that include characters and stories from Greek mythology. The Amazons are a very popular story today, and are included in many movies and tv shows, such as *Wonder Women*, *Xena: Warrior Princess* and *Clash of the Titans* (1984 and 2017). These reflect a growing interest with powerful, independent female figures. The 2017 *Wonder Woman* film, in particular, was extremely liked and highlighted a story where women are not only there to aid the male hero, but are the heroes themselves.

Conclusion

Greek mythology serves as a reflection of the ancient Greek patriarchal society, particularly in its representation of fear through female characters. The Greek language itself illuminates the nature of fear as well as the gods of fear, Deimos and Phobos. These deities' embodiment of fear shows how essential fear was as a tool during ancient Greece, hence the many portrayals and diction revolving around it. Mythical creatures such as Scylla, the Sirens, and Medusa exemplify how female figures were portrayed as symbols of dread and caution, reflecting deep-seated cultural anxieties surrounding female power. These monstrous figures, while having physical strength, symbolize the fear of women who defy societal norms and disrupt the established order, threatening the male way of life. The portrayal of goddesses and mortal women reinforces the patriarchal perspective of females being emotional and burdensome therefore reinforcing the authority of men at the time. Although there were a few exceptions to these portrayals, they weren't enough to alter the dominant lens in which women were viewed through, which characterized them as inferior.

Overall, Greek mythology reflects the cultural norms and expectations of the time which both females and males looked to as models. These myths influenced the lives and identities of both women and men in ancient Greece. In recent years, the movement to retell Greek mythology from a female perspective has emerged, reflecting a cultural shift toward recognizing

underrepresented voices and experiences. These novels aim to dig deeper into the motivations and backgrounds of all categories of female characters from Greek mythology.

Works Cited

- Alexander, Caroline . “The Dread Gorgon | Caroline Alexander.” *Lapham’s Quarterly*,
www.laphamsquarterly.org/fear/dread-gorgon.
- Atsma, Aaron J. “DEIMOS & PHOBOS - Greek Gods of Fear, Panic & Terror (Roman Metus,
Pavor).” *Www.theoi.com*, 2017, www.theoi.com/Daimon/Deimos.html.
- Barringer, Judith M. *Atalanta as Model*. University of California Press, 1996.
- Brooklyn Museum . “Brooklyn Museum: Amazon.” *Www.brooklynmuseum.org*,
www.brooklynmuseum.org/eascfa/dinner_party/place_settings/amazon.
- Cartwright, Mark. “Pandora.” *World History Encyclopedia*, World History Encyclopedia, 27 July
2015, www.worldhistory.org/Pandora/.
- Feldman, Thalia. “Gorgo and the Origins of Fear.” *Arion: A Journal of Humanities and the
Classics*, vol. 4, no. 3, 1965, pp. 484–94. *JSTOR*, <http://www.jstor.org/stable/20162978>.
Accessed 2 Sept. 2024.
- Gornichec, Genevieve. *The Witch’s Heart*. New York, Ace, 2021.
- Greenberg, Mike, and PhD. “Phobos and Deimos - the Complete Guide to the Greek Spirits.”
Mythology Source, 7 July 2020, mythologysource.com/phobos-deimos-greek-spirits/.
- Homer, and Robert Fitzgerald. *The Odyssey*. London, David Campbell Publishers, 2000.
- Homer. *The Odyssey*. Translated by W.H.D Rouse, University Press, 1946.
- Mactaggart, Catriona. “How Greek Mythology Is Being Rewritten through a Feminist Lens.”
The Independent, 1 June 2021,
www.independent.co.uk/how-greek-mythology-is-being-rewritten-through-a-feminist-lens/.
- Mark, Joshua J., and Paul Kriwaczek. “Mythology.” *World History Encyclopedia*, 31 October
2018, <https://www.worldhistory.org/mythology>. Accessed 15 September 2024.
- Miller, Madeline. *Circe*. Bloomsbury publishing, 10 Apr. 2018.
---. *The Song of Achilles*. London, Bloomsbury, 2012.
- Mayor, Adrienne . “Princeton University Press: The Amazons: Lives and Legends of Warrior
Women across the Ancient World.” *Princeton.edu*, 2024,
press.princeton.edu/books/hardcover/9780691147208/the-amazons. Accessed 18 Aug. 2024.
- McGreevy, Nora. “Men Have Feared Women for Millennia. Just Look at the Monsters of Greek
Mythology.” *Smithsonian Magazine*, 31 Mar. 2021,
www.smithsonianmag.com/arts-culture/meet-female-monsters-greek-mythology-medusa-
sphinx-180977364/.
- Mills, Alice. *Mythology*. Hodder & Stoughton, 1 Jan. 2003.
- Nagy, Gregory . “The Subjectivity of Fear as Reflected in Ancient Greek Wording.” *The Center
for Hellenic Studies*,
chs.harvard.edu/curated-article/gregory-nagy-the-subjectivity-of-fear-as-reflected-in-ancient-greek-wording/.
- O’Malley, Tam. *Mythological Animals*. eBook Partnership, 20 May 2021.
- Roger Lancelyn Green, et al. *The Tales of the Greek Heroes*. London, Puffin, 2015.
- Saint, Jennifer. *Ariadne*. Flatiron Books, 4 May 2021. ---. *Elektra*. Flatiron Books, 3 May 2022.
- Saint, Jennifer. “Ariadne.” *World History Encyclopedia*, 1 Apr. 2022,
www.worldhistory.org/Ariadne/.

Evolving Identities: The Impact of Climate Change on the San and Nubian Peoples

By Olumisimi Owolabi

Introduction

Like a bird's beak evolving to suit its environment, human cultures can transform in response to a changing climate. This idea frames the complex relationship between climate and culture, especially for African indigenous communities whose practices and sustenance are deeply tied to their natural environments. This study focuses on two such communities: the San people and the Egyptian Nubians. The San, a group renowned for their historic hunter-gatherer lifestyle – though only a tiny fraction still actively practice it – and the Nubians, with their long-standing agricultural practices along the Nile, both face unique challenges as they confront the warming climate.

This paper aims to answer the central question: How has climate change transformed cultural and social dynamics in Southern and Northern Africa? Climate change has significantly reshaped these dynamics within the San and Egyptian Nubian communities, particularly concerning culture, occupation, and gender roles. Despite their distinct geographical and historical contexts, both groups have adapted to climate challenges by altering their traditional practices and adopting new strategies to sustain their way of life. This paper seeks to demonstrate the importance of understanding how these factors intersect and shape strategy in response to climate change.

This research uses case studies on the San people and the Egyptian Nubians to explore the impacts of climate change on cultural and social dynamics, supported by various sources, including journals, articles, and interviews with experts. Furthermore, the study aims to assess how climate change may continue to transform human life and what strategies different communities adopt to combat these challenges. By evaluating case studies, we can learn how climate change disrupts and reshapes cultural identities and resilience strategies, guiding other vulnerable communities worldwide.

Methodology

This study utilizes a combination of scholarly sources and qualitative interviews to lay a solid foundation for better understanding the implications of climate change for both the San people and the Egyptian Nubians. These findings may also apply to other communities that face similar challenges with climate change, particularly in terms of their traditional ways of life and cultural practices. I conducted virtual interviews with academic experts, including anthropologist Professor Robert Hitchcock, who has worked with the San since 1975, and historian Professor James McCann, who has researched African communities' adaptations to environmental challenges. Throughout the interviews, I asked the experts a series of questions regarding ways climate change has affected the socio-cultural fabric of the respective communities. In addition to the interviews, this research draws upon reputable journals and articles from databases such as Google Scholar and JSTOR. I prioritized using reliable and academically rigorous sources,

which led to a continuous refinement of key terms to accurately encompass the research's complexity.

Literature Review

Climate change significantly impacts the water resources in Africa, profoundly affecting cultural, social, and occupational communities like the Egyptian Nubians and the San. In her book *Climate Change in Africa* (2009), Camilla Toulmin highlights water's significance in Africa, emphasizing its role in livelihood, subsistence, spiritual significance, and economic activities like irrigation and industrial processes. Toulmin documents how climate change intensifies the water cycle, leading to severe floods and droughts. According to climate models, global temperatures will increase by 3.2 to 3.6 degrees Celsius in the next 100 years. Northern and Southern Africa are expected to become drier, and East Africa is expected to become wetter (Toulmin 38). For the Egyptian Nubians, and many other groups, the Nile River has historically been a vital source of Egyptian Nubian culture. The Carnegie Museum of Natural History's article "Egypt and the Nile" emphasizes the Nile's role in providing essential water and silt for many aspects of Egyptian life, such as agriculture, transportation, and cultural and spiritual practices (Egypt and the Nile). These sources demonstrate how water works as a fundamental component of many African cultures and how many problems stemming from global warming affect the continent culturally and impact the water quantity and quality.

Gender dynamics add another dimension to the impacts of climate change. The book *Gender and Climate Change: South Africa Case Study* (2010) by Dr. Agnes Babugura explores the gendered impacts of climate change in South Africa, noting that it is not gender-neutral. Babugura's research dives into the differences between men's and women's societal roles, demonstrating that women make up over 70% of the workplace population in the area (Babugura 16). Similarly, the *National Strategy For Mainstreaming Gender In Climate Change In Egypt* (2011) also discusses gender dynamics and climate change. It dives into factors contributing to the vulnerability of women to climate-induced hazards in Egypt, including reasons such as biological differences, social norms, and inequitable distributions of aid and resources (Aguilar et al.). This highlights gendered impacts in Egypt and how cultural and social norms intersect with climate change, directly adding to how resilience strategies are implemented in these regions.

Occupational dynamics also contribute to the implications of climate change. The South African Medical Journal's article "Climate change and occupational health: A South African perspective" by T. Kjellstrom provides a deeper dive into the impacts of climate change and discusses more specific examples, such as people in the workplace with chemicals present, agricultural workplaces, and miners. However, most of these studies primarily focus on heat and global warming, which is different from the topic of water, which is closely related to previous research on cultural aspects (Kjellstrom). Additionally, the source "Climate Change In South Africa: 21 Stunning Facts About South Africa's Climate Breakdown" by the nonprofit Greenpeace discusses the vulnerabilities of specific occupational groups in South Africa and how

climate change might affect economic well-being. For example, Greenpeace uses examples of how reduced productivity, earnings, and health issues caused by climate change, especially intense heat, all contribute to problems within the workforce, such as forestry and tourism. It even highlights the World Bank's warning that climate change could push over 100 million more people into poverty by 2030 (Igamba).

The Intergovernmental Panel on Climate Change (IPCC) report on Africa (2022) provides a comprehensive overview of the impacts of climate change across the continent. This source, authored by different climate scientists and experts from various African cultures, discusses Africa's response to the climate crisis. This source touches upon multiple aspects, including food security, water resources, and cultural heritage. It also emphasizes the importance of cultural heritage and provides examples of affected traditions and cultural sites, including rock art sites, underwater heritage sites, rituals, and Indigenous engineering systems (Trisos et al. 1393). This source presents climate change's impacts in different regions of the continent, which is a significant consideration for the study. It also discusses ways aspects of cultures, such as African indigenous languages and knowledge, can combat climate change (Trisos et al.).

The existing literature provides a solid foundation for understanding the multifaceted impacts of climate change on water resources, cultural practices, gender, and occupational dynamics; however, several gaps remain. While research highlights general effects such as water scarcity and intensified weather patterns, there is limited focus on how these changes specifically affect the cultures of different communities. Additionally, the intersection of gender dynamics is often unexplored despite evidence indicating that women are particularly vulnerable in climate change-affected areas. Occupational health studies also tend to prioritize heat and general climate issues over their impacts on water quality and daily life for workers. Most notably, there is a significant lack of research focused specifically on the San people and the Egyptian Nubians. These gaps highlight the need for targeted research to explore these nuances in more detail and understand how climate change truly affects these communities.

Section 1: The San People

The San people, indigenous to the Kalahari Desert, have experienced significant shifts in their cultural lives due to climate change. Professor Robert Hitchcock has witnessed the evolving challenges the San have faced concerning their settlement patterns, livelihood strategies, and cultural practices. According to Hitchcock, the San have traditionally led a semi-nomadic lifestyle, adapting to the arid environment of the Kalahari. However, over the last fifty years, climate change has propelled significant changes in their daily lives. Compounding these climate challenges is the historical dispossession of San lands due to Western colonialism, which has drastically shaped their way of life. Hitchcock notes, "The biggest impact of Western colonialism has been the reduction of land for the San people. For example, the San used to occupy Namibia, and now they occupy a much smaller portion. Much of the land was given to German colonial farmers and later Afrikaners from South Africa" (Hitchcock). This reduction in land and territory, further exacerbated by climate change, has pushed the San into marginal areas less

suited to their traditional subsistence practices. As temperature rises and water levels decrease, the reduced land area becomes less suitable for the San's traditional practices. This stress on natural resources makes preserving cultural practices and traditional knowledge more challenging, as these new living conditions limited their access to traditional hunting and gathering grounds (Hitchcock).

Nutrition among the San people has also worsened due to climate change. Less nutritious alternatives have replaced their traditional, diverse diets (Hitchcock). Although the government tries to offset the problems of nutritional difficulties by providing some food or other resources like cash for food, these measures still do not fully address the nutritional gap. Hitchcock spotlights the problematic nature of this dietary shift: "The food that they provide, similar to the United States with Native Americans in the 19th century, is mostly maize meal. Not very diversified, they might get a little tin of fish and some milk powder, which is problematic regarding the diet" (Hitchcock). This dietary change has heightened the need for healthcare and traditional healing practices. However, these practices have also declined due to restricted mobility, the natural deaths of significant healers, and the lack of training being passed down to the younger generation in the community (Hitchcock). As a result, the accelerating effects of climate change undermine the San's ability to cope with both nutritional and health challenges.

Rainmaking also holds historical importance among the San people, yet its practice and meaning have evolved in response to climate change (Harvey). Historically, rainmaking rituals invited rain through ceremonial dances and spiritual practices, as depicted in ancient rock art. Contemporary San communities rarely engage in these rituals themselves. Instead, neighboring groups such as the Khoikhoi or Tswana are often called upon to perform the rainmaking ceremonies. However, as environmental conditions in the Kalahari have shifted, the frequency and effectiveness of these rituals have diminished (Hitchcock).

Climate change has also significantly impacted gender dynamics among the San people, worsening existing disparities and reshaping traditional roles. Historically, San society was known for its gender-equitable structure, with both men and women playing essential roles in activities such as gathering plant foods and engaging in agricultural practices, as noted by Hitchcock. However, climate change has intensified gender inequalities within the community. As environmental conditions deteriorate, women, primarily responsible for gathering food and managing agricultural tasks, face the brunt of climate change (Hitchcock). Erratic rainfall patterns and droughts have reduced crop yields and disrupted farming practices, diminishing women's roles and livelihoods within the community. The shifts have also increased male dominance in politics and decision-making, steering away from the San's traditional gender balance. To combat these issues, efforts to promote women's empowerment, led by internal and external organizations, aim to address these disparities and elevate women's status within San society. Initiatives focus on providing resources, training, and opportunities for women to participate more actively in decision-making processes and economic activities, fostering greater gender equality (Babugara 19).

Climate change has also significantly transformed the occupational landscape within San communities. Changes in disease patterns – such as increased malaria outbreaks – are influenced by changing environmental conditions and affect workforce availability and productivity. Traditional occupations such as agriculture and pastoralism face increased uncertainty due to unpredictable weather patterns, affecting livelihoods that rely heavily on natural resources. Furthermore, climate change has impacted tourism in Southern Africa, a sector critical for showcasing San cultural heritage and promoting economic prosperity. Declines in tourism numbers, exacerbated by global events like the COVID-19 pandemic, have reduced income opportunities for San communities reliant on tourism-related activities such as cultural tours and artisan crafts. The mining industry, particularly prevalent in countries such as Namibia, Botswana, and Zimbabwe, further complicates the socioeconomic life of the San. International companies involved in operations like fracking contribute to environmental hardships like dropping water tables and habitat destruction, which disproportionately affect indigenous populations like the San. These changes in the local environment and the global economy illustrate the difficulties of maintaining traditional occupations while adapting to new opportunities (Hitchcock). For example, some San communities have been granted hunting permits to preserve their traditional methods, like using spears and poison arrows, under Namibia's strict anti-poaching laws. Furthermore, eco-tourism initiatives such as Nhoma Camp and N/a'an ku sê Lodge have provided alternative income streams, allowing the San to engage in cultural tourism and conservation efforts. However, these efforts are constantly threatened by climate change, overdevelopment, and economic instability (Styles).

The San have discussed their deep connection to their land and environment with Hitchcock and are keen observers of change over time. They frequently discuss noticeable declines in rainfall, reduced plant growth, and diminished wildlife populations. These changes have significantly impacted their way of life, making it more difficult for them to sustain themselves through hunting, gathering, and agriculture (Hitchcock).

In response to these challenges, the San have had to resist the adverse effects of climate change. They have actively sought representatives in national climate change agencies and organizations, advocating for their rights and interests in decision-making processes. The forums help ensure that their unique perspectives and knowledge of the environment are integrated into policies and strategies addressing climate change. Moreover, the San also work to have their voices heard on global platforms. They engage with international organizations such as the European Union, U.S. Agency for International Development, and the World Bank to access funding for climate change initiatives in their communities. These efforts aim to leverage international partnerships to implement sustainable solutions that align with the San's cultural values and priorities. The resistance efforts of the San also extend beyond advocacy to include strategies at the community level, like preserving traditional knowledge, to mitigate and adapt to climate change. This means they leverage their ancestral knowledge of local ecosystems to adapt agricultural practices, hunting techniques, and water management strategies to changing climates. The San people also work on community-based adaptation projects. The San have

initiated projects that promote sustainable land use, water preservation, and renewable energy solutions within their communities. Lastly, the San educate. They educate the younger generation about climate change and the importance of environmental stewardship so they can continue their traditional practices (Hitchcock).

Additionally, the San have remained active in climate change and cultural identity discussions. Their advocacy efforts address historical injustices in securing land rights, with climate change adding complexity to these issues. For instance, the San are involved in promoting green energy initiatives, or renewable resources, to reduce the impact of environmental challenges. Green energy is designed to address environmental sustainability and promote economic development. Additionally, the San incorporated drought-resistant plant varieties and relocated to more viable areas to enhance their resilience. These efforts tackle immediate survival needs and aim to preserve and adapt their cultural heritages. By participating in discussions like these and taking action, the San continue to advocate for their cultural rights (Murphy).

Section 2: The Egyptian Nubians

The Egyptian Nubians, an ethnic group along the Nile River, have also experienced significant socio-cultural changes due to climate change, including shifts in traditional livelihoods, gender roles, and occupational dynamics. Professor James McCann, through his research, has observed how the Egyptian Nubians have adapted differently to their surroundings and how migration and trade have shaped these communities. The Egyptian Nubians have traditionally depended heavily on the Nile for their livelihood, relying on its water for agriculture, fishing, and transportation, now threatened by climate change. The Nile River is crucial to the survival and cultural practices of the Egyptian Nubians. The Nile is the primary water source for Cairo (McCann). The Blue Nile, in particular, is vital for subsistence, providing most of the water during the Ethiopian rainy season (Egypt and the Nile). According to McCann, the river holds symbolic weight for the Egyptian Nubians, evident in rituals where animals are sacrificed at the river's edge to establish a connection with a deity, demonstrating the water's significance as a symbol of holiness (McCann).

The Nubian region has relied on the Nile, with local populations having historically managed water for crop cultivation and adapted their political systems to accommodate the Nile's seasonal and long-term fluctuations. However, climate change has worsened challenges related to water scarcity, including increasing pressure on water resources needed for agriculture, trade, and socioeconomic development in Egypt. The strategic significance of the Nile has positioned Egypt as a focal point for European powers, most notably the British and the French. In the early 19th century, local leadership coexisted with European oversight, with Mohammad Ali expanding Egyptian influence into Nubia and East Africa while asserting control over Sudan and water resources. The construction of dams, like the Aswan Dam, aimed to manage water more efficiently but raised concerns about control and implications for downstream countries. With climate models predicting a significant decline in water availability in certain parts of

Africa and events surrounding the Aswan Dam highlighting Egypt's vulnerability, there are increasing concerns about Egypt's water scarcity (McCann).

In Egypt, climate impacts are amplifying pre-existing gender inequalities. According to the National Strategy for Maintaining Gender in Climate Change in Egypt, women, particularly in vulnerable communities, face disproportionate hardships due to climate change because women are more likely to be employed in sectors directly affected by climate change, such as agriculture and water management (Aguilar et al. 14). These sectors are heavily sensitive to climate fluctuations, and women's involvement in them makes them particularly vulnerable to climate-induced disruptions, including drought or flooding. Moreover, patriarchal norms and expectations that confine women to underpaid labor often limit women's access to crucial resources and decision-making power, compounding their exposure to these environmental risks and reducing their ability to adapt to these changing conditions (Aguilar et al. 14).

Furthermore, traditional gender roles and limited access to resources exacerbate women's vulnerability. Women traditionally play a crucial role in agricultural production, managing family diets, and food provision. Agriculture contributes over 13% of Egypt's GDP and provides over 30% of employment. However, many women in agricultural sectors work without wages or receive low pay compared to their male counterparts. They face illiteracy and lack land ownership, limiting their ability to respond to and manage climate-related challenges effectively. The National Strategy recommends that women become more involved in decision-making and have improved working conditions, including fair wages, health benefits, and protection from hazardous work environments (Aguilar et al. 15, 27).

Egyptian Nubian's occupational dynamics are also affected by climate change. As McCann notes, modern Egypt's economy and way of life have heavily depended on the Nile's water availability. However, water scarcity has increasingly affected Egypt's traditional irrigation systems. Changes in the Nile's flow and patterns have led to higher evaporation rates, diminishing the water resources needed for agriculture. This is compounded by rising temperatures and altered precipitation patterns, further impacting crop yields (McCann). According to an article titled "Climate Change in Egypt: Opportunities and Obstacles" by Amr Hamzawy, Mohammad Al-Mailam, and Joy Arkeh, approximately 55% of Egypt's labor force is engaged in agriculture, accounting for 11.3% of the country's GDP. Recently, climate change altered the Nile's flow, jeopardizing these resources. These environmental changes result in significant declines in agricultural productivity, with projections indicating a 12% reduction in wheat yields and a 15% decrease in maize yields by 2030. These reductions will likely worsen Egypt's already high cereal import dependency, which has escalated from 34% to 42% in the past decade (Hamzawy). Additionally, Guardian writer Ruth Michaelson notes that these changes have severe consequences for the Egyptian Nubians, whose livelihoods rely heavily on agriculture. The reduction in crop yields diminishes agricultural productivity, likely leading to fewer job opportunities. The Toshka project, which aimed to expand farmland and create new urban areas, has exacerbated these challenges. Initiated in the late 1990s and revived in 2015, the project sought to irrigate desert land and establish new agricultural and urban developments.

However, the project's failure, marked by high saline levels contaminating groundwater and slow progress, has undeveloped much of the land. The anticipated job creation from the project has not materialized, forcing many Nubians to migrate to urban areas in search of employment. This migration pressures urban job markets, which are already struggling to accommodate the influx of new workers (Michealson).

According to McCann, as climate change alters rainfall patterns and agricultural practices, the Egyptian Nubians have implemented various strategies to cope with these challenges. Traditional farming methods, which rely on predictable seasonal rains, have grown to be less effective due to the weather becoming more irregular. As a result, the Nubian farmers began adjusting their cultivation methods and crop choices. They are incorporating more drought-tolerant crops and adopting alternative irrigation methods, such as rainwater harvesting and soil moisture management, to handle the sporadic water supply better and maintain agricultural productivity. Nubians have also addressed water management by constructing small-scale irrigation systems and protecting water sources through culturally significant practices. Despite the benefits of new water infrastructure, such as dams intended to alleviate water shortages, these projects have led to displacement and loss of ancestral land. In response, the Egyptian Nubians have engaged in advocacy and negotiations to safeguard their rights and ensure that new projects consider their cultural and social needs (McCann).

Conclusion

Despite being located on opposite sides of Africa, the San people of Southern Africa and the Egyptian Nubians of Northern Africa share more similarities than differences regarding climate change's impacts on their cultural and social life. Both communities exhibit profound cultural connections to their environment, influencing their way of life. For the San, the Kalahari Desert is central to their cultural identity, while the Nile River is fundamental to the Egyptian Nubians. Additionally, displacement and land loss are common challenges faced by both communities. The San have been forced to abandon their traditional lifestyle due to climate shifts, including increased drought and water scarcity. Similarly, the Egyptian Nubians have been displaced by the construction of dams on the Nile, which has disrupted traditional flooding patterns essential for agriculture and complicated water management in the face of altered precipitation and evaporation rates. In response to these challenges, both communities have implemented adaptation strategies to mitigate the effects of climate change. The San have diversified their livelihoods by integrating small-scale agriculture and tourism into their traditional hunting and gathering practices. Meanwhile, the Nubians have adopted modern agricultural techniques to deal with the reduced water availability and changing environmental conditions affecting their farming.

However, it remains true that significant differences also exist between these communities, precisely in their historical context. Geographically, the San reside in the arid and semi-arid regions of the Kalahari, where drought and water scarcity are primary concerns. In contrast, the Egyptian Nubians live along the Nile River, where changes in water flow and

flooding patterns impact agricultural practices. These geographical contexts led to distinct occupational dynamics: the San traditionally relied on hunting and gathering, whereas the Nubians depended on agriculture. Consequently, their adaptation strategies differ, reflecting their unique environmental and economic challenges. Additionally, historical and colonial contexts further differentiate the experiences of these communities. The San have faced significant land dispossession due to colonialism in Southern Africa, impacting their traditional practices. Conversely, the Egyptian Nubians contend with the legacy of British colonialism, which involved large-scale infrastructure projects that altered the Nile River and displaced communities.

The experiences of the San people and the Egyptian Nubians illustrate the impact of climate change on culturally significant environments and traditional lifestyles. Despite their geographical and historical differences, both communities serve as examples of indigenous knowledge and cultural resilience's critical role in navigating climate challenges. The adaptation strategies employed by these groups highlight the necessity of incorporating diverse perspectives to develop effective and equitable climate solutions. Furthermore, the lingering effects of colonialism, gender disparities, and occupational dynamics exacerbate the vulnerabilities of marginalized groups, making it necessary to address historical injustices to fight against these environmental changes. Regional contexts, whether characterized by water scarcity or extreme weather events, influence communities' specific challenges and their ability to adapt. Organizations like the Intergovernmental Panel on Climate Change (IPCC) are crucial in integrating diverse viewpoints and providing comprehensive assessments to guide global policy and action. By acknowledging and addressing these complexities, we can work toward creating inclusive and sustainable solutions that support all communities in the face of climate change.

Works Cited

- Aguilar, Lorena, et al. *National Strategy for Mainstreaming Gender*. June 2011, genderclimatetracker.org/sites/default/files/Resources/2011_IUCN_Climate-Change-Gen-der-Action-Plan-Egypt.pdf. Accessed 26 Aug. 2024.
- Babugura, Agnes. *Gender and Climate Change: South Africa Case Study*. Heinrich Boll Foundation Southern Africa, 2010, www.boell.de/sites/default/files/assets/boell.de/images/download_de/ecology/south_africa.pdf.
- "Egypt and the Nile." *Carnegie Museum of Natural History*, carnegiemnh.org/egypt-and-the-nile/#:~:text=Every%20aspect%20of%20life%20in,brought%20life%20to%20the%20desert.
- Hamzawy, Amr, et al. "Climate Change in Egypt: Opportunities and Obstacles." *Carnegie Endowment for International Peace*, carnegieendowment.org/research/2023/10/climate-change-in-egypt-opportunities-and-obstacles?lang=en.
- Harvey, Chelsea, et al. "Rainmaking Experiments Boom amid Worsening Drought." *Scientific American*, 7 Feb. 2023, www.scientificamerican.com/article/rainmaking-experiments-boom-amid-worsening-drought/. Accessed 26 Aug. 2024.
- Hitchcock, Robert. Personal interview. 11 February 2024.
- Igamba, Joan. "Climate Change in South Africa: 21 Stunning Facts about South Africa's Climate Breakdown." *Greenpeace*, 30 Aug. 2023, www.greenpeace.org/africa/en/blogs/54171/climate-change-in-south-africa-21-stunning-facts-about-south-africas-climate-breakdown/. Accessed 26 Aug. 2024.
- Kjellstrom, T., et al. "Climate Change and Occupational Health: A South African Perspective." *South African Medical Journal*, vol. 104, no. 8, 2014, www.ajol.info/index.php/samj/article/view/107881.
- McCann, James. Personal interview. 6 March 2024.
- Michaelson, Ruth. "Egypt's Nubians Fight for Ancestral Land Earmarked for Mega-project." *The Guardian*, 13 Feb. 2017, carnegieendowment.org/research/2023/10/climate-change-in-egypt-opportunities-and-obstacles?lang=en. Accessed 26 Aug. 2024.
- Murphy, Devin. "San Communities Are Fighting to Protect Their Homes from Oil and Gas Exploration." *Rewild*, 11 Feb. 2022, <https://www.rewild.org/news/san-communities-are-fighting-to-protect-their-homes-from-oil-and-gas>
- Styles, Ruth. "San vs Wild: What the San People Can Teach Us about Living with Climate Change." *Ecologist*, 23 May 2012, theecologist.org/2012/may/23/san-vs-wild-what-san-people-can-teach-us-about-living-climate-change.

Toulmin, Camilla. "Introduction." Introduction. *Climate Change in Africa*, Bloomsbury Publishing, pp. 1-14,
https://dpl6hyzg28thp.cloudfront.net/media/CamillaToulmin_2009_3Water_ClimateChangeInAfrica.pdf

Trisos, C.H., et al. "Africa." *Climate Change 2022: Impacts, Adaptation, and Vulnerability*, edited by H.-O. Pörtner et al., Cambridge University Press, 2022, pp. 1285–1455.
https://www.ipcc.ch/report/ar6/wg2/downloads/report/IPCC_AR6_WGII_Chapter09.pdfdoi:10.1

A Comparative Analysis of the Economic Relief Efforts from the 1918 Spanish Flu and COVID-19 Pandemics By Hana Sultan, Alena Friedrich

Abstract

This project conducts a comparative analysis of the post-pandemic economic relief efforts from the 1918 Spanish Flu and COVID-19. The paper begins by contextualizing the impacts of the Spanish Flu with the overlapping effects of WWI, leading to difficulty in making definitive economic conclusions. It then discusses the importance of recognizing the differing mortality groups of the Spanish Flu and COVID-19 pandemics. After giving this background, the data on non-pharmaceutical intervention methods (NPIs) used during the 1918 pandemic in various U.S. cities is introduced, and the economic consequences of the pandemic are explored using financial data from the same period. Next, conclusions from the recent COVID-19 pandemic are explored, particularly ones related to the extensive use of social distancing. With data from the two time periods, parallels are drawn, and the economic consequences of implementing NPIs are discussed to inform future pandemic responses.

Introduction

In light of the recent COVID-19 pandemic, extensive intervention policies have been enacted by the U.S. government to curb economic shock. The involvement of economic outcomes with pandemic planning is a relatively new area of study and has involved exploring various pandemic responses. As COVID-19 spread in 2020, non-pharmaceutical interventions (NPIs) became globally enforced policies through measures such as mandatory social distancing, and four years later, their effects on economic outcomes can now be observed. Using data from the 1918 Spanish Flu pandemic (a global health crisis with a similar government response), we can see that the start and duration of NPI implementation were variable among cities during that time period, which allows for the evaluation of the effectiveness of such measures. Including research from both pandemics provides greater validity to our conclusions, as the health crises and enforced control measures were comparable in nature. This paper aims to evaluate the use of NPIs and data from the two time periods to inform future pandemic responses for the best economic outcomes.

Background

Approaching the end of World War I in 1918, the Spanish Flu began and spread rapidly through the United States. With record-high mortality rates, the disease quickly evolved into a global pandemic—one which is often compared to the recent COVID-19 pandemic. Infections with the influenza virus caused flu symptoms, such as high fevers and fatigue, quickly leading to respiratory failure due to the lack of available treatments. With an estimated 675,000 deaths in the U.S., approximately 0.5% of the nation's population at the time, this pandemic was considered one of the worst in history (Noymer & Garenne, 2000).

In 2019, COVID-19 arose as the next global pandemic, generating widespread fear due to its unknown nature and comparisons to previous epidemics. With over 1 million deaths, the coronavirus has had a significant impact but today accounts for less than a 0.02% loss in population globally (Patterson et. al 2021).

Using these two time periods, parallels can be drawn related to the implementation of NPIs and their economic effects. Although similar, it must also be recognized that the mortality rate during the Spanish Flu had a much more substantial effect on labor supply than COVID-19, as the largest mortality group was working-age men during the Spanish Flu compared to elderly men in the recent pandemic. Additionally, due to the overshadowing of WWI, there was little media coverage of the Spanish Flu, and economic involvement was nearly negligible. Numerous studies have been conducted to quantify the economic shocks caused by such a pandemic and to draw parallels between the Spanish Flu and the COVID-19 pandemic (Wheelock 2020). Recognizing the limitations in the data, we can use records from various cities that utilized NPIs during the 1918 pandemic and the effects of social distancing in the past years to explore the use of these interventions in improving economic outcomes.

Literature Review

In this paper, we aim to investigate the relationship between NPI implementation and post-pandemic economic outcomes. This topic has gained popularity with the recent outbreak of COVID-19, and emerging research on the economic effects of pandemics, namely the 1918 Spanish Flu and COVID-19 pandemics, aids our analysis. Basco, Domenech, and Roses (2021) use evidence from the Spanish Flu to determine pandemic-related economic shocks on demand and non-essential employment sectors. Lison, Banholzer, Sharma, et. al (2023) and Fair, Karatayev, Anand, et. al (2022) conducted assessments on the effectiveness of NPI measures during the COVID-19 pandemic, directing our analysis. Modeling techniques constructed by Acemoglu, Chernozhukov, Werning, et al. (2020) and Bognanni, Hanley, Kolliner, et. al (2020) also inform our final policy suggestions. Ultimately, we rely most substantially on a paper by Bootsma and Ferguson (2007) that models a correlation between NPI intervention during the 1918 Spanish Flu pandemic and differing mortality rates in 16 major U.S. cities- the same which we conduct our analysis upon, except observing economic variables rather than mortality rates.

We use the papers highlighted above to gather evidence on the economic influence of pandemics and response measures, but we also contribute to the analysis of NPI implementation, specifically on improving economic outcomes. Using our own data collection, we use statistical methods to provide greater comprehension of the economic aftermath of NPI measures, extending economic exploration of specific pandemic controls. Our analysis stimulates further discussion on comprehensive pandemic-response measures to avoid economic downfall following future pandemics.

Data

In order to evaluate the correlation between earlier NPI intervention start dates and economic effects, we constructed graphs to visualize the relationships between these variables of interest. Using manufacturing records pulled from the United States Census Bureau's 1920 Census, we examined data collected from a variety of cities from 16 states. We looked at three time periods with a five-year gap before and after the Spanish Flu to evaluate a correlation between earlier/longer NPI implementation and change in wages and cost of materials. The census data is collected through a variety of sources, including surveys, businesses, and federal and state governments. It serves to provide an accurate representation of national statistics and inform future decision-making.

Methodology

To examine the economic effects of the Spanish flu and COVID-19 pandemics, we used a quantitative research approach. By utilizing records from a five-year period before and after the Spanish Flu pandemic in 16 major U.S. cities, we investigated a correlation between the implementation of NPIs during the pandemic and significant economic effects. Specifically, we looked at the duration of NPI interventions and start dates as independent variables and differences in change of wages and cost of materials (COM) from 1909-1914 and 1914-1919 as outcome variables. By incorporating the differences-in-differences statistical technique, or looking at the change in outcomes between a pre-pandemic and post-pandemic period, we may establish stronger claims about the causal relationship between variables.

Using the collected data, we constructed a variety of multiple linear regression models to examine relationships between the variables of interest to relate our findings to prior research. Additionally, we ran similar regressions for the cost of living to conclude the more holistic economic effects of the pandemic. Our focus was on the first intervention, as variables pertaining to the start of a second intervention, including total excess mortality, were not significant and therefore excluded from the analysis. Using the collected data, we constructed a variety of linear regression models to examine relationships between the variables of interest. We found results that were statistically significant at or below a 0.10 p-value threshold and discussed the implications of our findings in the existing conversation around pandemic policy responses in the field of economics.

Results

Intervention Start vs. Differences in % Wage/COM Change

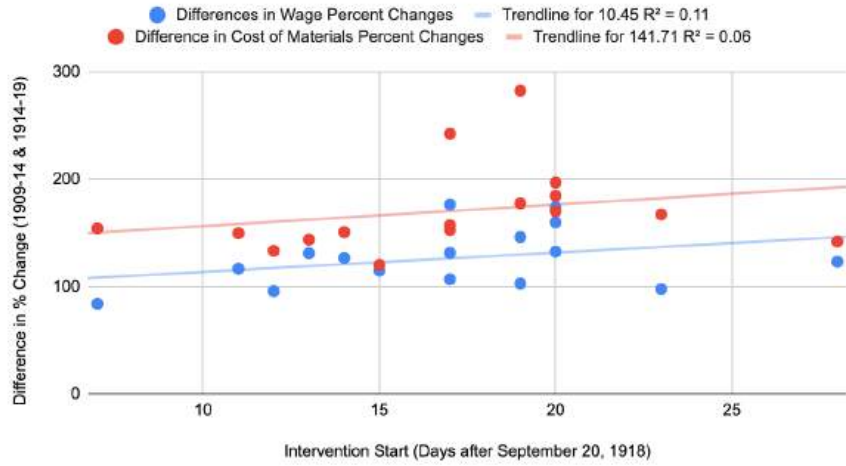


Figure 1. This figure depicts a scatterplot showing a weak positive relationship between the start of the first intervention in 16 U.S. cities (measured in days after September 20, 1918) and the differences in percent change in wages and cost of materials from two time periods (1909-1914 and 1914-1918).

Intervention Duration vs. Differences in % Wage/COM Change

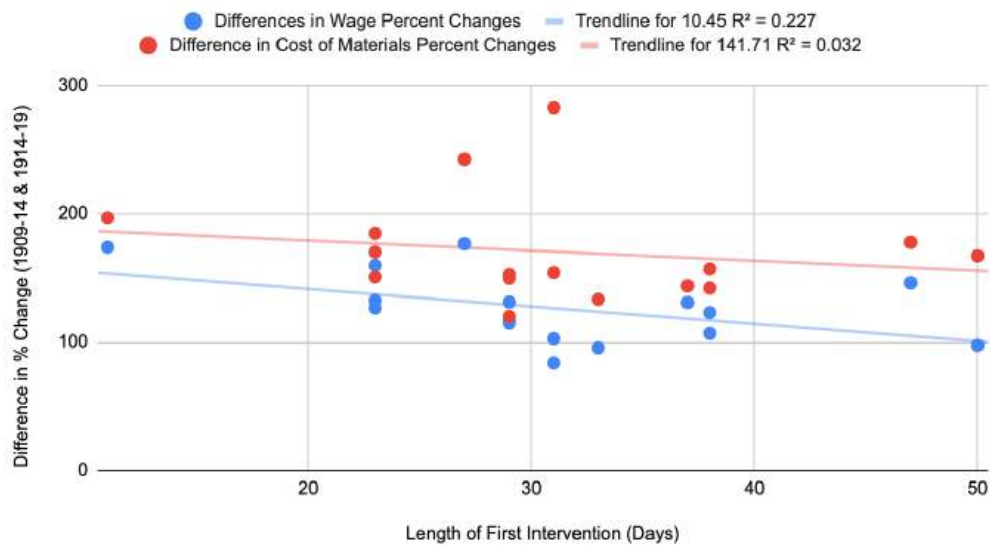


Figure 2. This figure depicts a scatterplot illustrating a weak negative relationship between the length of the first intervention (in days) in 16 U.S. cities and the differences in percent change in wages and cost of materials between two time periods (1909-1914 and 1914-1918).

Difference in Wage Percentage Change Between 1909-1914 and 1914-1919				Difference in Cost of Materials Percentage Change Between 1909-1914 and 1914-1919			
	Estimate	Standard Error	P-value		Estimate	Standard Error	P-value
Intercept	135.9273	25.9402	0.00016***	Intercept	161.4770	48.1259	0.00517***
Days Past Sept 20th of Start of Intervention	2.2830	1.1728	0.07352*	Days Past Sept 20th of Start of Intervention	2.3161	2.1759	0.30651
Total Length of First Intervention (Days)	-1.5558	0.6183	0.02579**	Total Length of First Intervention (Days)	-0.9789	1.1471	0.40891
Adjusted R: 0.3096 F Statistic: 4.364 P-Value: 0.03548 Number of Observations: 16				Adjusted R: -0.02722 F Statistic: 0.8013 P-Value: 0.4697 Number of Observations: 16			
Significance Codes: '***' p≤0.01; '**' p≤0.05; '*' p≤0.1; ' ' p>0.1				Significance Codes: '***' p≤0.01; '**' p≤0.05; '*' p≤0.1; ' ' p>0.1			

Figure 3. The tables present the multiple regression analysis results for the differences in wage percent change (left panel) and the cost of materials percent change (right panel) across two time periods (1909-1914 and 1914-1919).

After constructing linear regression models of the variables of interest, we found significant results associated with earlier start dates of NPI interventions, length of initial interventions, and difference in percent change of wages. In Figure 1, a scatterplot of intervention start dates and changes in wages and cost of materials is shown. It plots the days since September 20, 1918 along the x-axis and measures the difference in percent change in the two outcome variables between 1909-1914 and 1914-1919 along the y-axis. A weak positive correlation is seen between intervention start dates and both variables, but the low values of R^2 indicate a low explanatory power of these relationships—particularly for the cost of materials. In Figure 2, we then investigated the relationship between the first intervention length and the difference in percent wage and COM change between 1909-1914 and 1914-1919. The correlation found between intervention duration and cost of materials is notably weaker than the relationship between intervention duration and wage change, suggesting that wage impacts are the only factors notably affected.

Having found a slight correlation between both intervention start dates, length of the first intervention, and difference in change of wages and cost of materials, a multiple regression analysis was then conducted to determine the joint economic impact of both the time and duration of NPIs, as seen in Figure 3. In Figure 3, the start date and duration of the first interventions were used as income variables, and the difference in percent wage/COM change from 1909-1914 and 1914-1919 were used as separate outcome variables.

In the right table, the results of the multiple regression on the relationship between the start of the first intervention and duration and the difference in COM change between the two 5-year periods are shown. The high p-values of both input variables indicate a statistically insignificant relationship, and the negative R^2 leads us to conclude the insignificance of the explanatory variables. We were not able to find data indicative of a strong relationship between the two input variables and the difference in percent cost of material change, leading us to then explore the difference in wage change as the next outcome variable.

As depicted in the left table, the multiple regression model produced an R^2 value of 0.3096 and a joint p-value of 0.03548 for the difference in percent wage change. These results indicate a strong correlation between earlier intervention start and longer duration with decreased change in wage difference compared to pre-pandemic economic paths.

Furthermore, with a p-value of about 0.04, we can be confident that the correlation is not due to chance. In the table, we use the statistics to conclude that for every additional intervention start date past September 20th, wage growth between 1914-1919 compared to wage growth between 1909-1914 was 2.28 percentage points higher, ultimately revealing that delayed intervention meant significant change in wage growth. This result is statistically significant at the 10th percent level, meaning there is a 90% confidence in our yielded results. In addition, the data tells us that for every additional day of a city's first NPI intervention, wage growth between 1914-1919 compared to wage growth between 1909-1914 was 1.56 percentage points lower, meaning longer interventions meant slower wage growth. This result is statistically significant at the 5 percent level.

The multiple linear models regressing the same independent variables on the cost of materials (which we treated as a proxy for the cost of living) finding insignificant results suggest that we can reliably claim that impacts were felt from wage changes without significant cost changes. Moreover, we highlight using data pertaining to the first intervention, as the second intervention results were not statistically significant. The total length of interventions was not as impactful as the length of the first intervention, leading to the recognition of the first intervention as the vital period for economic influence.

Discussion

Post-1918 pandemic, U.S. cities observed a notable recession, partly caused by many cities mandating against large gatherings and observing a temporary lockdown. A sharp decline in economic activity resulted from reduced wages, a decline in consumption, and a loss of non-essential businesses. As written by the Wall Street Journal in October 1918, this impact was quantifiable by an approximately 50% decrease in production, loss of trade, and impairment of efficiency during the pandemic (Correia & Luck 2020). The occurrence of pandemics represents both a supply and demand shock. Specifically, high mortality and reduced hours result in a decline in supply, while a decrease in consumption and an increase in unemployment reduces demand (Basco et al. 2021). The pandemic's inherent contribution to a national recession

supports that earlier and longer implementations of NPIs are not conducive to swift economic recovery.

After conducting various regressions, we determined a significant relationship between earlier and longer NPI interventions on differences in percent wage change. The data's statistical significance has led us to conclude that cities during the 1918 Spanish Flu pandemic that implemented their first intervention both earlier and for a longer duration observed worse economic outcomes post-pandemic. A larger difference in percent wage change corresponds to larger wage change post-pandemic compared to pre-pandemic paths and indicates a more negative economic outcome. Having found insignificant results associated with a difference in percent COM change, we deduced that the course of cost of living did not change substantially in correlation with the input variables (having used the cost of materials as a proxy for cost of living). Therefore, the significant difference in wage growth signifies a disproportional change with COM and an inability for consumers to meet the new prices demanded by businesses. Significant wage changes in contrast to overall consistent COM suggest that NPI implementation consequentially left consumers worse off.

During the COVID-19 pandemic, extensive implementation of NPIs was seen through government-mandated social distancing orders. Beginning in 2020, state and federal governments imposed "6-feet" distancing rules, as well as shutdowns of non-essential businesses and large crowds. These measures, enacted quickly and aggressively, aimed to reduce the spread of the virus, but a sharp decline in economic activity followed.

Our findings on the 1918 Spanish Flu pandemic were analogous to current research on the economic effects of NPI interventions during the COVID-19 pandemic. A paper published by ClinicoEconomics and Outcomes Research wrote that NPI measures implemented during the pandemic, such as widespread lockdowns, had unintended repercussions, leading to a decline in economic growth (Richards et al. 2022). In a study of NPI implementation in European and Central Asian countries, NPIs were found to have led to "a decline of about 10% in economic activity", suggesting NPIs cause severe economic consequences (Demirguc-Kunt 2021).

Furthermore, the supply and demand shock generated by the pandemic not only affected both suppliers and consumers alike but generated greater consequences due to socioeconomic inequality. Lower-income groups are more sensitive to the restrictions imposed by NPIs, resulting in disproportionate rates of unemployment and decreased mobility- the Council of Economic Advisors to the U.S. Government predicted an average reduction of disposable income of more than 10% in Q1 of 2020 for those in the low-income bracket (CEA 2020).

The overall data is in agreement with our findings, with a pandemic-related drop in employment of 13.6 percent between March and April 2020, the height of the COVID-19 pandemic (U.S. Bureau of Labor Statistics 2022). We use this research to draw parallels between the economic consequences of harsh NPI implementation during both the Spanish Flu and COVID-19 pandemics. Strict lockdown measures, enacted quickly and aggressively, can be tied to efforts to reduce a pandemic's predicted economic shocks but ultimately come at the detriment of workers, especially those in low-income communities.

In spite of economic consequences, the health benefits of NPI implementation should be taken into account. The extensive use of NPIs during the COVID-19 pandemic contributed to reduced mortality rates, feasibly contributing to lower economic costs from loss of labor supply. An article by the Proceedings of the National Academy of Sciences of the U.S.A. highlights the effects of NPIs on mortality rates, claiming a significant correlation between peak mortality and timing of interventions (PNAS 2007). A correlation analysis of 16 U.S. cities (the same used in our models) during the Spanish Flu Pandemic supports this claim, discovering that “attack rates (and thus mortality) can be reduced by 30-40% through transitory controls”. Likewise, during the COVID-19 pandemic, the predicted peak mortality was said to have been significantly higher should lockdown measures not have been implemented, and mandates were able to significantly reduce the number of cases and deaths (Fair et al. 2022). Given the pronounced health benefit of NPI implementation, we may consider the impact on labor supply and productivity and, thus, a reduced demand shock during a pandemic.

Taking into consideration both the short-term economic costs and reduction in mortality rates, we look to inform future policy by exploring the economic outcomes of policy implemented in response to the post-COVID recession. The economic shock resulting from the Spanish Flu is speculated to have been more significant, but an inadequate economic response was executed due to a lack of economic involvement in pandemic planning before the late 1900s (Boianovsky 2021).

In light of the repercussions of harsh NPI implementation, the U.S. economy contracted by 31.4% in Q2 of 2020, and government policy response was crucial to recovery. Substantial relief and recovery measures in 2020 and 2021 fueled an economic recovery “that made the deepest recession post-World War II era also the shortest” (Center on Budget and Policy Priorities 2022). Six laws under both the Biden and Trump administrations were considered key pandemic relief programs: The Coronavirus Preparedness and Response Supplemental Appropriations Act, The Families First Coronavirus Response Act (2020 & 2021), The Coronavirus Aid, Relief, and Economic Security Act (CARES), Paycheck Protection Program and Health Care Enhancement Act, and The American Rescue Plan of 2021. The Federal Reserve took additional monetary policy steps, cutting its target interest rate and initiating large-scale asset purchases to control inflation. The enactment of an estimated \$5 trillion in pandemic funding led to a recovery that was stronger and faster than initial projections: a 5.8% fall in GDP in 2020 and a 1.1% fall in 2023 compared to the predicted 12% and 9% fall, respectively (CBPP 2022). A sharp increase in employment and a decrease in inflation was also observed in 2023, an outcome of the implemented policies.

Recognizing the success of stimulus programs, we use this data to inform future policy recommendations. While the pandemic response contributed to the nation’s economic recovery, we advise thorough control measures earlier in the pandemic to prevent a severe recession. Using the data and research collected, we understand the implications of harsh, widespread NPI measures, and therefore propose targeted NPIs along with stimulus packages to ensure the best economic outcomes.

Interactions between economic and epidemiological decisions are necessary to target policies that mitigate economic downfall while also saving lives. A working paper by the National Bureau of Economic Research suggests an “epi-econ” model, integrating a model of the economy during an epidemic to create a response strategy. The paper shows that “targeted policies combined with measures that reduce interactions between groups and increase testing and isolation of the infected can minimize both economic losses and deaths” (Demirguc-Kunt 2021). Policymakers should adjust to different social groups and predicted economic outcomes to pass measures that target the most vulnerable while providing the safety net of stimulus packages. In doing so, NPI implementation may result in a lower economic cost, and therefore reduce government spending in response to a post-pandemic economic crisis.

Due to the cyclical nature of pandemics, pandemic planning and preparedness is essential to maintain economic stability, and we recommend further research and development of epi-econ models. Targeted policy also requires observational data on selective NPIs on vulnerable groups to allow for evidence-based implementation before the next pandemic, which we suggest in addition to the refinement of the Econ-SIR Model by the Federal Reserve (Bognanni 2020). With the integration of targeted policy and stimulus packages, we wish to progress future pandemic responses to mitigate economic recession.

Limitations

Our results face the same limitations as many similar studies in the field, specifically regarding difficulty with claiming definite causality and potential interference of confounding variables. With the end of World War I coinciding with the 1918 Spanish Flu pandemic, we recognize the possible influence of confounding variables on our data. However, our difference-in-differences approach allows us to make stronger claims on causality, absolute claims on a changing economy cannot be established. When drawing parallels between the 1918 Spanish Flu and COVID-19 pandemics, we must also acknowledge the differing peak mortality groups to understand differences in supply and demand shocks. Lastly, the focus of this paper is on the short-term economic outcomes of NPI implementation and does not address possible differences in short-term versus long-term outcomes. More research is necessary on the economic influences of pandemics, and our paper serves to contribute to the conversation on the short-run effects.

Conclusion

The recent COVID-19 pandemic has prompted discussion on pandemics’ economic consequences, leading to deeper research on economic outcomes of past pandemics and effectiveness of control measures. In our paper, we evaluate the use of NPIs during the 1918 Spanish Flu and COVID-19 pandemics with the purpose of drawing parallels between their economic effects in order to inform future pandemic planning. Using the differences-in-differences statistical technique to strengthen our causal correlations, we conducted various regression analyses on our variables of interest. We looked specifically at the

duration and length of the first NPI interventions in 16 major U.S. cities as our input variables, and cost of materials (which we treated as a proxy for cost of living) and wages as outcome variables. We discovered a statistically significant relationship between differences in wage change and first intervention length and duration, leading us to conclude harsh NPI implementation has substantial economic consequences. This was in agreement with the research we conducted on the effects of NPIs during the COVID-19 pandemic and the ensuing recession. That said, we also explored NPIs' success in reducing mortality rates, highlighting the health-economy trade-off of strict NPI implementation.

We then discussed data on economic recovery measures implemented post-COVID pandemic to explore the benefits of stimulus packages in mitigating the economic crisis. In combination, we brought attention to epi-econ modeling of pandemic outcomes in order to construct recommendations for future pandemic policy.

In an age of greater economic global interdependence than ever before, having strong control measures to reduce recession during future pandemics is crucial in order to maintain economic stability. The economic consequences of health crises are now widely recognized, and we hope our findings can influence future pandemic planning for improved economic outcomes.

Works Cited

- Acemoglu, D., Chernozhukov, V., Werning, I., Whinston, M., Doyle, J., Ellison, G., Fenichel, E., Johnson, S., Mongey, S., Shimer, R., We, A., & Yi, S. (2020). *NBER WORKING PAPER SERIES A MULTI-RISK SIR MODEL WITH OPTIMALLY TARGETED LOCKDOWN*. https://www.nber.org/system/files/working_papers/w27102/revisions/w27102.rev0.pdf?mod=article_inline
- and, E., Health, Board, & on, F. (2018, October 15). *The Economics and Modeling of Emerging Infectious Diseases and Biological Risks*. Nih.gov; National Academies Press (US). <https://www.ncbi.nlm.nih.gov/books/NBK534894/>
- Avery, C., Bossert, W., Clark, A., Ellison, G., & Ellison, S. F. (2020). An Economist's Guide to Epidemiology Models of Infectious Disease. *Journal of Economic Perspectives*, 34(4), 79–104. <https://doi.org/10.1257/jep.34.4.79>
- Barro, R. J. (2022). Non-pharmaceutical interventions and mortality in U.S. cities during the great influenza pandemic, 1918–1919. *Research in Economics*, 76(2). <https://doi.org/10.1016/j.rie.2022.06.001>
- Basco, S., Domènech, J., & Rosés, J. R. (2021). The redistributive effects of pandemics: Evidence on the Spanish flu. *World Development*, 141, 105389. <https://doi.org/10.1016/j.worlddev.2021.105389>
- Boianovsky, M., & Erreygers, G. (2021). How Economists Ignored the Spanish Flu Pandemic in 1918–1920. *Erasmus Journal for Philosophy and Economics*, 14(1). <https://doi.org/10.23941/ejpe.v14i1.549>
- Bootsma, M. C. J., & Ferguson, N. M. (2007). The effect of public health measures on the 1918 influenza pandemic in U.S. cities. *Proceedings of the National Academy of Sciences*, 104(18), 7588–7593. <https://doi.org/10.1073/pnas.0611071104>
- Center on Budget and Policy Priorities. (2023, April 13). *Chart Book: Tracking the Recovery From the Pandemic Recession*. Center on Budget and Policy Priorities. <https://www.cbpp.org/research/economy/tracking-the-recovery-from-the-pandemic-recession>
- Correia, S., Luck, S., & Verner, E. (2020, November 12). *Fight the Pandemic, Save the Economy: Lessons from the 1918 Flu*. The Economic Historian. <https://economic-historian.com/2020/11/1918-flu/>
- Crafts, N., & Fearon, P. (2010). Lessons from the 1930s Great Depression. *Oxford Review of Economic Policy*, 26(3), 285–317. <http://www.jstor.org/stable/43664566>
- Demirgüç-Kunt, A., Lokshin, M., & Torre, I. (2021). The sooner, the better: The economic impact of non-pharmaceutical interventions during the early stage of the COVID-19 pandemic. *Economics of Transition and Institutional Change*, 29(4), 551–573. <https://doi.org/10.1111/ecot.12284>
- Deng, Y., Meng, L., & Zhou, Y. (2022). The effectiveness and costs of nonpharmaceutical interventions for COVID-19 containment: A border discontinuous

- difference-in-difference approach. *China Economic Review*, 75, 101849.
<https://doi.org/10.1016/j.chieco.2022.101849>
- Fair, Kathryn, et al. “Estimating COVID-19 Cases and Deaths Prevented by Non-Pharmaceutical Interventions, and the Impact of Individual Actions: A Retrospective Model-Based Analysis.” *Epidemics*, vol. 39, June 2022. Elsevier,
<https://doi.org/10.1016/j.epidem.2022.100557>.
- Haw, D. J., Morgenstern, C., Forchini, G., Johnson, R., Doohan, P., Smith, P. C., & Hauck, K. D. (2022). Data needs for integrated economic-epidemiological models of pandemic mitigation policies. *Epidemics*, 41, 100644–100644.
<https://doi.org/10.1016/j.epidem.2022.100644>
- International Monetary Fund. (2021, July 2). *Policy Responses to COVID19*. International Monetary Fund.
<https://www.imf.org/en/Topics/imf-and-covid19/Policy-Responses-to-COVID-19#U>
- Lison, Adrian , et al. “Effectiveness Assessment of Non-Pharmaceutical Interventions: Lessons Learned from the COVID-19 Pandemic.” *The Lancet Public Health*, vol. 8, no. 4, Apr. 2023, [www.thelancet.com/journals/lanpub/article/PIIS2468-2667\(23\)00046-4/fulltext](http://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(23)00046-4/fulltext).
- Morens, D. M., & Fauci, A. S. (2007). The 1918 Influenza Pandemic: Insights for the 21st Century. *The Journal of Infectious Diseases*, 195(7), 1018–1028.
<http://www.jstor.org/stable/30086034>
- Noymer, A., & Garenne, M. (2000). The 1918 Influenza Epidemic’s Effects on Sex Differentials in Mortality in the United States. *Population and Development Review*, 26(3), 565–581.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2740912/#:~:text=The%20estimated%20population%20of%20the>
- Pangallo, M., Aleta, A., del Rio-Chanona, R. M., Pichler, A., Martín-Corral, D., Chinazzi, M., Lafond, F., Ajelli, M., Moro, E., Moreno, Y., Vespignani, A., & Farmer, J. D. (2023). The unequal effects of the health–economy trade-off during the COVID-19 pandemic. *Nature Human Behaviour*, 8, 1–12. <https://doi.org/10.1038/s41562-023-01747-x>
- Richards, F., Kodjamanova, P., Chen, X., Li, N., Atanasov, P., Bennetts, L., Patterson, B. J., Yektashenas, B., Mesa-Frias, M., Tronczynski, K., Buyukkaramikli, N., & Houry, A. C. E. (2022). Economic Burden of COVID-19: A Systematic Review. *ClinicoEconomics and Outcomes Research*, 14, 293–307. <https://doi.org/10.2147/CEOR.S338225>
- Segarra-Blasco, A., Teruel, M., & Cattaruzzo, S. (2021). The economic reaction to non-pharmaceutical interventions during Covid-19. *Economic Analysis and Policy*, 72, 592–608. <https://doi.org/10.1016/j.eap.2021.10.006>
- Siklos, P. L. (2022). Did the great influenza of 1918–1920 trigger a reversal of the first era of globalization?*. *International Economics and Economic Policy*, 19.
<https://doi.org/10.1007/s10368-021-00526-1>
- The Council of Economic Advisers*. (2020).
<https://trumpwhitehouse.archives.gov/wp-content/uploads/2020/08/Evaluating-the-Effects-of-the-Economic-Response-to-COVID-19.pdf>

- The Six Laws that Funded Pandemic Relief Programs | Pandemic Oversight.* (n.d).
Www.pandemicoversight.gov.
<https://www.pandemicoversight.gov/about-us/pandemic-relief-program-laws>
- Tomes, N. (2010). “Destroyer and Teacher”: Managing the Masses During the 1918–1919 Influenza Pandemic. *Public Health Reports*, 125(Suppl 3), 48–62.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2862334/>
- Wheelock, D. (2020, June 1). *What Can We Learn from the Spanish Flu Pandemic of 1918-19 for COVID-19?* | *St. Louis Fed.* Wwww.stlouisfed.org.
<https://www.stlouisfed.org/on-the-economy/2020/june/what-can-we-learn-from-the-spanish-flu-pandemic-of-1918-for-covid-19>

The Impact of Social Media on Teen Anxiety By Naiesha Shah

Abstract

This study explores the link between Instagram usage and anxiety levels among Indian teenagers. With growing concerns about social media's influence on mental health, especially in adolescents who are highly sensitive to peer interactions, this research addresses a gap in the literature by focusing on Indian teens, a group often underrepresented in studies of social media impact. The primary objective was to investigate whether the frequency and type of Instagram use were associated with increased anxiety levels in teens. The study aimed to explore how Instagram engagement, including the time spent on the platform and the nature of content consumed, could influence psychological well-being. A sample of 30 participants aged 13-18 was recruited through convenience sampling. Participants completed a non-standardized survey on Instagram usage and a standardized anxiety assessment using the DASS-Y test. Data was collected via Google Forms, and participants' anxiety levels were analyzed in relation to their social media habits. The findings revealed no statistically significant correlation between Instagram usage and anxiety. However, a notable result was the strong correlation between self-reported anxiety levels and actual DASS-Y scores, indicating that participants had a high level of self-awareness about their mental health. In conclusion, while Instagram use did not directly correlate with increased anxiety, the study underscores the complexity of social media's impact on mental health. Future research should consider larger sample sizes and focus on different types of content engagement to gain a more comprehensive understanding of the relationship between social media and anxiety in teens.

Introduction

Teen anxiety, particularly in the context of social media, has emerged as a pressing global issue impacting various aspects of teenagers' lives, from social interactions and academic achievements to mental well-being. Previous research suggests that adolescents experience significant neurological changes during their teen years, particularly in brain regions related to social interactions, making them more sensitive to acceptance and rejection, especially in the context of social media interactions⁷² Moreover, adolescents' biological sensitivity to social cues and peer acceptance/rejection plays a crucial role in their anxiety levels, with platforms encouraging passive consumption, significantly affecting self-esteem and perception of social approval⁷³. However, previous research has not thoroughly examined the precise relationships between the kinds of accounts teenagers follow on social media and the ensuing anxiety levels, particularly when considering Indian teens.

⁷² <https://www.nature.com/articles/s41467-018-03126-x.pdf>

⁷³ <https://www.sciencedirect.com/science/article/pii/S0747563216304682>: and <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9966679/>

Previously, research found that excessive or dysfunctional use of mobile phones, often termed Problematic Mobile Phone Use (PU), is associated with various psychological variables. For instance,⁷⁴ I discovered that chronic stress, low emotional stability, female young age depression, and extraversion are linked with PU such as excessive texting. Similarly, excessive use of Instagram (hours spent on the platform) was associated with increased levels of anxiousness for boys and poor self-declarations of body-image for girls . In this respect, excessive Instagram use has been shown to frequently amplify peer pressure and social expectation, putting tremendous pressure on the teens to conform to unrealistic standards set either by their peers or by influencers. Other threats associated with cyberbullying on Instagram include greater exposure of victims to such psychological problems as depression, social anxiety, and suicidal thoughts. Accordingly, repeated use of Facebook has been linked with psychological stress among adolescents, where passive use, namely news feed browsing, leads to envy and low well-being.⁷⁵ Furthermore, the societal challenge of digital addiction, as highlighted by research on Problematic Mobile Phone Use , suggests that its overuse may result in psychological problems like chronic stress and depression . Such findings imply understanding not only psychological but also social factors of severe use of Instagram or mobile phones to arrive definitely at providing some interventions to foster healthier technology use habits among teens.

When taking into account the standpoints regarding the study of the psychological aspects of problematic mobile phone use behavior,⁷⁶ found that frequent use of Facebook directly related to the tendency towards more psychological distress, with an indirect effect via increased communication overload and lower self-esteem. Such findings highlight the need for context-appropriate individualized approaches to target such variables in promoting mindful and balanced use of technology among youth.

The current research gap regarding Instagram and anxiety among adolescents, particularly with reference to specific demographics such as Indian adolescents needs to be addressed. Past studies fail to consider the cultural contexts and specific role social media plays in influencing mental health outcomes with growing nuances related to collectivistic cultures like India. To be more specific, the extensive relationship between the time spent and type of content consumed on Instagram and their effect on teen anxiety levels, for Indian adolescents, is poorly researched.

Modern adolescents' use of Instagram was explored thoroughly throughout the study, accompanied by explanations stating how they interact with Instagram and possible correlations to their anxiety. By investigating the frequency, duration of social media usage, and types of content, the study intends to make major gains toward an understanding of the relationship between Instagram and teen anxiety. The hypothesis revolves around investigating the relationship between excessive social media use, particularly Instagram among teenagers, and heightened anxiety.

⁷⁴ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10310995/>

⁷⁵ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9955439/>

⁷⁶ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9966679/>

By doing so, it wins on a few accounts: It can reveal how digital environments affect adolescent mental well-being, provide insights on the difficulties faced by Indian teens, as well as guide policymakers in crafting effective interventions and supportive mechanisms in the local Indian context. What is more, by filling in these boroughs, this research could enhance global awareness and advocacy efforts in the service of healthier digital habits and mental health benefits for adolescents.

Methodology

Participants

All the participants of this study were 13 to 18 years old, composed of both males and females, all from India. Participants were recruited using a convenience sampling method; friends and associates were contacted via WhatsApp. Survey was shared with 40 participants using convenience sampling by recruiting them via WhatsApp. Survey was shared with 40 participants using convenience sampling by recruiting them via whatsapp. The inclusion criteria required participants to be within the specified age range and active Instagram users, ensuring relevance to the study's focus on social media usage. No compensation or incentives were provided to the participants.

Materials

The study utilized a non-standardized survey and an anxiety assessment as its primary instruments for data collection. To quantify the variables being studied, specifically anxiety levels and Instagram usage, an anxiety test was employed alongside a non-standardized survey developed to capture detailed information about participants' Instagram habits.

For the anxiety assessment, a standardized test was used - DASS-Y, with background information and psychometric data provided to ensure its appropriateness for measuring anxiety in this context. Cronbach's alpha values were $\alpha = 0.89$ for depression, $\alpha = 0.84$ for stress. This was written by Szabó and Lovibond in 2022⁷⁷. An example question from the survey - "I got upset about little things"

The survey, administered through Google Forms, included questions about the number of followers and followings, the ratio of known to unknown followers, weekly time spent on social media, time allocation among different Instagram activities (reels, stories, posting), categories under which content is saved, engagement with Instagram ads, completion rate of watched reels, and criteria for accepting unknown followers. Additionally, participants were asked about the number of Instagram accounts they maintain. The survey was specifically created for this study to comprehensively capture various aspects of Instagram use, enabling a thorough examination of instagram usage and its relationship with anxiety. The purpose of these questions was to determine the extent of Instagram use and its potential correlation with anxiety levels. An

⁷⁷ Szabo, Marianna, and Peter F. Lovibond. "Development and Psychometric Properties of the DASS-Youth (DASS-Y): An Extension of the Depression Anxiety Stress Scales (DASS) to Adolescents and Children." *Frontiers in Psychology*, vol. 13, 2022, <https://doi.org/10.3389/fpsyg.2022.766890>.

example question from the survey was “On average, how many hours per day do you spend on Instagram?:

Procedure

The study initiated with a general, non-standardized questionnaire prepared for the collection of relevant data. This questionnaire was well designed to be very clear and effective in collecting information required by the study. Besides the non-standardized questionnaire, the standardized anxiety test was also conducted with participants. The anxiety test played a significant role in the systematic clinical assessment of participants' anxiety levels.

Once the survey and the anxiety test were finalized, they were administered to the targeted participants via Google Form. Upon receiving consent and assent from teenagers and their parent/guardian, prior to participating in data collection, the participants were required to both answer each question in the survey and the anxiety test.

Participating involved mostly online data collection through Google Forms to best enable convenient accessibility for participants. This would allow the participant to take on the survey and anxiety test at their own pace and in their own environment with minimal interference, so that each of the participants had comfortable settings that were also well-known. The participant was still clearly instructed to complete the survey and the anxiety test as candidly as possible.

The research design for the current study is correlational in nature-to assess the association between the use of Instagram and anxiety levels of the participants. The independent variable herein is the use of Instagram, whereby its assessment can be answered by questions on how frequently they use it, the amount of time spent using it, engagement levels, and user behavior. In this regard, the dependent variable herein will be anxiety levels, which was measured by a standardized scale so that measurement would be consistent and reliable.

To mitigate potential biases and ensure the validity of the findings, there were exclusion criteria specifically controlling for age, citizenship and residency. This approach helped isolate the effect of Instagram usage on anxiety by accounting for age and demographic related differences that could influence the results.

Results

This study aims to shed light on the association between social media and teen anxiety by examining variables such the frequency and intensity of instagram usage, the kinds of information consumed, and its influence on depression and anxiety levels.

Variable	Total
Number of participants	30
Number of females	18
Number of males	12
Mean age of participants	16.73 years

Table 1: Descriptive Statistics For All Participants

Participants were asked to fill up a survey about social media use and their stress, anxiety and depression levels which were measured by the DASS-Y (Depression Anxiety Stress Scales - Youth). There were a total of 30 participants. The mean age of respondents was 16.73 (SD = 1.08). The male to female ratio of girls to boys was 3:2.

Descriptive Statistics					
	Number of Followers on Instagram	Number of Following on Instagram	Minutes Spent on Instagram Per Day	On a scale of 1 to 10, how often do you experience symptoms of anxiety	Total
Valid	30	30	30	30	30
Mean	723.000	834.267	156.200	4.467	19.467
Std. Deviation	836.617	504.557	93.958	2.636	14.441
Minimum	25.000	1.000	0.000	1.000	0.000
Maximum	5000.000	2465.000	420.000	8.000	54.000

Table 2: Descriptive Statistics For Instagram Usage, Frequency, and Self-Reported Anxiety

Participants reported their instagram data based on the number of following, followers, minutes spent and a self-report likert scale of 1-10 on how often they experience symptoms of anxiety. The mean number of followers teenagers had on instagram was 723 (n = 30), SD = (834). Figure 1 shows the number of people that the participants are following on instagram. The minimum number of followers on Instagram was 25 and the maximum number was 5000. Figure 2 shows the number of following on instagram minimum was 1 maximum was 2465. Figure 3 shows the number of minutes spent on instagram per day with the minimum being 0.0 and maximum 420 minutes.

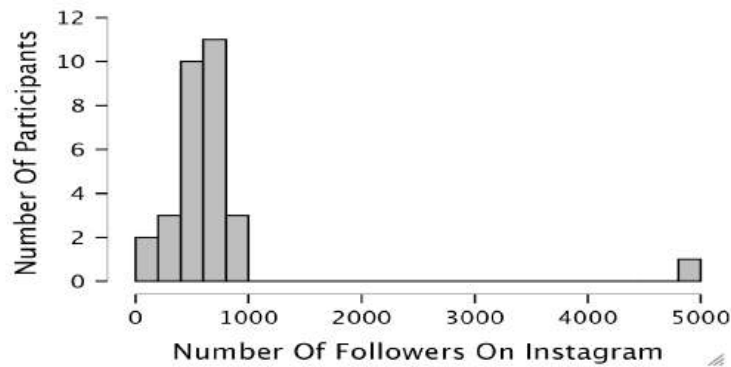


Figure 1: Distribution Plot for Number Of Followers On Instagram

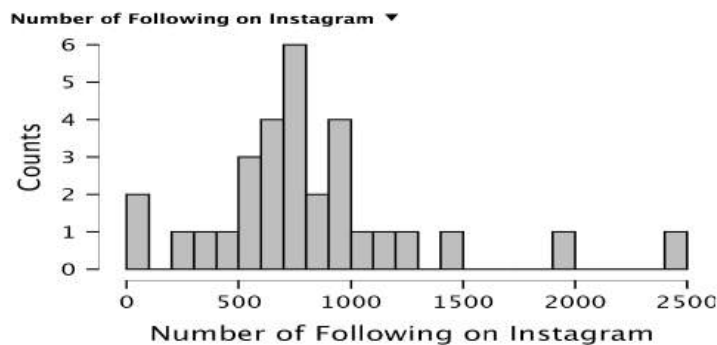


Figure 2: Distribution Plot for Number Of Following On Instagram

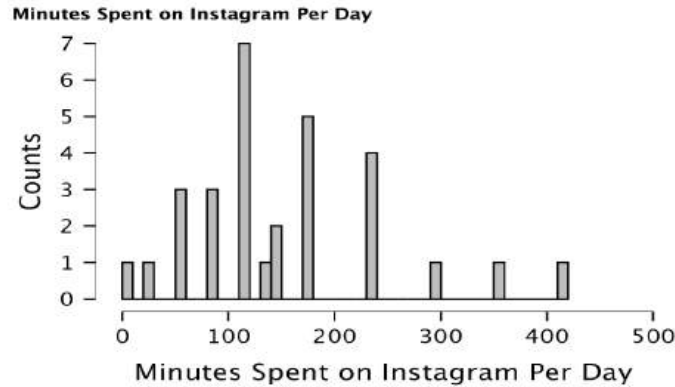


Figure 3: Distribution Plot for the Minutes Spent on Instagram Per Day

There is a moderate correlation between the stress score, as measured by the DASS-Y, and participants' self-reported experiences of anxiety ($r = 0.642$, $p < 0.001$) (Table 4). The stress score is moderately correlated to the anxiety score ($r = 0.762$, $p < 0.001$) (Table 4) and moderated to the depression score ($r = 0.541$, $p < 0.01$) (Table 4). Lastly, the anxiety score which is correlated to the depression score ($r = 0.488$, $p < 0.01$) (Table 4). Unfortunately, there was no significant correlation between instagram usage and teenage anxiety. None of the variables related to the frequency and intensity of instagram usage were correlated to anxiety levels as measured by the DASS-Y. While the majority showed a negative correlation none of the results were statistically significant.

Variable	On a scale of 1 to 10, how often do you experience symptoms of anxiety	Stress Score	Anxiety Score	Depression Score	Total
1. Number of Followers on Instagram	Pearson's r				
2. Number of Following on Instagram	Pearson's r				
3. Minutes Spent on Instagram Per Day	Pearson's r				
4. On a scale of 1 to 10, how often do you experience symptoms of anxiety	Pearson's r	—			
5. Stress Score	Pearson's r	0.642***	—		
6. Anxiety Score	Pearson's r	0.361	0.762***	—	
7. Depression Score	Pearson's r	0.336	0.541**	0.488**	—
8. Total	Pearson's r	0.545**	0.919***	0.860***	0.782***

* p < .05, ** p < .01, *** p < .001

Variable	Number of Followers on Instagram	Number of Following on Instagram	Minutes Spent on Instagram Per Day
1. Number of Followers on Instagram	Pearson's r	—	
2. Number of Following on Instagram	Pearson's r	0.213	—
3. Minutes Spent on Instagram Per Day	Pearson's r	0.099	-0.197
4. On a scale of 1 to 10, how often do you experience symptoms of anxiety	Pearson's r	-0.154	-0.021
5. Stress Score	Pearson's r	-0.126	-0.173
6. Anxiety Score	Pearson's r	-0.125	-0.134
7. Depression Score	Pearson's r	-0.173	-0.239
8. Total	Pearson's r	-0.164	-0.213

* p < .05, ** p < .01, *** p < .001

Table 4: Correlation Matrix For Instagram Usage, Frequency, and Performance on DASS-Y

Discussion

This study examined the correlation between Instagram usage and anxiety levels among Indian teenagers aged 13-18 years. The research aimed to explore how both the time spent on Instagram and the type of content consumed could contribute to anxiety in this demographic. The study sought to address gaps in existing literature by focusing specifically on Indian teens, who may face unique social and cultural pressures from social media compared to their global peers. By investigating Instagram usage patterns, the study aimed to provide insights into how social

media impacts mental health, particularly anxiety, in the context of Indian youth. This research offers a deeper understanding of the psychological effects of Instagram on a population under increasing digital and societal pressures.

The study could not reject the null hypothesis as no statistically significant correlation was found between Instagram usage and anxiety levels in teenagers. Although participants who reported higher Instagram usage showed slightly elevated anxiety levels, this correlation was not statistically significant. The most interesting thing we found was how closely DASS-Y test results related to self-reported anxiety levels. This suggests that they had a lot of insight into their mental health. This demonstrated that they had the ability to match their anxiety levels with their perceptions, so they might have some insight into their emotional states and be able to articulate them. Although the impact of Instagram use on anxiety is immeasurable, this study is significant in a sense that it reminds us how essential self-awareness is for mental health. Our results suggest the need to further study the mutual effects among social media, individual dispositions and mental health outcomes in teenagers. Studying Instagram use among teenagers may not conclusively link as a cause of increased anxiety but it definitely shows the odds ratio leans toward further increasing rates. And of course, correlation does not imply causation so just using more Instagram cannot be the reason higher anxiety is happening. Another interesting result was that self-reported anxiety levels were strongly correlated with the actual score on the DASS-Y test. This shows that participants were very high in mental health literacy. However, the fact that teenagers can so accurately judge their own anxiety levels shows they are not oblivious to this emotional perspective -- at least when it comes to social media.

Strengths Limitations & Future Research

The paper has some important strengths, in particular a single focus on Indian adolescents as a demographic very much peripherally examined in social media studies. This focus provides new insights into the experiences of a demographic which can often be overlooked within current literature. The large number of participants and the high correlation between self-reported anxiety and actual self-anxiety show that this is a valid study and the subjects are good judges of their own mental health status.

Of course, the study does have its limitations. These limitations of a relatively small sample size and use of self-reported data may introduce biases that affect the findings. Furthermore, the cross-sectional nature of our design precludes us from making any inferences about causality between Instagram use and anxiety. The study may have missed nuances that could have affected results, by not distinguishing between types of content being consumed on Instagram. These limitations may partially explain the numerous non-significant associations observed in the study and additional research with better power to detect subtler effects of these processes on health is warranted.

Future research examining how active (e.g., engaging with the app) and passive (e.g., merely browsing) Instagram engagement relate to anxiety levels may also prove beneficial. Given the relationship between self-awareness and anxiety, some form of self-monitoring in

mental health interventions for adolescents might be of benefit. These areas, if focused on in future research, may provide a more nuanced view of the complicated relationship between social media use and mental health, offering pathways to better approaches for coping with anxiety.

Conclusion

This research was done to find how Instagram usage is linked with increased anxiety in India adolescents, but the study failed to reject the null hypothesis. But it did show that teens are aware of their mental health, with there being a positive correlation between self-reported and actual anxiety scores. Although the study does not support the idea that social media use causes anxiety, its findings suggest that different people will have a range of reasons for either good or bad reactions to using SNSs and underline some new avenues that are worth investigating in future research.

The results further our knowledge of mental health in adolescence within the context of social media-related behaviors, and suggest a direction for future research. By exploring more nuanced aspects of social media engagement and its impact on mental health, future research can offer more targeted insights that inform clinical practices and educational programs. The study underscores the need for more tailored interventions to support the well-being of teens in the digital age.

Works Cited

- Elmer, Timon, et al. "Students under Lockdown: Comparisons of Students' Social Networks and Mental Health before and during the COVID-19 Crisis in Switzerland." *Frontiers in Psychiatry*, vol. 11, article no. 508595, 2020, Accessed 19 Sept. 2024.
<https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2020.508595/full>.
- Ferreira, Beatriz, et al. "Decision-Making Predictors of Pro-Environmental Behavior: The Role of Emotional Intelligence and Critical Thinking." *International Journal of Environmental Research and Public Health*, vol. 20, no. 3, 2023, Article PMC9966679, doi:10.3390/ijerph20031774.
- Jones, Sara K., et al. "The Impact of Air Pollution on Cognitive Decline: A Systematic Review and Meta-Analysis." *Journal of Alzheimer's Disease*, vol. 94, no. 1, 2023, Article PMC10310995, doi:10.3233/JAD-220456.
- Karpinski, Andrew, et al. "The Relation between Facebook Use and Academic Performance." *Cyberpsychology, Behavior, and Social Networking*, vol. 16, no. 6, 2013, pp. 423-427, Accessed 18 Sept. 2024.
<https://pubmed.ncbi.nlm.nih.gov/21290162/>.
- Liu, Zongchao, et al. "Machine Learning Approaches in Cancer Immunotherapy Prediction: Current Applications and Challenges." *Frontiers in Immunology*, vol. 13, 2023, Article PMC9955439, doi:10.3389/fimmu.2023.1045671.
- Luo, Jianbo, et al. "Superlubricity of Graphite in Inert Gas Environment." *Nature Communications*, vol. 9, article no. 1231, 2018, Accessed 15 Sept. 2024.
<https://www.nature.com/articles/s41467-018-03126-x.pdf>.
- Panza, Giulio A., et al. "Effectiveness of Exercise Interventions on Anxiety Outcomes in Patients: A Systematic Review and Meta-Analysis." *PLOS ONE*, vol. 18, no. 1, 2023, Accessed 17 Sept. 2024.
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9966679/>.
- Primack, Brian A., et al. "Social Media Use and Perceived Social Isolation Among Young Adults in the U.S." *American Journal of Preventive Medicine*, vol. 53, no. 1, 2017, pp. 1-8, Accessed 14 Sept. 2024.
<https://pubmed.ncbi.nlm.nih.gov/23745614/>.
- Przybylski, Andrew K., and Netta Weinstein. "Can You Connect with Me Now? How the Presence of Mobile Communication Technology Influences Face-to-Face Conversation Quality." *Computers in Human Behavior*, vol. 71, June 2017, pp. 106-112, Accessed 16 Sept. 2024.
<https://www.sciencedirect.com/science/article/pii/S0747563216304682>.

Binary Black Hole Merger Models: Common Envelope Theory vs Chemically Homogenous Theory By Sophia Lu

Abstract

Binary black hole mergers affect many different things inside the galaxy. Only recently, with LIGO developments, have we observed binary black hole mergers from observing gravitational wave events. Both common envelope theory and the chemically homogenous theory are theories that explain how the binary black holes merge. It is unclear as to which theory is the true explanation for the mergers. Throughout the paper, the two models, how they work, their flaws, and evidence in support of them are explored. The paper aims to answer the question: What evidence is there to support the 2 different types of merger models?

Black holes and the evolution of stars

Black holes are regions of space so packed with matter to the extent that their gravity prevents anything from escaping. They are formed through two main channels: the death of stars and the direct collapse of gasses. As massive stars age, their unstable centers eventually collapse and compress into themselves. At the end of this process, the stars explode and create a black hole. Black holes that are formed through the collapsing of gasses are more massive. This pathway is believed to have operated more commonly in the early universe (Tillman and Dobrijevic). Inside black holes, the immense intense gravitational pressure creates an escape velocity greater than the speed of light, preventing light and matter from escaping. There are two main parts of a black hole: the event horizon and the singularity. The event horizon of a black hole is the boundary at which light cannot escape and is dependent on the amount of matter the black hole accretes. Black holes can accrete gas, cosmic dust, other stars, and more, and exist at the center of all galaxies. The singularity of a black hole is “the point in space-time where the mass of the black hole is concentrated” (Tillman and Dobrijevic). Black holes can be classified by their spin, charge, and mass. The four main categories of black holes are: Kerr (spinning and uncharged), Schwarzschild (non-spinning and uncharged), Kerr-Newman (spinning and charged), and Reissner-Nordstrom (non-spinning and charged). (Baird) The most common Kerr black hole is known to have two event horizons and a ring-shaped singularity at the center of the black hole (Németi).

The radiation and winds from a black hole impacts nearby planets and their general environments. Black hole winds can drive atmospheric escape, lead to ozone depletion, and heat up the atmosphere of nearby planets (Lingam and Perlman). The radiation emitted from black holes can erode atmospheres and supply dangerous amounts of UV radiation (Randall). Black holes can also affect the formation of galaxies, as an association was found between mass of the black hole at the center of a galaxy and the size of the galaxy, and they can release energy to surrounding gas (Lewton).

GW150914

LIGO (Laser Interferometer Gravitational-Wave Observatory) is an observatory founded in 1984 to directly detect the gravitational waves that were described by Einstein's General Theory of Relativity. LIGO relies on laser interferometry to measure the ripples in spacetime caused by cosmic events like the merging of black holes.

In 2015, LIGO detected the first ever observed gravitational waves with their ultrasensitive detectors (Abbott et al.). This gravitational wave event, GW150914, originated from the most prominent binary black hole merger to be observed, around 1.3 billion lightyears away. GW150914 was observed to have a signal increase in both frequency and amplitude from 35 to 150 Hz in about 8 cycles and a little over 0.2 seconds. The most likely explanation for this observation is that GW150914 originates from the inspiral of two orbiting masses—a binary black hole merger. Utilizing generic transient search and binary coalescence search, it was determined GW150914 was a statistically significant and verifiable event, with a false alarm probability of $< 2 \times 10^{-7}$, or 5.1σ .

GW150914 originates from a binary star system of two very massive stars that eventually collapsed and formed a pair of stellar mass black holes. The two initial stars had metallicities less than half of the Sun's and the spin of the GW150914 was very low (LIGO). The merging of these two black holes released energy equivalent to 8 solar masses and created an even larger black hole of around 142 solar masses (Chu). The pathway of how this merger event and following observed black hole mergers is heavily debated and studied. The mass, metallicity, and spin of this GW150914 merger event provided researchers with essential information when coming up with merger models.

This GW150914 observation established that binary black holes can naturally form and merge in a Hubble time, the estimated age of the universe, and opened the field up to questions about how such an event could have happened. It is significant that these events can happen within a Hubble time because most black holes formed in that time. Showing that a binary black hole system can form in a Hubble time shows that it is possible for it to exist and be observed. GW150914 challenged our understanding of the life and death of stars, how mass, composition, and rotation rates affect the evolution of stars, and provides insight on the beginning of the universe. In the years following the first observation of black hole mergers, LIGO targeted a sample of binary black hole mergers and inferred their intrinsic properties like their merger rates, their masses, and their inspiral spins. This data is used to further support and explain the different proposed merger models.

Common Envelope Evolution model and Chemically Homogenous Evolution model

The chemically homogenous evolution model was proposed because researchers had argued that the black holes produced through the common envelope model did not have a large enough mass and very little metal, which was what was observed with GW150914. Because of this, researchers looked into different models to see which could better fit the features of the observed binary black holes (Wolchover et al.). The chemically homogeneous process includes rapidly rotating binary stars spinning at specific speeds and conditions that allow the stars to

evolve chemically homogeneously as an explanation for the binary black hole merging. Internal mixing from fast stellar rotation influences the evolution of binary black holes by transporting material from the hydrogen-rich envelope to the central burning regions of the star and vice versa. If this rotation is efficient enough to prevent a build-up of a chemical gradient which separates the core of the star from the envelope, the stars can evolve homogeneously. As this process continues, hydrogen is exhausted in both the center and throughout the envelope of each star, resulting in the two stars contracting towards their helium centers, staying within their Roche lobes, the region around a star in a binary system within which any material is gravitationally bound to that particular star (“Roche Lobe”). The two close stars then evolve into two separate black holes, neither overflowing their Roche lobes nor initiating mass transfer like with common envelope evolution. These two black holes spin together and eventually merge 4-11 Gyr after they are formed. This model finds that more massive stars with lower metallicity are more prone to evolving chemically homogeneously. At higher metallicity, angular momentum loss slows down the stars and shuts off rotational mixing. More massive stars are more prone to chemical homogenous evolution as they have a weakened entropy barrier due to radiation pressure (Mandel and de Mink).

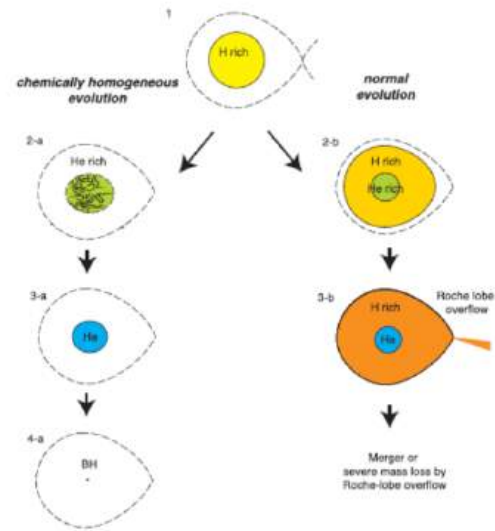


Figure 1

Common envelope evolution model proposes a process that emphasizes unstable mass transfer as an explanation for the binary black hole merging. With this model, the two stars initially spin with a very wide orbit. As one star evolves, its hydrogen envelope enlarges and forms a red supergiant. The envelope of this red supergiant then interacts with and enters the gravitational field of the smaller star, sucking the smaller star away. This process draws the two stars together until the smaller star overgrows the red supergiant and the supergiant collapses into a black hole. Now, when the second star forms a red supergiant, it engulfs the black hole in a common envelope. The second star then collapses into a black hole as its hydrogen envelope is lost to space. The two black holes are then close enough to eventually merge. It is assumed that a binary enters a common envelope phase if both binary components fill their Roche lobes, if the mass transfer rate exceeds $0.1M y^{-1}$, or if the photon trapping radius of a compact object accretion extends beyond the Roche lobe radius (Mandel and de Mink).

This paper aims to answer the question: What evidence is there to support the 2 different types of merger models? Researching the different types of models is important because it can give us potential answers to many different things relating to the beginning of the universe. It also provides us with more information and context for other events that occur in space.

Evidence for the two models

Chemically homogeneous model

Chemically homogeneous evolution is difficult to investigate because massive, rapidly rotating stars with low metallicity are very rarely observed. This in itself is a major flaw of this model. Majority of the merger events include stars that do not spin very rapidly, so there is little hard evidence for chemically homogenous evolution. However, there are some observational clues and simulations that can be interpreted as evidence for this model.

In 2016, Mandel and de Mink performed Monte Carlo simulations of expected merger rates for a model that underwent chemically homogenous evolution. The process went as follows: estimating the rate of local mergers to serve as a sanity check for the Monte Carlo simulation results, creating a model of massive binary populations over cosmic time that accounts for various factors, performing the Monte Carlo simulation, and analyzing the results. While simulating massive binary populations, a few assumptions were made:

- 1) The primary mass m_1 follows a Kroupa initial mass function
- 2) The distributions are separable and measured with fair approximation for the distribution of binary properties at higher redshift and low metallicity.
- 3) The minimum equatorial velocity is as follows:

$$\omega_c = \begin{cases} 0.2 + 2.7 \times 10^{-4} \left(\frac{m}{M_\odot} - 50 \right)^2 & \text{for } m < 50 M_\odot, \\ 0.2 & \text{for } m \geq 50 M_\odot. \end{cases}$$

- a)
 - 4) Mass loss driven by stellar winds and envelope ejection during the final explosion is accounted for by adopting “ $f_{MS} = 0.1$ for the fraction of mass that is lost during the main-sequence evolution, and $f_{WR} = 0.25$ for the fraction of mass lost during post-main-sequence evolution as a Wolf–Rayet star”
 - 5) Changes in orbit due to wind-driven/supernova mass loss are accounted for by assuming the mass loss is both fast and spherical compared to the orbital motion and the angular momentum of mass lost is equal to the orbital angular momentum of the star.
 - 6) The binary remains circular throughout its evolution.

A graph presented by Mandel and de Mink of the star formation rate (SFR) versus redshift shows the SFR for different metallicities (fig 2). The solar metallicity predicted for the chemically homogenous model is $Z = 0.0134$ (Asplund et al. 2009) with a mean metallicity of $\sim 1.06 \times 10^{-0.15} Z$ with a standard deviation of $\sim 0.38Z$, with significant uncertainty. This metallicity corresponds to the red dotted

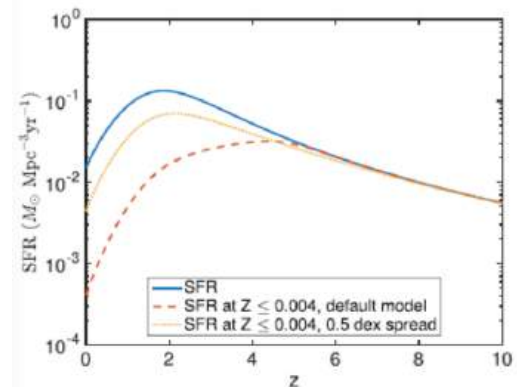


Figure 2

line in the graph and shows that the star formation rate of the chemically homogenous model increases the further back in time we look and that there is a very small formation rate in recent years. This is significant because it shows that as we progress, we are more likely to observe gravitational wave events that are the result of chemically homogenous mergers, as it takes a sizable amount of time before we can observe events in the past.

The results of the Monte Carlo simulation find that out of 10^8 simulated binaries, 500 binaries satisfy the conditions for chemically homogenous mixing. The merger rate is found to increase the further we go back in cosmic time, before dropping again as we approach present day, with a peak merger rate of $20 \text{ Gpc}^{-3} \text{ yr}^{-1}$ at $z < 0.5$ as shown in figure 3. Within figure 3, the solid black line represents the total binary black hole merger rate, while the colored, smaller lines represent the 500 sample binary systems. This simulation shows that within sufficiently massive tight binaries, mixing processes can indeed give rise to stars evolving chemically homogeneously.

The results of this simulation serve as evidence for the significance of the chemically homogenous model. Systems of close massive binary black stars can give rise to significant merger rates with peaks of $20 \text{ Gpc}^{-3} \text{ yr}^{-1}$. This merger rate is also consistent with the lack of LIGO detections that placed upper limits ($70\text{-}170 \text{ Gpc}^{-3} \text{ yr}^{-1}$) above the predicted merger rate of $10 \text{ Gpc}^{-3} \text{ yr}^{-1}$. By avoiding the common-envelope phase, the chemically homogenous model is able to avoid the uncertainties that are associated with that formation channel and does not produce short time-delay mergers.

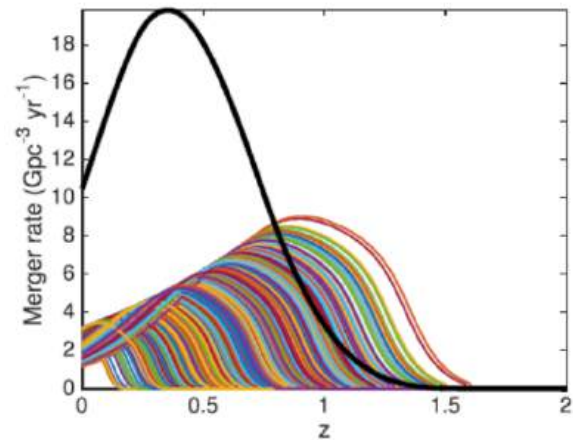


Figure 3

Common Envelope Model

Olejak & Belczynski showed that fast-spinning black holes in merging binary systems can be formed by tidal spin up and a common envelope phase, forming two equal mass black hole components. In low metallicity environments, a large number of massive stars are predicted to survive common envelope evolution (Zevin et al.).

F. Lyu et al analyzed data from GW190814 and showed that GW190814 and similar events have all the initial conditions for common envelope evolution. They identify that GW190814 was more likely formed through common envelope evolution than chemically homogenous evolution.

The graph of the result of the simulations shows that with a short enough orbital period ($< 0.5\text{d}$), events that go through common envelope evolution can end up with spins that match with the observed black hole spin ($x > 0.3$). As shown in Figure 4, as the orbital period decreases, the black hole spin starts to increase. This provides evidence for common envelope evolution by

being able to match up observational data about the spin of the black hole with simulated spins of resultant common envelope merger black holes.

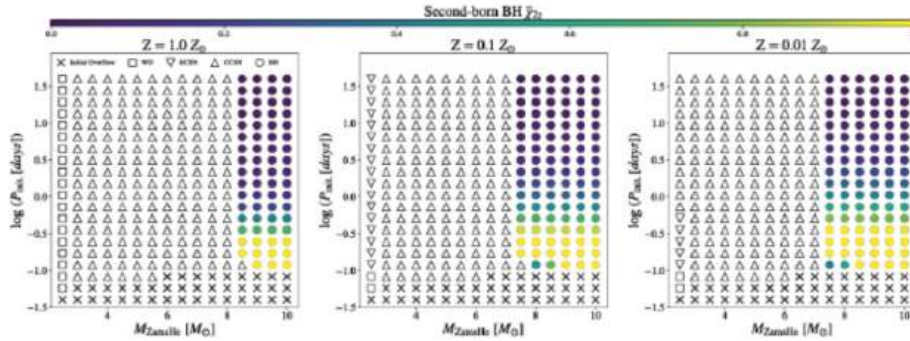


Figure 4

In 2021, Zevin et al. modeled five different channels: late-phase common envelope, stable mass transfer, chemically homogenous evolution, metal-poor globular clusters (GCs), and formation in neutral star clusters (NSC). Looking at a network of LIGO Hanford, Livingston, and Virgo operating at signal to noise threshold of $p=10$, they found the detection probabilities of the different models. Unlike many other investigations, they accounted for spin, which is said to have greatly affected their results. Looking at their results, they determined that the common envelope channel dominates the underlying binary black hole population in their models. When considering only the common envelope and metal poor GC models, they found that around 90% of the underlying population comes from the common envelope channel. They also found that there was a preference for highly inefficient common envelopes ($a = 0.2$) by a Bayes factor of >10 , with a marginal preference for larger values ($a = 5.0$) by a Bayes factor of 5 (Figure 5).

Table 1
Bayes Factors $\log_{10}(B)$ across χ_b Models (Columns) and α_{CE} Models (Rows)

		χ_b				
		0.0	0.1	0.2	0.5	
α_{CE}	0.2	-0.63	-0.56	-1.24	-1.71	-0.35
	0.5	-0.06	-0.58	-0.96	-1.11	0.00
	1.0	$\equiv 0$	-0.77	-1.02	-1.29	$\equiv 0$
	2.0	0.34	0.05	-1.19	-1.15	0.42
	5.0	0.56	0.39	-0.54	-0.87	0.70
		$\equiv 0$	-0.27	-1.11	-1.35	

Figure 5

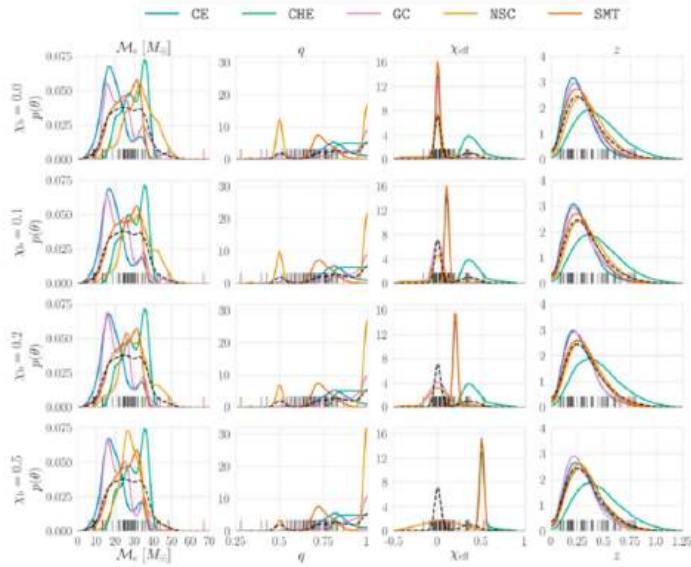


Figure 6

The results of considering all five formation channels with varying natal spin prescriptions is shown in Figure 6. When testing for chirp mass, at lower natal spins, common envelope evolution and chemically homogenous evolution are competitive, with the largest probability peaks. As natal spin increases however, chemically homogenous evolution and NSC dominate common envelope evolution. When testing for redshift, common envelope evolution proves to be dominant for all lower natal spins, only being overtaken by GC at $X=0.5$.

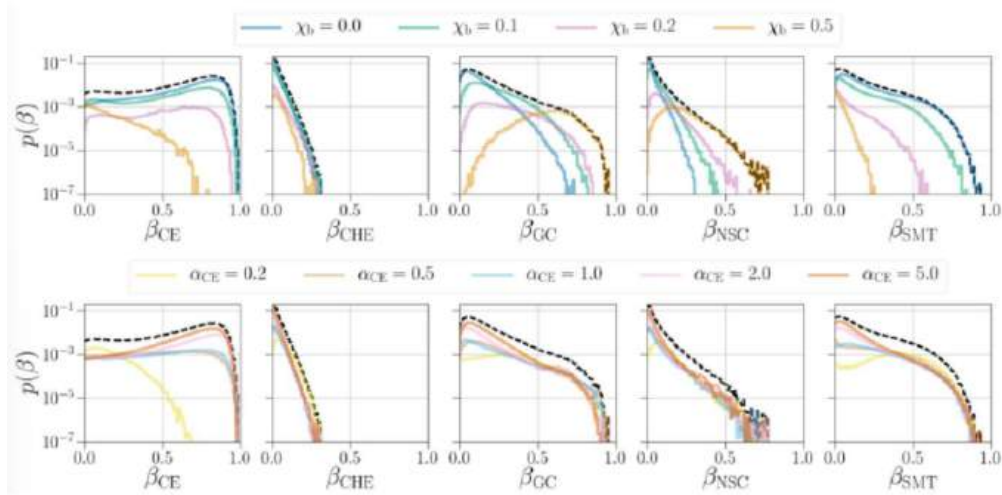


Figure 7

When looking at the inferred branching fractions for all five channels shown in Figure 7, common envelope evolution dominates the field channels which make up the majority of the underlying population.

Differences between the two models

The chemically homogeneous model, unlike the common envelope, requires rapid spin and rotation for instabilities in the chemical makeup of stars to prevent the build up of a chemical gradient. It also requires low metallicity and larger initial star mass. The common envelope model, on the other hand, requires overflowing of the roche lobe, and initiation of unstable mass transfer with rates above $0.1 M \text{ y}^{-1}$. When approaching the question of the more dominant model, we need to consider the features of the typical binary black hole and how they align with the characteristics required from either model.

Other potential channels

The Marchant et al. investigation showed that the stable mass transfer channel could produce more mergers within a Hubble time than common envelope evolution could. This opens up the field to more possibilities than just the two channels of common envelope and chemically homogenous evolution. Zevin et al. provided us with channels like NSC and GC that were competitive with both common envelope and chemically homogenous evolution. When looking at inferred branching fractions for all five channels, the peaks of the NSC and SMT channel are even larger than the peaks of common envelope evolution, showing that they are competitive and worth further examining.

Additionally, Zevin et al.'s simulations were unique in that they considered different numbers of channels. They considered all five, just three channels, and just two channels, and saw that the results were different. This shows the importance of considering other possibilities and that going forward, for more accurate conclusions, it is necessary to take into consideration other formation channels.

Flaws

The chemically homogeneous model requires rotational rates of 20-30% of the Keplerian velocity. Because this can only be achieved when two nearly equal mass stars are close to filling their Roche lobes, there is a very small parameter space window for chemically homogenous evolution. "The spin periods of wide-period systems are too low for chemically homogenous evolution while short-period systems would have already overflowed their Roche lobes at zero age" (Mandel and de Mink). The window for chemically homogenous evolution, shown in figure 8 is significantly smaller than the window for normal evolution, and the chances of a system adhering to the dotted line, the binaries that satisfy a more stringent threshold on chemically homogeneous evolution, are even smaller. The likelihood of having the right mass ($>40 M$) and orbital period (1.5-2.5d) is very low and therefore puts the chemically homogenous model at a disadvantage when considering the more

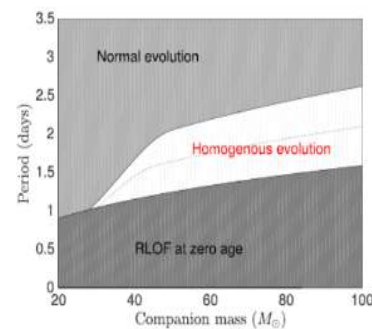


Figure 8

dominant model (Mandel and de Mink).

There are significant uncertainties with the chemically homogenous channel, including the efficiency of the mixing processes in tidally locked binaries which could potentially close off this channel entirely. Additionally, there have been observational electromagnetic constraints. In principle, it may be possible to observe chemically homogenous evolution, but the phase is short lived and observations can only be done in environments where low metallicity is rare. This constrains the possible galaxies that we can examine to look for traces of observational evidence of chemically homogenous evolution. Predictions regarding the merger rates of chemically homogenous evolution are not sensitive to Roche lobe overflow, because of how poorly understood it is, and common envelope phases, while being very sensitive to uncertain internal mixing processes (Mandel and de Mink).

Common envelope evolution has large uncertainties surrounding its efficiency, the mass boundary at which it terminates, and the conditions required for it to initiate. Marchant et al. investigated the role of common envelope evolution in the formation of merging binary black holes by computing binary simulations using MESA code of a 30 M donor star in low metallicity environments with a black hole companion. They computed mass transfer rates, modeled mass transfer and the common envelope phase, and applied their model to 2 other models where there is an increased overshooting and increased mass loss rates. The results of their simulation of the overshooting model show that there were no cases of successful envelope ejection from the common envelope phase below initial orbital periods $\log_{10}(\pi/d) < 3.2$. Below this range, the systems undergoing stable mass transfer and merging through common envelope evolution varies. They find that common envelope evolution only happens for black hole masses that are less than 5 M, but that stripping the hydrogen envelope of a 30M star already results in a 14M helium core. Additionally, the models with efficiency $\alpha = 1$ that do manage to eject their envelopes during common envelope evolution result in wide black hole binaries which cannot merge within a Hubble time.

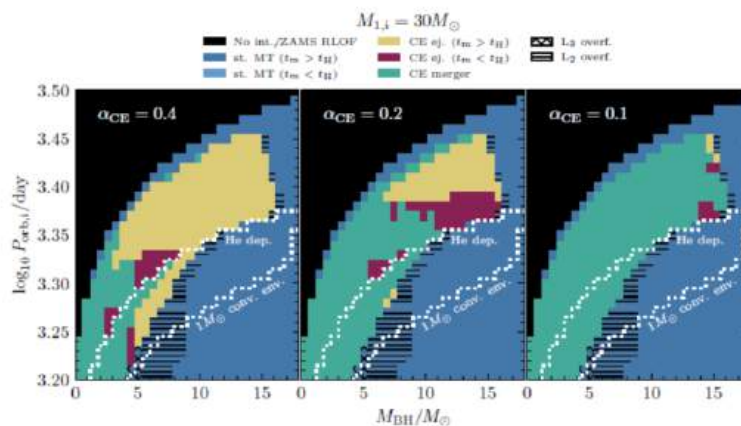


Figure 9

The outcome of the simulations with CE efficiencies of 0.4, 0.2, and 0.1 are represented in figure 9. This graph shows that most systems that produce wide binary black holes would merge during common envelope evolution, with the chances of merging increasing as efficiency decreases. With all three cases of different efficiencies, envelope ejection cannot be found below an initial orbital period $\log_{10}(\pi/d) < 3.2$. The majority of common envelope ejections observed resulted in mergers that took greater than a Hubble time merge. Even with the extreme cases that a system that undergoes common envelope evolution can produce a black hole that merges in less than a Hubble time, the ratio between the systems formed through common envelope evolution to the systems formed via stable mass transfer is 0.35. When examining the model with an efficiency of $\alpha=1$, we see that the ratio is even smaller, at 0.017.

The results of this investigation show that cases of successful envelope ejection in common envelope evolution is very rare, and the cases that were successfully simulated were unable to merge within a Hubble time. The results also show that the channel of stable mass transfer seems far more dominant.

Moving forward

As the detected population of binary black holes increases, the diversity of potential formation channels also increases. Moving forward, people plan to include additional formation channels and make more comprehensive analysis including the new data provided. There are significant uncertainties with the two common envelope and chemically homogenous evolution models that can only be clarified with research and understanding of other factors like the Roche lobe. More observational evidence and information like the spin of black holes in the binaries, their metallicities, and masses is also necessary to make arguments for either model. Because most mergers happened very long ago, in the future, we will observe more gravitational wave events of binary black hole mergers, and the data from LIGO detectors can be used to support different models. There are uncertainties regarding the efficiency of common envelope evolution and mixing processes in chemically homogenous evolution that people plan to investigate to clear up the picture on what the true dominant channel, or channels, of formation is.

Works Cited

- Abbott, B. P., et al. "Observation of Gravitational Waves From a Binary Black Hole Merger." *Physical Review Letters*, vol. 116, no. 6, Feb. 2016, <https://doi.org/10.1103/physrevlett.116.061102>.
- Baird, Christopher S. "Are There Different Types of Black Holes?" *Science Questions With Surprising Answers*, 2 Aug. 2022, www.wtamu.edu/~cbaird/sq/2022/08/02/are-there-different-types-of-black-holes/#:~:text=If%20we%20just%20focus%20on,and%20do%20have%20a%20net.
- Chu, Jennifer. "A 'Bang' In LIGO and Virgo Detectors Signals Most Massive Gravitational-wave Source Yet." *MIT News | Massachusetts Institute of Technology*, 2 Sept. 2020, news.mit.edu/2020/ligo-virgo-gravitational-wave-0902.
- "Colliding Black Holes Tell New Story of Stars | Quanta Magazine." *Quanta Magazine*, 9 Nov. 2022, www.quantamagazine.org/colliding-black-holes-tell-new-story-of-stars-20160906.
- Lewton, Thomas. "What Drives Galaxies? The Milky Way's Black Hole May Be the Key. | Quanta Magazine." *Quanta Magazine*, 23 Aug. 2022, www.quantamagazine.org/what-drives-galaxies-the-milky-ways-black-hole-may-be-the-key-20220823.
- Lyu, F., et al. "Revisiting the Properties of GW190814 and Its Formation History." *Monthly Notices of the Royal Astronomical Society*, vol. 525, no. 3, Aug. 2023, pp. 4321–28. <https://doi.org/10.1093/mnras/stad2538>.
- Mandel, Ilya, and Selma E. De Mink. "Merging Binary Black Holes Formed Through Chemically Homogeneous Evolution in Short-period Stellar Binaries." *Monthly Notices of the Royal Astronomical Society*, vol. 458, no. 3, Feb. 2016, pp. 2634–47. <https://doi.org/10.1093/mnras/stw379>.
- Tillman, Nola Taylor, and Daisy Dobrijevic. "Black Holes: Everything You Need to Know." *Space.com*, 19 May 2023, www.space.com/15421-black-holes-facts-formation-discovery-sdcmp.html.
- Németi, István. "Figure 11.38. A Slowly Rotating (Kerr) Black Hole Has Two Event..." *ResearchGate*, [www.researchgate.net/figure/A-slowly-rotating-Kerr-black-hole-has-two-event-horizons-and-a-ring-shape-singularity_fig20_267551352#:~:text=38.-,A%20slowly%20rotating%20\(Kerr\)%20black%20hole%20has%20two%20event%20horizons,plane%20of%20axes%20x%2C%20y.](http://www.researchgate.net/figure/A-slowly-rotating-Kerr-black-hole-has-two-event-horizons-and-a-ring-shape-singularity_fig20_267551352#:~:text=38.-,A%20slowly%20rotating%20(Kerr)%20black%20hole%20has%20two%20event%20horizons,plane%20of%20axes%20x%2C%20y.)
- Randall, Ryan. "Impact of black hole winds, radiation examined in new study." *Phys.org*, 31 Mar. 2022, phys.org/news/2022-03-impact-black-hole.html.
- Parshall, Allison. "Colliding Supermassive Black Holes Discovered in Nearby Galaxy." *Scientific American*, 20 Feb. 2024, www.scientificamerican.com/article/colliding-supermassive-black-holes-discovered-in-nearby-galaxy/#:~:text=Logically%2C%20these%20giant%20black%20holes,processes%20in%20their%20host%20galaxies.
- "Roche Lobe." *Oxford Reference*, <https://doi.org/10.1093/oi/authority.20110803100425362>.
- Zevin, Michael, et al. "One Channel to Rule Them All? Constraining the Origins of Binary Black Holes Using Multiple Formation Pathways." *The Astrophysical Journal*, vol. 910, no. 2, Apr. 2021, p. 152. <https://doi.org/10.3847/1538-4357/abe40e>.

Investigation of the Transmembrane Regions of Neurotrophin Receptors

By Angela Chou

Abstract

The transmembrane domains of neurotrophin receptors are essential for their structural integrity and function, yet their hydrophobic nature makes them challenging to study. This research focuses on the preparation and characterization of the p75 neurotrophin receptor (p75NTR) TM domain (TMD) for structural analysis. We developed an efficient method to produce isotopically labeled p75NTR TMD by expressing the hydrophobic domain in *E. coli* as a fusion protein, sequestered in inclusion bodies. Using a series of purification steps, including Ni-NTA affinity chromatography, cyanogen bromide cleavage, and reverse-phase HPLC, we achieved high-purity TMD samples. These samples were then reconstituted into lipid bicelles to mimic the native membrane environment, allowing for successful structural studies via NMR spectroscopy. My optimized protocol for TMD preparation and reconstitution provides a robust platform for investigating the oligomerization and signaling mechanisms of neurotrophin receptors, particularly for future drug screening applications.

1. Implication of receptor transmembrane domains in receptor function and disease

The single-pass transmembrane (TM) proteins account for the vast majority of signaling receptors on the cell surface; currently over 1300 of them have been identified serving many essential cellular processes, including signal transduction, cell adhesion and transmembrane transport of materials¹. In general, these receptors comprise an extracellular domain (ECD) that binds ligands, a transmembrane domain (TMD) that traverse the cell membrane once, and an intracellular domain (ICD) that interacts with intracellular proteins to activate downstream cascade of signaling pathways (Fig. 1A). In the canonical view of ligand-induced activation of receptors, ligand binding to the receptor ectodomain causes the receptor to oligomerize, which then triggers high order assembly of downstream signaling complexes to fully activate signaling^{2,3} (Fig. 1A). The TMDs of these receptors have been commonly thought of as the membrane anchors but mounting experimental data have indicated that they often play critical roles in receptor signaling, and the diversity of mechanism with which the TMD can promote signaling is beyond the traditional views in receptor biology⁴⁻⁶.

In 1986, a study from Robert Weinberg's lab at MIT reported that the *neu* receptor (rat equivalent of the human receptor tyrosine kinase HER2) harbors an oncogenic mutation within the TMD of the receptor⁷. This mutation – Val664 replaced by Glu – resulted in ligand-independent activation of the growth factor receptor and thus cancer (Fig. 1B). Analogous oncogenic TMD mutation V659E in the human receptor HER2 has also been identified to be a rare oncogenic driver of lung adenocarcinomas and is associated with poor prognosis in non-small lung carcinoma patients. This finding suggested, for the first time, that the receptor TMD is not merely a membrane anchor but plays an active role in receptor activation. However,

due to the technical challenges of structural and functional investigation of the TMDs, the precise mechanism of this mutation remains unclear to this day.

In addition to growth factor receptors, non-cysteine disease-causing mutations in the TMD have been found in several cytokine receptors. A S505N mutation in the thrombopoietin receptor (TpoR) results in constitutive TpoR activation and essential thrombocythemia, a hereditary disease characterized by excessive production of platelets^{9,10} and this mutation is believed to promote a mode of TpoR TMD dimerization that activates downstream signaling independent of thrombopoietin¹¹.

More recently, receptor TMDs have been brought under spotlight again by a couple of publications suggesting that fast-acting antidepressants such as ketamine (Fig. 1C) and lysergic acid diethylamide (LSD) may act on the neurotrophin receptor TrkB by directly binding to the receptor TMD and facilitating receptor activation^{12,13} (Fig. 1D). Ketamine and LSD are recreational drugs that cause healthy people to hallucinate but can provide treatment benefits for patients with severe depression. But the receptor target of these psychedelic compounds relevant to elicited anti-depressant effects still remains elusive.

The above studies have generated renewed interest to characterize the structural and functional roles of TMDs in neurotrophin receptor signaling and to examine if these roles can be specifically modulated by the psychedelic drugs. I hypothesize that homotypic or heterotypic interactions amongst receptor TMD, which is the control relay between the ECD and ICD, plays a specific and active role driving receptor signaling. This active role could be TMD dimerization, trimerization, or even higher order clustering, and disease mutations or small molecule drugs could influence receptor signaling by modulating TMD oligomerization.

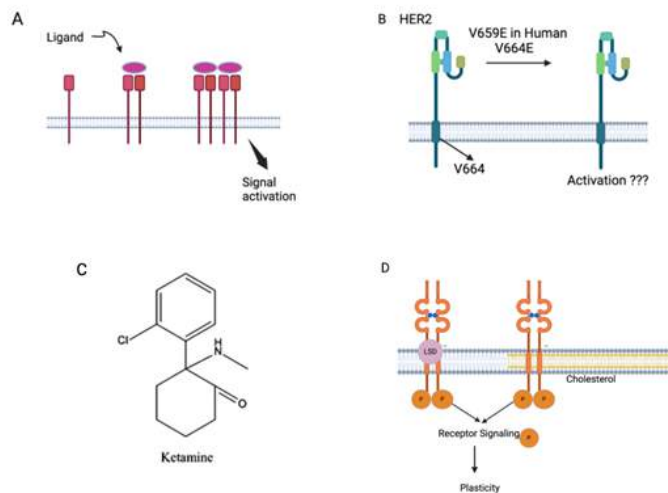


Figure 1. Conceptual models for ligand induced receptor activation for single-pass transmembrane receptors. A) Schematic illustration of ligand binding induced receptor clustering and activation, showing sequential formation of dimers and tetramers after ligand binding that activates downstream cascade. B) Schematic illustration of how a single oncogenic mutation within the TMD of HER2 causes ligand independent activation of HER2. C) Chemical structure of ketamine. D) A proposed mechanism by which ketamine or LSD activates TrkB by binding and causing conformational changes in the TMD of TrkB.

2. Existing methods for structural studies of receptor TMDs

Structural studies of single-pass TMD oligomerization have historically been extremely difficult. The rapid advance in cryo-electron microscopy (cryo-EM) in the last ten years has renewed efforts to resolve single-pass TM receptors that include the TMD. But these goals are still exceptionally challenging. The TMD alone is too small for cryo-EM, but when being examined as a part of the full-length receptor, random reorientation between the large ectodomain and the TMD can pose a fundamental problem for cryo-EM particle averaging¹⁴.

Despite the challenge, scientists have used several methods to obtain structural information of TMD dimerization. One is biochemical method, involving scanning mutagenesis and reading out oligomerization using for example the bacteria two-hybrid assay (schematically illustrated in Fig. 2A)¹⁵. This approach, however, can only provide very low resolution information. Lipidic cubic phase crystallography is another method for solving atomic resolution structures¹⁶, although the likelihood of obtaining well diffracting crystals is very low (Fig. 2B). Solution NMR application has been the most successful for solving structures of the TMDs, exemplified by the original work of Engelman and coworkers on the TMD dimer of the glycophorin A protein (GpA)¹⁷. They solved the structure of the GpA TMD dimer in detergent micelles by NMR (Fig. 2C) and revealed the glycine-XXX-glycine motif for TM helix dimerization in the membrane. NMR has since been widely applied to investigate the structures of small TMDs in detergent micelles and bicelles^{4,18-26}.

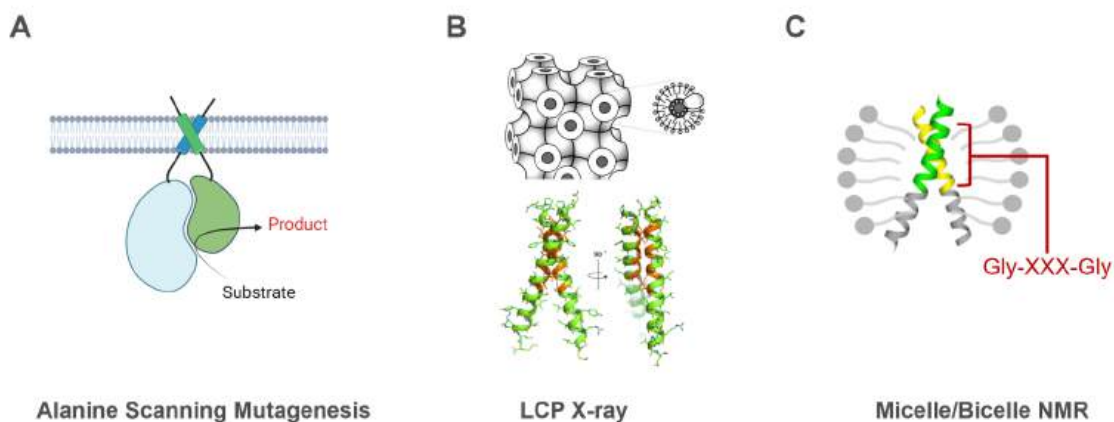


Figure 2. Known methods for studying the structures of TMD oligomerization. A) Schematic illustration of a bacterial two-hybrid assay for reporting TMD-TMD interaction. Two complementary domains of an enzyme were fused, separately, to the TMDs under investigation. Specific TMD-TMD interaction would bring the two enzyme domains together to have catalytic activity. B) Crystallization of TMD in the lipidic cubic phase for structure determination by X-ray diffraction. C) Reconstitution of TMD in detergent micelles for structure determination by solution NMR.

3. Preparation of neurotrophin receptor TMD for drug-binding characterization

The TMDs of neurotrophin receptors are very hydrophobic, which makes cellular protein expression and purification extremely difficult. Since NMR studies require the polypeptide chain to be uniformly labeled with stable isotopes such as ¹⁵N and ¹³C, chemical synthesis of TMDs would be prohibitively expensive. Here, I describe the adaptation of an effective TMD

production method²⁷ to produce isotopically labeled TMD of the p75 neurotrophin receptor of high quality.

The p75 neurotrophin (NT) receptor, p75NTR, was identified more than a decade ago as a TM protein capable of binding the nerve growth factor (NGF)²⁸. After many years of study, p75NTR emerged as an important receptor which can signal independently as well as modify the signaling capacities of receptors in the tropomyosin-related kinase (Trk) family of receptor tyrosine kinases. Like other cytokine and growth factor receptors, p75NTR comprises the ECD, TMD, and ICD. The predicted TMD is from residues Leu252 – Phe273, and hence we aimed to produce a TMD-containing p75NTR fragment. (Fig. 3A and 3B). As calculated by the program ProtParam (<https://web.expasy.org/protparam/>), this sequence yields a very hydrophobic peptide with a GRAVY score of 0.969. The GRAVY score for a typical TMD is 0.5 – 1.5; less than 0.5 means hydrophilic, and greater than 1.5 means extremely hydrophobic.

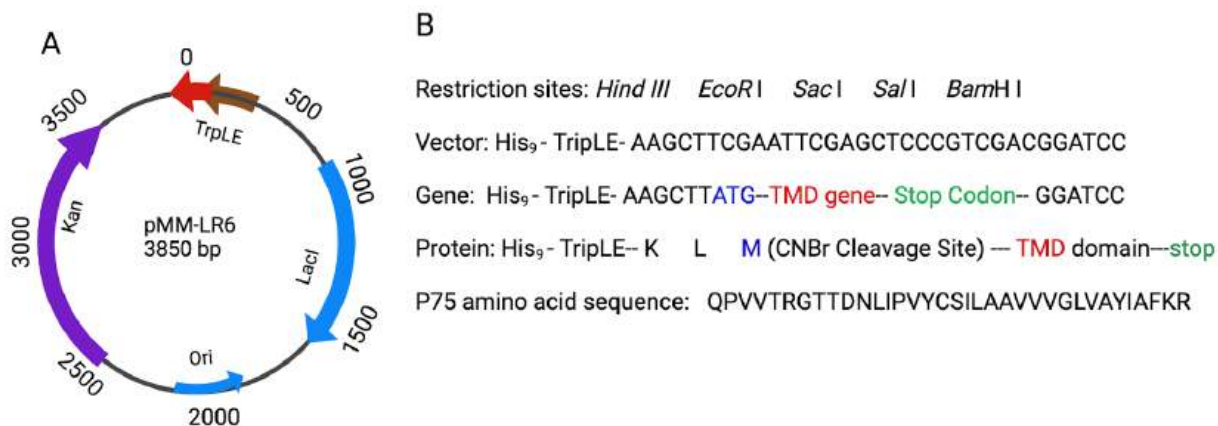
Since the GRAVY score indicates high hydrophobicity, the p75 peptide is toxic to cells as it will insert into lipid bilayers and thus rupture the cell membrane. Together with senior researchers, I expressed the small hydrophobic TMDs in *E. coli* as C-terminal fusion to TrpLE, an unfolded protein that drives inclusion body formation upon synthesis by the ribosomes (Fig. 3A). As such, the toxic TMDs were sequestered in the inclusion bodies and thus could be expressed at a high level. This was done using a DNA plasmid named pMM-LR6^{29,30}. Based on the pMM-LR6 plasmid, a His₆-tag was added to the N-terminus of TrpLE to facilitate purification. A methionine was inserted between the TrpLE and the p75 TMD to enable cleavage at this position with Cyanogen Bromide (CNBr) so that the two fragments can be separated (Fig. 3B). Hence, the TM sequence cannot contain any methionine. The purification of the TMD can be achieved with three major steps, all under denaturing conditions: 1) purification of the TrpLE-TMD fusion from inclusion bodies by Ni-NTA affinity, 2) CNBr cleavage to separate TrpLE and TMD, 3) purification of the TMD by reverse phase high-pressure liquid phase chromatography (HPLC).

Specifically, the TrpLE-(p75 TMD) fusion protein was expressed in *E. coli*. Protein expression was confirmed by SDS-PAGE analyses of the cell lysates before and after induction by IPTG (Fig. 3C). Moreover, cleavage by CNBr separated the p75 TMD from TrpLE. After cleavage, the p75 TMD was separated by HPLC using the Zorbax SB-C3 column. The HPLC profile and SDS-PAGE analysis both indicated high protein purity (Fig. 3D). SDS-PAGE analysis of the TMD powder also resolved the protein as a monomer (~3.5 kDa), consistent with the expected MW of the TMD.

The target TMDs were then reconstituted in bicelles that are sufficiently large to mimic the bilayer environment of the membrane. Bicelle is a self-assembled, soluble lipid disc in which the lipid bilayer region comprises long-chain lipids and short-chain detergents that solubilize the lipid patch by forming a rim around it³¹. The bicelles composed of 1,2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) and 1,2-dihexanoyl-sn-glycero-3-phosphocholine (DHPC) have been extensively characterized. When the lipid/detergent ratio (q) is greater than 0.5, DMPC/DHPC bicelles are known to

become disc-like in which the lipids and detergents are largely segregated³²⁻³⁴. When a protein is reconstituted in such bicelles, typically with $0.5 \leq q \leq 0.6$, it is essentially in a lipid bilayer environment while still amenable for high resolution NMR spectroscopy. To incorporate TMDs into bicelles, the purified and lyophilized TMD is dissolved in hexafluoro-isopropanol (HFIP) with a suitable amount of lipids, followed by drying of the solution under nitrogen stream to achieve thin films. The thin films are then dissolved in an 8 M urea solution containing a calculated amount of detergent. Reconstitution begins as the denaturant is removed by dialysis. The q value of the sample was accurately measured using 1D ^1H NMR.

After reconstitution into bicelles, we recorded a two-dimensional NMR spectrum known as ^1H - ^{15}N HSQC on a 900 MHz NMR spectrometer (Fig. 3E). This spectrum correlates the ^1H frequency (X-axis) and ^{15}N frequency (Y-axis) of the backbone amide proton and nitrogen, respectively, of each of the amino acids in the protein. Thus, each peak in the 2D spectrum corresponds to a residue in the polypeptide sequence. In the NMR spectrum, there are 30 peaks, consistent with 32 non-proline residues of the p75 TMD. Moreover, the peaks are well dispersed due to different local electronic environments, indicating that the TMD adopts a specific conformation.



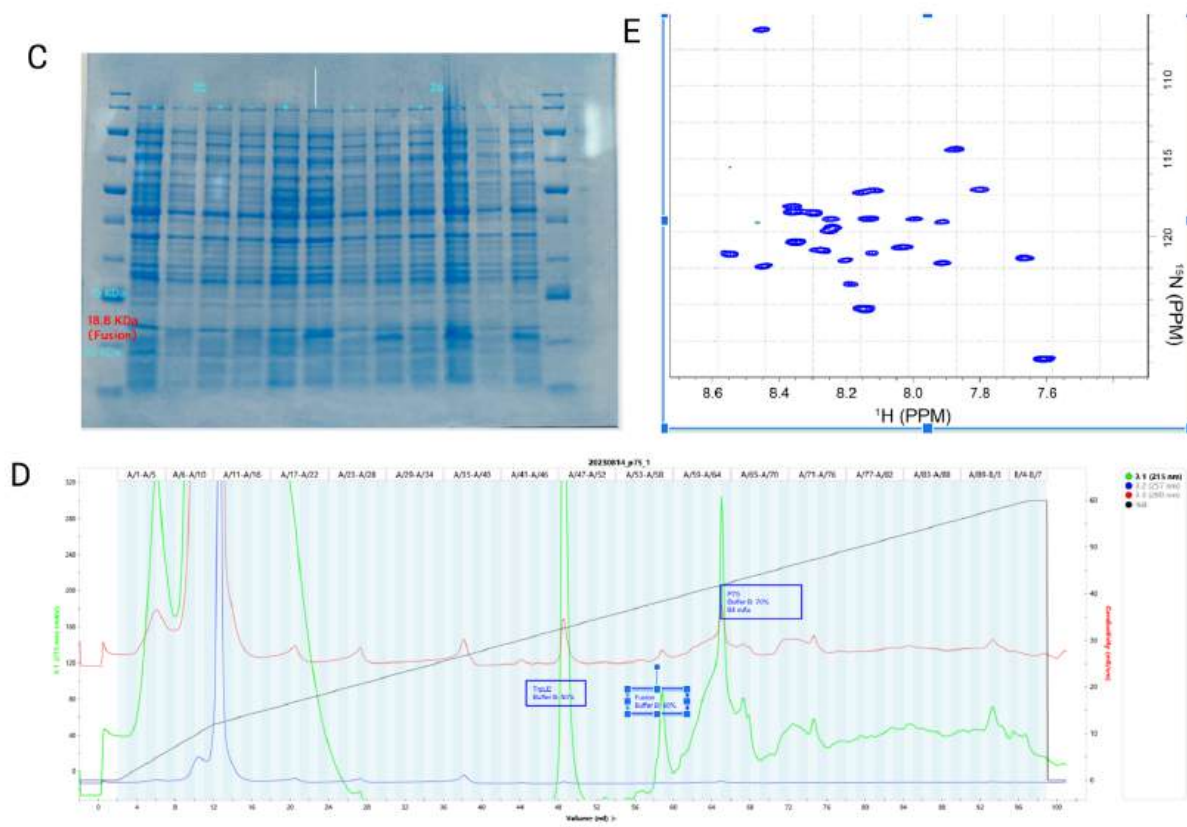


Figure 3. Expression, purification, and NMR analysis of the transmembrane domain of p75NTR.

(A) Drawing of the pMM-LR5 plasmid for inducing *E. Coli* to express the TrpLE-TMD fusion protein. B) Details of the expression plasmid including the restriction sites, vector, gene, and protein makeups to indicate the locations of cyanogen bromide cleavage, and the amino acid sequence of the TMD-containing fragment of p75NTR. C) SDS-PAGE analysis of cells before (-) and after (+) induction by IPTG, showing noticeably more TrpLE-TMD fusion protein (kDa 18.8 band) after induction. D) HPLC chromatogram (Zorbax SB-C3 column) showing the elution peaks of TrpLE, fusion protein (TrpLE-TMD), and TMD. E). 2D ^1H - ^{15}N TROSY-HSQC spectrum of p75NTR TMD reconstituted in DMPC-DHPC bicelles, recorded at 303 kelvin and ^1H frequency of 600 MHz.

Conclusion

We have shown that by using targeted expression into inclusion bodies in bacteria along with highly optimized affinity and high-pressure reverse-phase lipid chromatography, the hydrophobic TMDs of neurotrophin receptors can be produced to high purity. The peptides can also be uniformly isotopically labeled for NMR measurements. In the protocol, we reconstitute the peptides into bicelles that are sufficiently large to mimic the membrane, because the small

TM domains often require a lipid bilayer environment to form stable oligomeric complexes. We found that when using bicelles with molar ratio of lipid/detergent (q) greater than or equal to 0.5 (e.g., $0.5 \leq q \leq 0.6$; the bilayer region size is between 44 and 50 Å), with which the bicelles are very close to lipid discs, the TMD still yield high quality NMR spectra feasible for structural studies and drug screening. Finally, my protocol should be useful for studying TMD oligomerization of other neurotrophin receptors, because for many of these transmembrane membrane receptors the TMD plays essential roles in receptor assembly and possibly in higher order receptor clustering.

Works Cited

- Bugge, K., Lindorff-Larsen, K. & Kragelund, B. B. Understanding single-pass transmembrane receptor signaling from a structural viewpoint-what are we missing? *Febs j* **283**, 4424-4451, doi:10.1111/febs.13793 (2016).
- Lemmon, M. A. & Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **141**, 1117-1134, doi:10.1016/j.cell.2010.06.011 (2010).
- Schlessinger, J. Ligand-induced, receptor-mediated dimerization and activation of EGF receptor. *Cell* **110**, 669-672, doi:10.1016/s0092-8674(02)00966-2 (2002).
- Endres, N. F. *et al.* Conformational coupling across the plasma membrane in activation of the EGF receptor. *Cell* **152**, 543-556, doi:10.1016/j.cell.2012.12.032 (2013).
- Pan, L. *et al.* Higher-Order Clustering of the Transmembrane Anchor of DR5 Drives Signaling. *Cell* **176**, 1477-1489 e1414, doi:10.1016/j.cell.2019.02.001 (2019).
- Cai, T., Lenoir Capello, R., Pi, X., Wu, H. & Chou, J. J. Structural basis of γ chain family receptor sharing at the membrane level. *Science* **381**, 569-576, doi:10.1126/science.add1219 (2023).
- Bargmann, C. I., Hung, M. C. & Weinberg, R. A. Multiple independent activations of the neu oncogene by a point mutation altering the transmembrane domain of p185. *Cell* **45**, 649-657, doi:10.1016/0092-8674(86)90779-8 (1986).
- Ou, S. I. *et al.* HER2 Transmembrane Domain (TMD) Mutations (V659/G660) That Stabilize Homo- and Heterodimerization Are Rare Oncogenic Drivers in Lung Adenocarcinoma That Respond to Afatinib. *J Thorac Oncol* **12**, 446-457, doi:10.1016/j.jtho.2016.11.2224 (2017).
- Onishi, M. *et al.* Identification of an oncogenic form of the thrombopoietin receptor MPL using retrovirus-mediated gene transfer. *Blood* **88**, 1399-1406 (1996).
- Ding, J. *et al.* Familial essential thrombocythemia associated with a dominant-positive activating mutation of the c-MPL gene, which encodes for the receptor for thrombopoietin. *Blood* **103**, 4198-4200, doi:10.1182/blood-2003-10-3471 (2004).
- Matthews, E. E. *et al.* Thrombopoietin receptor activation: transmembrane helix dimerization, rotation, and allosteric modulation. *FASEB J* **25**, 2234-2244, doi:10.1096/fj.10-178673 (2011).
- Moliner, R. *et al.* Psychedelics promote plasticity by directly binding to BDNF receptor TrkB. *Nat Neurosci* **26**, 1032-1041, doi:10.1038/s41593-023-01316-5 (2023).
- Casarotto, P. C. *et al.* Antidepressant drugs act by directly binding to TRKB neurotrophin receptors. *Cell* **184**, 1299-1313.e1219, doi:10.1016/j.cell.2021.01.034 (2021).
- Huang, Y. *et al.* A molecular mechanism for the generation of ligand-dependent differential outputs by the epidermal growth factor receptor. *Elife* **10**, doi:10.7554/eLife.73218 (2021).
- Battesti, A. & Bouveret, E. The bacterial two-hybrid system based on adenylate cyclase reconstitution in *Escherichia coli*. *Methods* **58**, 325-334, doi:10.1016/j.ymeth.2012.07.018 (2012).

- Caffrey, M. Membrane protein crystallization. *J Struct Biol* **142**, 108-132, doi:10.1016/s1047-8477(03)00043-1 (2003).
- MacKenzie, K. R., Prestegard, J. H. & Engelman, D. M. A transmembrane helix dimer: structure and implications. *Science* **276**, 131-133 (1997).
- Call, M. E. *et al.* The structure of the zeta/zeta transmembrane dimer reveals features essential for its assembly with the T cell receptor. *Cell* **127**, 355-368, doi:10.1016/j.cell.2006.08.044 (2006).
- Bocharov, E. V. *et al.* Spatial structure of the dimeric transmembrane domain of the growth factor receptor ErbB2 presumably corresponding to the receptor active state. *The Journal of biological chemistry* **283**, 6950-6956, doi:M709202200 [pii] 10.1074/jbc.M709202200 (2008).
- Lau, T. L., Kim, C., Ginsberg, M. H. & Ulmer, T. S. The structure of the integrin alphaIIb beta3 transmembrane complex explains integrin transmembrane signalling. *EMBO J* **28**, 1351-1361, doi:10.1038/emboj.2009.63 (2009).
- Barrett, P. J. *et al.* The amyloid precursor protein has a flexible transmembrane domain and binds cholesterol. *Science* **336**, 1168-1171, doi:336/6085/1168 [pii] 10.1126/science.1219988 (2012).
- Fu, Q. *et al.* Structural Basis and Functional Role of Intramembrane Trimerization of the Fas/CD95 Death Receptor. *Mol Cell* **61**, 602-613, doi:10.1016/j.molcel.2016.01.009 (2016).
- Dev, J. *et al.* Structural basis for membrane anchoring of HIV-1 envelope spike. *Science* **353**, 172-175, doi:10.1126/science.aaf7066 (2016).
- Chen, W. *et al.* Familial Alzheimer's mutations within APPTM increase A β 42 production by enhancing accessibility of epsilon-cleavage site. *Nat Commun* **5**, 3037, doi:10.1038/ncomms4037 (2014).
- Call, M. E., Wucherpennig, K. W. & Chou, J. J. The structural basis for intramembrane assembly of an activating immunoreceptor complex. *Nature immunology* **11**, 1023-1029, doi:10.1038/ni.1943 (2010).
- Lee, J. *et al.* Structure of the Ebola virus envelope protein M^{PER}/TM domain and its interaction with the fusion loop explains their fusion activity. *Proc Natl Acad Sci U S A* **114**, E7987-E7996, doi:10.1073/pnas.1708052114 (2017).
- Fu, Q., Piai, A., Chen, W., Xia, K. & Chou, J. J. Structure determination protocol for transmembrane domain oligomers. *Nature protocols* **14**, 2483-2520, doi:10.1038/s41596-019-0188-9 (2019).
- Hempstead, B. L. The many faces of p75^{NTR}. *Curr Opin Neurobiol* **12**, 260-267, doi:10.1016/s0959-4388(02)00321-5 (2002).
- Blacklow, S. C. & Kim, P. S. Protein folding and calcium binding defects arising from familial hypercholesterolemia mutations of the LDL receptor. *Nat Struct Biol* **3**, 758-762 (1996).
- North, C. L. & Blacklow, S. C. Solution structure of the sixth LDL-A module of the LDL receptor. *Biochemistry* **39**, 2564-2571 (2000).

- Sanders, C. R., Hare, B. J., Howard, K. P. & Prestegard, J. H. Magnetically-Oriented Phospholipid Micelles As A Tool For The Study Of Membrane-Associated Molecules. *Progress In Nuclear Magnetic Resonance Spectroscopy* **26**, 421-444 (1994).
- Glover, K. J. *et al.* Structural evaluation of phospholipid bicelles for solution-state studies of membrane-associated biomolecules. *Biophysical journal* **81**, 2163-2171 (2001).
- Piai, A., Fu, Q., Dev, J. & Chou, J. J. Optimal Bicelle Size q for Solution NMR Studies of the Protein Transmembrane Partition. *Chemistry* **23**, 1361-1367, doi:10.1002/chem.201604206 (2017).
- Caldwell, T. A. *et al.* Low- q Bicelles Are Mixed Micelles. *J Phys Chem Lett* **9**, 4469-4473, doi:10.1021/acs.jpcelett.8b02079 (2018).

The Neurological Effects Of Music And How It Can Be Used To Improve Patient Well-Being By Narayan Nair

Abstract

Research has shown that music can be used to treat a variety of issues, ranging from psychological health to chronic pain. This paper will investigate the effects of music on the human body, with a particular focus on its influence on pain, both physical and emotional. Some questions that will be explored in this paper include what the effects of music on pain are, if the emotions conveyed by a song enhance or diminish the impact of music, and if music has a similar influence in the long term as it does in the short term??. The research approach involved doing a simple literature review. In addition, a survey study will be conducted to assess the relationship between music and emotions. Participants will be asked about their current music taste, and what emotion they think the songs convey. This survey study will analyze whether certain songs have the ability to change one's emotional state or elicit certain emotions. This research is important because of its potential implications on the medical field. If music can positively influence a patient's mood, it may lead to better cooperation with the doctor's instructions, leading to a smoother experience for both patients and healthcare providers.

Keywords Behavior and social sciences; neuroscience; medicine; music; music therapy

Introduction

Music has been used for therapeutic purposes since the 1800s, with one of the first references to music therapy being Bruno Goergen's incorporation of music in psychological fields¹. Ever since then, even more research has been done in several aspects, including its potential impacts on the medical field. However, one area with a lackluster amount of research is how music's effect on emotions may play a role in helping patients manage pain. In fact, the topic of music cognition was seen as something not worthy of researching by Steven Pinker, and that music cognition is simply a result of evolution and adaptation for human language. While music has been tested to relieve pain, the exact method of how this occurs is largely unknown, and its use in medical fields has not been adequately tested². The correlation between the emotional response to music and the effectiveness of music as therapy has not yet been tested. The objective of this research paper is to explore the potential benefits music could have on physical and psychological health, with extra emphasis put on pain management. If music can help in regulating emotions, then music has the potential to help regulate pain, even in the long term, and the emotions that music tries to convey can affect the result. If this hypothesis is correct, then it could increase the usefulness of music therapy in medical fields. If playing music can improve patients' mood, they are more likely to be cooperative, improving the experience for both healthcare providers and patients alike. Additionally, the music may provide a calming effect for the patient, which may allow them to be more willing to undergo more painful procedures.

History of Music

Music has been composed since the dawn of humanity, with some of the very first references being found 40,000 years ago, and evidence to suggest that it has been around for longer³. As such, music acts as a universal language, and allows people from multiple cultures to connect. However, it took several thousand years for music therapy to be a suggested idea in academia. The pioneers of music therapy were Edwin Atlee, and Samuel Mathews, who were the first to publish dissertations on the potential medical uses of music⁴. Shortly after, the first use of music therapy in a professional setting was recorded in Blackwell's Island, New York, as well as the first systematic experiment concerning music theory⁴. Music therapy continued to be a subject of interest up to and past the 1900s, where several institutions based on music therapy would be founded and subsequently disbanded after only a few years, including the National Society for Music Therapists, and the National Foundation for Music Therapy⁵. Then, in the 1940's Michigan State University was the first to establish music therapy as an academic program⁵. The University of Kansas, Chicago Musical College, Alverno College, and College of the Pacific would follow MSU's lead⁵. Since then, much research has been conducted on how music affects patient's brains, and it is possible to supplement the research of the past with neurological knowledge from the present.

The Neurological Factors of Music

In order to understand how to use music in assisting with medical treatments, it is imperative to understand how music affects the brain. Otherwise, certain untrue assumptions may be made, and attempting to expand research done with those assumptions may yield incorrect information. The brain first processes the audio in the auditory nuclei, after which it derives meaning from the sound in the primary auditory cortex, and temporal lobe⁶. The temporal lobe breaks down what emotion is derived from the music, and its connection to memory creates associations with the song and certain parts of a person's life, leading to an even stronger emotional connection to the music⁶. There are eight attributes that can be varied in music otherwise known as pitch, rhythm, timbre, tempo, meter, contour, loudness, and spatial location. Different aspects of music often cause different associations in a person's brain thanks to the temporal lobe⁶. Higher, shrill sounds are more likely to be perceived as annoying, while lower pitched sounds are more calming for example⁶. In contrast, soothing music, often characterized by lower pitch, slower tempo is associated with higher levels of relaxation⁶. This research indicates another possible use for music on top of soothing patients and allowing them to undergo painful procedures: stress management.

Music Therapy

Music therapy is a type of therapy that uses music to improve people's well being. This section will discuss music's effects on stress, anxiety, pain, and sleep. Specifically, it will focus on how music can be used to soothe patients emotionally, how it affects the brain, and other potential applications of music therapy.

Effects of Music on Stress

Stress is known to negatively impact a person's health, including physically, emotionally, and mentally. Research can help examine the positive benefits of music on different forms of stress. Emotional stress, what this paper intends to focus on, will most likely follow this pattern and be positively affected by soothing music. The brain is the emotional center of the human body and where stress originates. It naturally follows that if music is able to influence stress, then it would do this by influencing certain brain structures. Music has been shown to engage cognitive functions, increase attention span, concentration, imagination, creativity and more⁷. It does so by having an effect on almost all areas of the brain. Playing music has been shown to affect most parts of the brain, including the cerebellum, the frontal lobe, and even the amygdala, and listening to music has a similar effect⁶. Posttraumatic stress disorder (PTSD) is a condition that is often caused by a highly traumatic memory in a person's life, and is identified by avoidance, an inability to forget certain distressing memories, and increased vigilance⁸. While more research still needs to be conducted, music therapy has been linked with improving function in people with PTSD, and helping those who suffer from it reintegrate into society⁸. Although music therapy has not been tested enough to come to a definite conclusion regarding its effect on stress, connections can be drawn by learning more about the neuroscience of music. It has already been shown that music has a positive effect on people with PTSD, because of the way music affects the brain, especially the amygdala, the emotional center of the brain. This implies music could produce similar results on patients who are suffering with excessive stress that is not trauma-related. Overall, music has the potential to positively affect those who suffer from different types of stress.

Effects of Music on Anxiety

Music does not have to be exclusively used to assist in stress management, as it may also be effective for dealing with anxiety. However, as with stress, it has yet to be thoroughly researched. This section discusses how music could be used to treat anxiety alongside stress. Considering how the two terms are often linked, this research should yield similar results to the stress section, and the same research questions for stress can apply here. Music's effect on emotion also comes into play here. Due to how music influences the amygdala, patients with anxiety experience similar effects to those who suffer with stress⁶. Music also had positive effects on patients with schizophrenia, who had improved symptom scores after multiple music therapy sessions⁹. Additionally, music has been linked with better moods for people with depression and anxiety¹⁰.

This research implies that music has a similar effect on anxiety and stress, however, more research needs to be conducted to further cement the relationship between the two. The way music affects the amygdala is key to understanding the way music contributes to mitigating these disorders. With the research that has been currently conducted, music therapy has been shown to have at least a mildly positive effect on those suffering from emotional disorders. Such as schizophrenia, depression, anxiety etc.

Effects of Music on Pain

As stated before, music has been shown to have a significant effect on pain, especially when it is soothing to listen to. The research shows a strong link between listening to music, and a reduction in pain. However, music as a treatment for pain can be explored further, and doing so will improve our understanding of this topic.

Organized music therapy has been correlated with numerous health benefits, including lower pain, heart rate, emotional distress, and even opioid intake¹¹. This is most likely due to it affecting much of the brain, including deactivating the amygdala, which helps regulate emotions⁶. Improving the amygdala's ability to control emotional responses makes it less likely for people to engage in self destructive behavior, and allows them to handle higher pain thresholds. This research may prove useful in medical treatments. Patients have been shown to have more positive attitudes towards medical treatments that are less painful. Hence, since music has already been linked to lower pain levels, music should make people's perception of medical procedures more favorable¹². By having a more positive perception of medical treatments, people might be more likely to cooperate with doctors, providing a smoother experience for all parties involved.

By affecting the amygdala, music helps control the emotional response elicited by pain. This treatment reduces the unpleasant feeling for patients and makes them more cooperative with medical procedures. Not to mention the numerous other health benefits that music provides.

Effects of Music on Sleep Quality

Sleep quality is one of music therapy's many health benefits. It especially deserves to be studied further because of its implications for people's health. Soothing music, as with other types, affects most of the brain, including the hypothalamus, the part of the brain responsible for sleep⁶. By stimulating the hypothalamus, music may induce drowsiness in a person, prompting them to fall asleep easier. This hypothesis has been tested, and results have suggested a positive correlation between application of music therapy, and sleep quality¹³. Not only does the research suggest an improvement in general sleep quality, but also an improvement in those who struggle with chronic sleep disorders¹³. This study implies significant things for sleep improvement interventions. Music, by nature, is very accessible thanks to the advent of sites like Youtube and Spotify, and is additionally extremely unlikely to harm the listener. As such, informal music therapy used as a treatment for insomnia is an incredibly low-risk option for those who need a more accessible way to treat sleep conditions. Music as a treatment can be used to great effect to improve sleep quality. Even chronic conditions can be somewhat alleviated by music therapy. Additionally, informal music therapy's accessibility makes it a great tool for those who cannot access a doctor to get a formal diagnosis for a condition.

Effects of Combining Music with Mindfulness

More soothing music is not just useful for improving sleep quality, it also has potential uses for mindfulness. Mindfulness has additional health benefits associated with it, which

compound with those that are provided by music¹⁴. Mindfulness is the ability to be in the present moment, without holding on to thoughts from the past, or worries for the future¹⁴. Often, mindfulness works by stimulating the brain in order to make it attuned to the present moment, which combines well with music, which also involves stimulation. Mindfulness can have positive effects on the brain, including the deactivation of the amygdala, which can help to regulate emotions¹⁵. This effect is similar to the effects of soothing music, making them combine well. One technique commonly used in conjunction with music is meditation, where music is used in order to create a feeling of peace inside the meditator¹⁴.

Music and Emotions

Music can be used for more than just creating peace, as it also has different effects on people based on the genre of the song. Specific genres are certainly likely to correspond to specific emotions due to how they often vary in tempo, pitch, and especially the instruments used. This section intends to find what emotions correspond to which genres, how personality and physiological characteristics may play a part in how music is perceived, and how this could be used to benefit humanity as a whole. Music genres have been heavily linked with certain emotions. Jazz music, for example, is associated with nostalgia, while reggae is more linked to relaxation¹⁶. This makes sense considering music's link with the amygdala¹⁵. Additionally, a person's past experiences may have affected their perception of music, especially in the case of jazz, as nostalgia is often an emotion driven by memories. Music's effect on the brain can be recorded in physical effects on the body as well. Music has been shown to be able to increase heart rate, skin temperature, and affect the endocrine system¹⁷. Not only that, but other research shows that music can affect one's perception of themselves. When asked to evaluate their own characteristics, people listening to happier songs evaluated themselves more positively than those who were listening to sad music¹⁸.

Research Survey

Along with conducting a review on the connections made between music and emotions, I will also be doing my own research. Little research has yet been done into how people perceive and understand music, and I intend to reinforce the research that has already been done. My study will hopefully provide some more support to other studies similar to it in nature, and will reinforce the current knowledge about music theory. Due to personal budget constraints, the experiment will have to be more like a survey instead of a more formal experiment where I can manipulate the variables.

Methods

A public survey will be created and posted on social media sites via Google Forms . The survey will ask the participant to complete a quick questionnaire to detail their current mood, after which they will be asked to listen to 4 different songs that are meant to incite anger, happiness, sadness, and anxiety. After each song, the participant will be asked to fill out the same

questionnaire to determine if the music actually had a significant impact on their mood. A total of 10 participants completed the survey. Due to how the survey was conducted, many of them came from the United States of America, and were of adolescent age, primarily in high school.

Song 1: Morrakka

Link: <https://www.youtube.com/watch?v=ielNlsacdk8>

Song 2: Ussewa

Link: <https://www.youtube.com/watch?v=Qp3b-RXtz4w>

Song 3: What was I made for?

Link: <https://www.youtube.com/watch?v=cW8VLC9nnTo>

Song 4: Super Battleworn Insomniac

Link: https://www.youtube.com/watch?v=jRx_in09nyM

Data Analysis

To calculate the average for each emotion, ratings were recorded after each song. Afterwards, baseline emotions were rated again and served as the baseline score for that emotion for the next song.

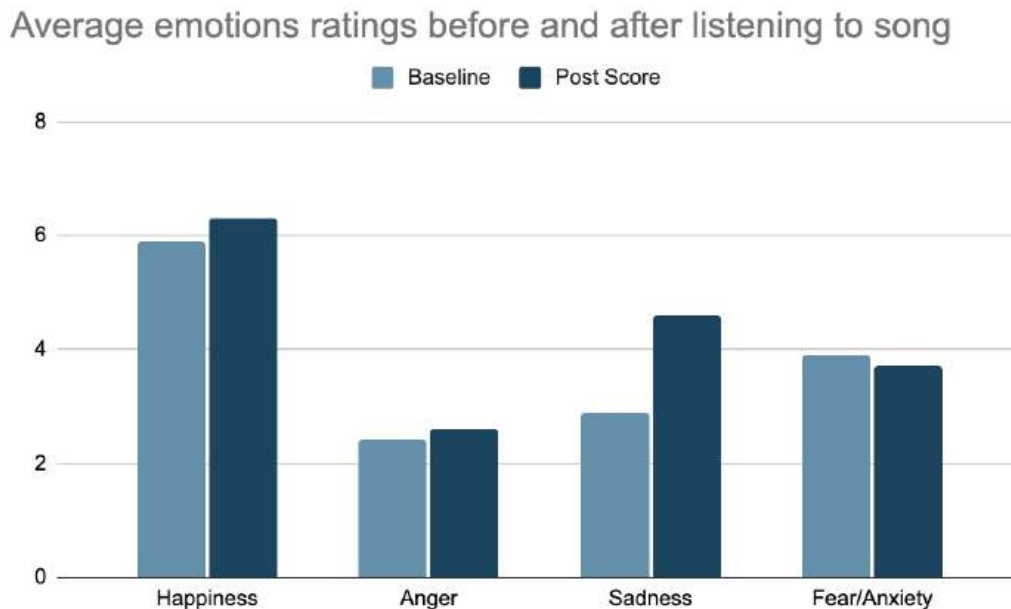


Figure 1, The results of the experiment.

Results

The average baseline happiness score was 5.90, and the happiness score of the participants was 6.30 with a SD of 1.84 after listening to the happy song, showing an average increase of 0.40. For anger levels, the baseline score was 2.40, and it only slightly increased after listening to the song meant to invoke rage, at 2.60 and an SD of 1.43. This implied that the song did not cause a big change in emotional expression of the participants, and the reaction that some had seeing a song that they were familiar with caused anger to decrease. However, the initial anger scores were already considerably lower compared to all other emotions studied, making it harder to assess major possible decreases. Sadness levels were relatively variable before the experiment, with a baseline score of 2.9. After listening to the sad song, there was an increase in sadness, with an average score of 4.6 and an SD of 2.24. These results are to be expected, naturally, sad songs would cause one to feel sadder. Finally, baseline fear/anxiety levels were low at 3.9, but slightly decreased after listening to the song meant to evoke it, at 3.7, with an SD of 2.19. This is most likely due to the fact that fear is hard to encapsulate in a song. None of the participants claimed to know the song, so it leaves that as the only explanation.

Discussion

The objective of this research was to explore the potential benefits music could have on physical and psychological health. What the literature shows is that music is a cost-effective, versatile alternative to what can often be costly medicine, linked to stress and anxiety reduction, as well as better sleep quality and pain management. However, these are merely links, and more research has to be done to figure out how strong they actually are. My research attempted to fill this gap, however due to unforeseen circumstances, there were significantly less responses than expected, so while there is a mild link between what it was trying to measure, the relationship between music and emotions, there were nowhere near enough responses to prove anything. Due to the demographics of the sample population, some considerations must be acknowledged. There is potential for bias since the sample came from a wealthy population which skewed towards adolescent age ranges. This could have impacted the results because music tastes at that age are naturally much different than those in older or younger generations, which could impact results.

Music therapy is a type of therapy that uses music to improve people's well being. Although music therapy has not been tested enough to come to a definite conclusion in its effect on stress, connections can be drawn by learning more about the neuroscience of music. It has already been shown that music has a positive effect on people with PTSD, because of the way music affects the brain, especially the amygdala, the emotional center of the brain⁶. This implies music could produce similar results on patients who are suffering with excessive stress that is not trauma-related. Overall, music has the potential to positively affect those who suffer from different types of stress. By affecting the amygdala, music helps control the emotional response that pain elicits. This treatment reduces the unpleasant feeling for patients and makes them more cooperative with medical procedures. Not to mention the numerous other health benefits that music provides.

Soothing music was shown to induce drowsiness and induce sleep by affecting the hypothalamus, the part of the brain responsible for controlling sleep, even help with insomnia¹³. Music therapy as a treatment for sleep is an excellent homeopathic way to improve sleep quality. It is a growing area in the fields of psychology and mental health, meditation by combining music and mindfulness can increase positive feelings by activating the amygdala. Music has been linked to changes in mood and physiology. People who listened to happier songs also rated their own attribute higher than those who listened to sad songs. As for mood, it is much more dependent on personal tastes, and people's perspective on the type of music will heavily impact people's reactions to it.

Limitations

Due to the nature of the survey, I was only able to ask for about 15 minutes worth of time from people, and the severe lack of responses suggests that even this was too much. If I had some more funding, I may have been able to provide economic incentive to participate, or create a longer, more thought out experiment, however my personal financial restraints were limiting in that regard. Additionally, I did not have the resources necessary to create a more exhaustive study, and do not currently possess the ability to acquire any. This idea, while extremely interesting, needs more research to be proven, and my study was not the best way to measure it. While the study was cost effective, and easy for people to respond to, it did not have as big of an outreach as I had hoped, and that proved to be quite the sticking point. Additionally, the places where I had posted it, mainly through social media, mainly appeal to teenagers, which caused an increase in teenage responses compared to other age groups, and those teenagers also happened to know me to some level, which may have affected how they respond.

Summary of Findings

However, even with the lack of data, some conclusions can be pulled. When listening to happier songs, there was a general increase in happiness levels and a decrease in sadness levels. Sad songs cause a decrease in fear and anxiety, and a small increase in sadness levels. Songs meant to be scary did not tend to invoke many substantial emotions, and angry songs weren't effective in capturing much anger in people who were not already angry. What is arguably the most interesting finding of this survey however, is that when people reported listening to a song before, their happiness levels were higher after listening to it again. This oddity can most likely be explained as familiarity. When people are more familiar with a song, they have a more positive reaction to it. This in particular is very interesting, as it means that by playing more familiar music to patients for example, it may calm them, and put them in a more positive state.

Directions for Future Research

Based on this study, I would suggest future research to look more into how familiarity might affect the influence music has on people. There is a big gap that would be extremely interesting and valuable to fill based on this study. Along with that, having more financial

resources and outreach would also help fix some biases, and potentially make this study more organized.

Conclusion

Besides my study, other facets of music have also been explored. It has been linked to a decrease in pain both physical and emotional, as well as an improvement in anxiety, sleep quality, and linked with emotions. This information is very useful due to how they can be used in the medical field. By using emotionally soothing music, patients would be more likely to comply with medical instruction, and feel less pain during more intensive procedures. While there is still much to explore in the field of music, it has proven to have some use in both medical and more informal settings.

Acknowledgments

I would like to thank Ana Corra for helping me with writing this research paper, as well as Kathryn Wall for helping me polish and market the paper. I would also like to thank the higher ups at Polygence for matching with these people, and for helping me figure out that I wanted to write a research paper for this project. I could not have done any of this without the assistance of the

Works Cited

- 1: Korenjak A. From moral treatment to modern music therapy: On the history of music therapy in Vienna (c. 1820-1960). *Nord J Music Ther.* 2018 May 29;27(5):341-359. doi: 10.1080/08098131.2018.1467481. PMID: 30393512; PMCID: PMC6199691.
- 2: Bernatzky G, Presch M, Anderson M, Panksepp J. Emotional foundations of music as a non-pharmacological pain management tool in modern medicine. *Neurosci Biobehav Rev.* 2011 Oct;35(9):1989-99. doi: 10.1016/j.neubiorev.2011.06.005. Epub 2011 Jun 16. PMID: 21704068.
- 3: Killin A. The origins of music: Evidence, theory, and prospects - Anton Killin, ... [Internet]. 2018 [cited 2024 Apr 3]. Available from: <https://journals.sagepub.com/doi/full/10.1177/2059204317751971>
- 4: Davis WB. Music therapy in 19th Century America [Internet]. Oxford University Press; 1987 [cited 2024 Apr 3]. Available from: <https://academic.oup.com/jmt/article-abstract/24/2/76/1175691>
- 5: Taylor DB. Music in general hospital treatment from 1900 to 1950. *J Music Ther.* 1981 Summer;18(2):62-73. doi: 10.1093/jmt/18.2.62. PMID: 10298285.
- 6: Warren J. How does the brain process music? *Clin Med (Lond).* 2008 Feb;8(1):32-6. doi: 10.7861/clinmedicine.8-1-32. PMID: 18335666; PMCID: PMC4953706.
- 7: Speranza L, Pulcrano S, Perrone-Capano C, di Porzio U, Volpicelli F. Music affects functional brain connectivity and is effective in the treatment of neurological disorders. *Rev Neurosci.* 2022 Mar 24;33(7):789-801. doi: 10.1515/revneuro-2021-0135. PMID: 35325516.
- 8: Landis-Shack N, Heinz AJ, Bonn-Miller MO. Music Therapy for Posttraumatic Stress in Adults: A Theoretical Review. *Psychomusicology.* 2017;27(4):334-342. doi: 10.1037/pmu0000192. PMID: 29290641; PMCID: PMC5744879.
- 9: Talwar N, Crawford MJ, Maratos A, Nur U, McDermott O, Procter S. Music therapy for in-patients with schizophrenia: exploratory randomised controlled trial. *Br J Psychiatry.* 2006 Nov;189:405-9. doi: 10.1192/bjp.bp.105.015073. PMID: 17077429.
- 10: Aalbers S, Fusar-Poli L, Freeman RE, Spreen M, Ket JC, Vink AC, Maratos A, Crawford M, Chen XJ, Gold C. Music therapy for depression. *Cochrane Database Syst Rev.* 2017 Nov 16;11(11):CD004517. doi: 10.1002/14651858.CD004517.pub3. PMID: 29144545; PMCID: PMC6486188.
- 11: Lee JH. The Effects of Music on Pain: A Meta-Analysis. *J Music Ther.* 2016 Winter;53(4):430-477. doi: 10.1093/jmt/thw012. Epub 2016 Oct 19. Erratum in: *J Music Ther.* 2021 Aug 24;58(3):372. PMID: 27760797.
- 12: Abu Alhajja ES, Aldaikki A, Al-Omairi MK, Al-Khateeb SN. The relationship between personality traits, pain perception and attitude toward orthodontic treatment. *Angle Orthod.* 2010 Nov;80(6):1141-9. doi: 10.2319/012710-59.1. PMID: 20677967; PMCID: PMC8929492.

- 13: Wang CF, Sun YL, Zang HX. Music therapy improves sleep quality in acute and chronic sleep disorders: a meta-analysis of 10 randomized studies. *Int J Nurs Stud*. 2014 Jan;51(1):51-62. doi: 10.1016/j.ijnurstu.2013.03.008. Epub 2013 Apr 9. PMID: 23582682.
- 14: Hernandez-Ruiz E, Dvorak A. Comparison of music stimuli to support mindfulness meditation - ... [Internet]. 2020 [cited 2024 Apr 3]. Available from: <https://journals.sagepub.com/doi/10.1177/0305735619878497>
- 15: Wheeler MS, Arnkoff DB, Glass CR. The Neuroscience of Mindfulness: How Mindfulness alters the brain and facilitates emotion regulation - mindfulness [Internet]. Springer US; 2017 [cited 2024 Apr 3]. Available from: <https://link.springer.com/article/10.1007/s12671-017-0742-x>
- 16: Verling E. Investigating the impact of different music genres on emotions and motivation [Internet]. 2021 [cited 2024 Apr 3]. Available from: <https://norma.ncirl.ie/id/eprint/4876>
- 17: Särkämö T, Tervaniemi M, Huotilainen M. Music perception and cognition: development, neural basis, and rehabilitative use of music. *Wiley Interdiscip Rev Cogn Sci*. 2013 Jul;4(4):441-451. doi: 10.1002/wcs.1237. Epub 2013 Mar 20. PMID: 26304229.
- 18: Scherer K., Zentner M. Emotional effects of music: Production rules | semantic scholar [Internet]. 2001 [cited 2024 Apr 3]. Available from: <https://www.semanticscholar.org/paper/EMOTIONAL-EFFECTS-OF-MUSIC:-PRODUCTION-RULES-Scherer-Zentner/9a1768e05ba383bd8589e1378ced4e9625762f1e>

Technologies to Combat Ocean Acidification By Diya Sammanna

Abstract

As humans continue to burn more fossil fuels and increase the amount of atmospheric carbon dioxide, the pH of the ocean is expected to continue decreasing. Ocean acidification comes with consequences: marine organisms are at a threat of losing their shells, fish are struggling to survive, and human nutritional health and food sources are at risk. Bipolar Membrane Electrodialysis (BMED) is hypothesized to become a more common technique in increasing the alkalinity of the ocean. This review starts with the chemistry and effects of ocean acidification and introduces three different technologies to increase the pH of the ocean: BMED, the addition of slaked lime, and increasing the growth of algae. Then, the most ideal method and the future of ocean acidification and technologies are hypothesized.

Introduction

Ocean acidification is the process in which the pH of the ocean decreases, hence, leading to a more acidic ocean. Because of the negative logarithmic scale used to measure pH, a small change in pH has a large impact: reducing the pH by 1 unit leads to 10 times more acidity (Guinotte and Fabry, 2008). The ocean's pH has decreased by 25% today than before the industrial revolution, changing from a pH of 8.2 to 8.1 (Longphuir et al., 2010).

Ocean acidification is caused by the increased amount of atmospheric carbon dioxide. The carbon dioxide from the atmosphere diffuses into the ocean, which increases the acidity of the ocean (Prentice et al., 2001). The major factors of the increased carbon dioxide are burning fossil fuels like coal and gas and deforestation (House et al., 2002). For example, transportation such as traveling in planes and cars burn a significant amount of fossil fuels.

When carbon dioxide from the atmosphere diffuses into the ocean, the carbon dioxide reacts with the ocean water (H_2O). The product of this reaction is carbonic acid (H_2CO_3). The carbonic acid dissociates into a bicarbonate ion (HCO_3^-) and a hydrogen ion (H^+). A bicarbonate ion can then dissociate into two hydrogen ions and a carbonate ion (CO_3^{2-}). The set of these reactions can move both in the forwards and reverse direction (Archer, 1999; Moon et al., 2019).

Specifically, the proportion of the reactants or products determines whether the reaction will move in the forwards direction, to synthesize more products, or in the reverse direction, to synthesize more reactants (Moon et al., 2019).

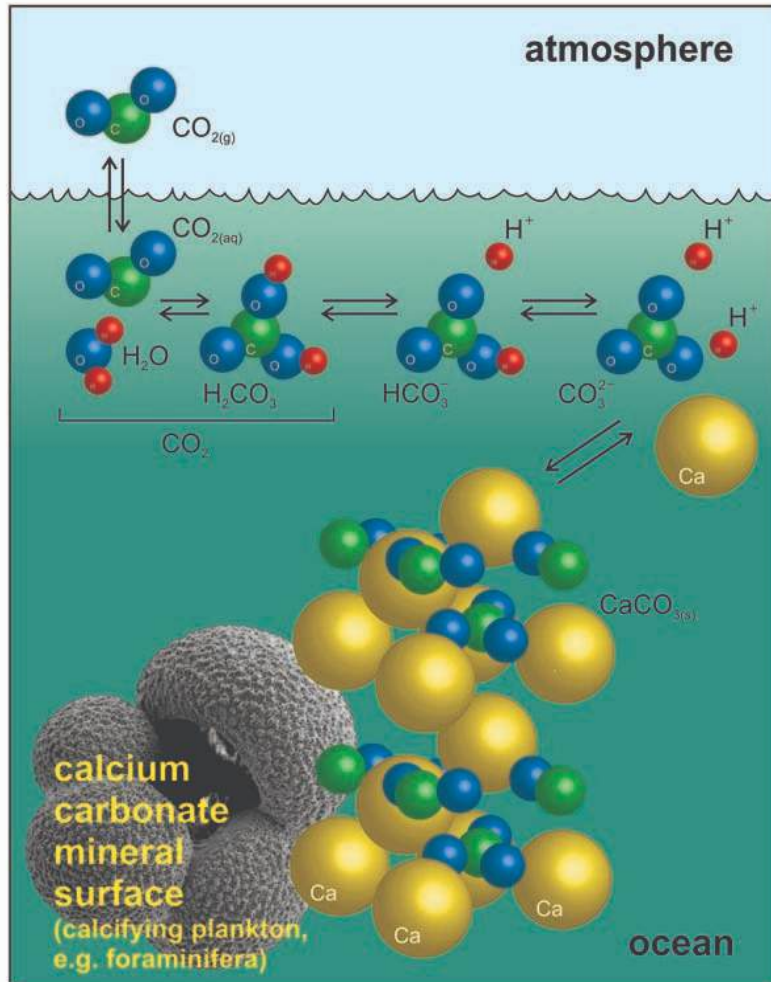


Fig 1: The reactions and dissociations of carbon dioxide in the ocean (Barker et al., 2012)

Like calcium ions (Ca^{2+}), hydrogen ions are attracted to carbonate ions to form bicarbonate ions (Jensen et al., 2018). However, hydrogen ions have a stronger pull towards carbonate ions compared to calcium ions. Therefore, this leads to the production of more bicarbonate ions, and a reduction to the amount of carbonate ions and calcium carbonate. This has a devastating effect on marine species that rely on carbonate ions and calcium ions to build their calcium carbonate (CaCO_3) shells (Gazeau et al., 2013). Some organisms, like coral, have shells that are directly exposed to the ocean which are at risk of dissolving (Buth, 2016). Other organisms, such as mollusks, have tissue buffers between their shells and the ocean; however, these organisms have to use more energy to maintain their shells which reduces the amount of energy available for metabolic processes (Eberlee and Story, 1984; Sokolova et al., 2012).

Ocean acidification can even pose a threat to humans: fish are a stable source of nutrition for humans (Guinotte and Fabry, 2008). In fact, the average amount of fish consumed each day in the United States was 15.65 grams of fish per person (Jacobs et al., 2006). However, because of ocean acidification and a higher amount of carbon dioxide in the ocean, fish are struggling in

the early stages of their life and have fewer amounts of key proteins such as omega-3 fatty acid (Guinotte and Fabry, 2008).

This review outlines the novel technologies that can reduce the effects of ocean acidification and the future for ocean acidification technological research.

Bipolar Membrane Electrodialysis (BMED)

Bipolar Membrane Electrodialysis is a technology that is used to trap carbon dioxide from the ocean (Khoiruddin et al., 2024). An anode (positively charged) and a cathode (negatively charged) are added on either side of the ocean water (Eisaman et al., 2012). From the cathode to the anode, there is an alternation between negatively charged cation-exchange layers (CEL) and positively charged anion-exchange layers (AEL) respectively (Titorova et al., 2020). An electrical current is applied to the cathode and the anode, which causes water to split into hydrogen ions and hydroxide ions (OH⁻): one molecule of water dissociates into one hydrogen ion and one hydroxide ion (Tee et al., 2017). The process of water splitting creates a pH gradient which allows negatively charged molecules to move through the AELs and the positively charged molecules to move through the CELs. Ocean water is passed in between one of the AELs and one of the CELs. Because bicarbonate ions are negatively charged, bicarbonate ions are attracted to the positively charged electrical current coming from the anode. As bicarbonate ions move through the AEL, bicarbonate ions are trapped with the CEL (Khoiruddin et al., 2024). BMED has enough acidity to form carbon dioxide dissolved in water from the bicarbonate (Aliaskari et al., 2024; Khoiruddin et al., 2024). Therefore, carbon dioxide can be examined further through for example, dissociation reactions (Khoiruddin et al., 2024).

BMED is used in direct ocean capture (DOC) to remove carbon dioxide from the air and direct air capture (DAC) to remove carbon dioxide from the ocean, and therefore allows for carbon dioxide from the atmosphere to diffuse in the ocean (Khoiruddin et al., 2024).

BMED has the potential to be converted to a large-scale, is rather environmentally safe, and traps carbon dioxide at a fast rate. However, the bipolar membranes (BPMs) used in BMED are expensive and BMED uses a lot of energy (Khoiruddin et al., 2024). Additionally, the energy used to supply BMED needs to come from renewable resources such as wind and solar energy (Thomaidis et al., 2016).

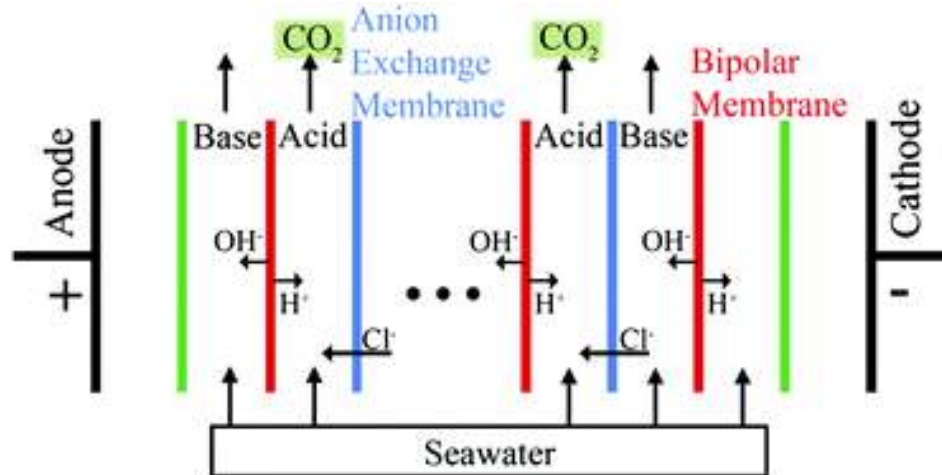


Fig 2: The processes and components of BMED (Eisaman et al., 2012)

Geoengineering: Adding Slaked Lime to the Oceans

Slaked lime has a chemical formula of $\text{Ca}(\text{OH})_2$ (NIH, 2024). Slaked lime reacts with two molecules of carbon dioxide to synthesize one calcium ion and two bicarbonate ions. This reaction uses carbon dioxide, which decreases the concentration of carbon dioxide in the atmosphere: the reaction of water and carbon dioxide moves in the forward direction, allowing more carbon dioxide from the atmosphere to dissolve in the ocean. Additionally, one hydroxide ion from slaked lime can react with one hydrogen ion in the ocean to synthesize water, reducing the concentration of hydrogen ions and hence, increasing the overall pH (Varliero et al., 2024).

Slaked lime fulfills both the need to reduce the amount of carbon dioxide in the water and to reduce the acidity of the ocean (Foteinis and Renforth, 2024). However, slaked lime comes from extraction and mining for limestone, which releases carbon dioxide in the atmosphere (Lackner, 2002). In addition, increasing the pH of the ocean through slaked lime uses a lot of materials and money: 1.4 tons to 1.7 tons of limestone are required to remove one ton of carbon dioxide from the ocean (Renforth et al., 2013).

Bioremediation: Growing Algae

Algae is a producer and therefore, goes through the process of photosynthesis, which reduces the concentration of carbon dioxide and acidity in the ocean. Six molecules of carbon dioxide and six molecules of water react to form one glucose molecule ($\text{C}_6\text{H}_{12}\text{O}_6$) and six oxygen molecules (O_2) (Sáenz et al., 2012). Similar to adding slaked lime, the reaction between water and carbon dioxide moves in the forward direction, allowing more carbon dioxide from the atmosphere to dissolve in the ocean (Moroney and Somanchi, 1999). Currently, the concentration of carbon dioxide in the ocean is increasing at a faster pace than the algae can remove it (Moreira and Pires, 2016). If humans can grow algae and then plant them into ocean zones, the carbon dioxide levels could decrease in these locations (Haoyang, 2018).

But, when there is an overgrowth of algae, the algae use a lot of oxygen and sunlight which negatively affect other marine plants. Furthermore, harmful algal blooms can be triggered from overgrowth: algae synthesize chemicals that are dangerous to other animals and even humans (Fu et al., 2012).

The Best Solution: Bipolar Membrane Electrodialysis (BMED) with Modification

As described earlier, BMED is a costly process, using a high amount of energy and expensive BPMs. However, if wind turbines were used to supply the necessary energy to BMED, this process could become more practical: wind turbines are a renewable and affordable resource with a price of about 1.9 cents per kilowatt-hour (Thresher et al., 2007).

The cost of the BPMs can also be reduced: Do-Hyeong Kim and Moon-Sung Kang from Sangmyung University found that using porous substrate films in BPMs were less costly but still worked well (Kim and Kang, 2017).

Using BMED with renewable resources like wind turbines and cheaper BPM materials such as porous substrate films provides a way to practically apply the principles of BMED to reduce the concentration of carbon dioxide in the ocean.

The Future

Ocean acidification is likely to persist and the ocean is expected to become even more acidic in the future (T. Abbasi and S. A. Abbasi, 2011). This is because there is a higher concentration of carbon dioxide in the atmosphere with the increasing emissions of fossil fuels (House et al., 2002). With a decreased ocean pH, marine organisms, especially with shells and human nutrition are at risk (Guinotte and Fabry, 2008).

To solve these problems, funding is being given to explore how the pH of the ocean can be increased; in 2023, the National Oceanic and Atmospheric (NOAA) Ocean Acidification Program allotted 24.3 million dollars for group efforts to decrease the amount of carbon dioxide in the ocean. From this fund, head researchers from Scripps Institution of Oceanography at University of California, San Diego Andreas Andersson and Andrew Dickson were granted approximately 1.45 million and 1 million respectively. For their approach to minimize the oceanic carbon dioxide, Andreas Andersson researches seaweed farming and Andrew Dickson focuses on the addition of bases. This is the first collaborative and widespread effort in the field of oceanic carbon dioxide removal (Fox, 2023).

As more research is being done in this field, BMED is hypothesized to become more commonly used in the future as more information is learned about the cost and energy saving methods that can be applied to this process (Luo et al., 2022).

Acknowledgements

I would like to thank my mentor, J Fearon, for guiding me throughout this project and the University of California, Santa Cruz Science Internship Program (UCSC SIP) for giving me the resources to carry out this project.

Works Cited

- Abbasi, Tasneem, and S. A. Abbasi. "Ocean acidification: The newest threat to the global environment." *Critical Reviews in Environmental Science and Technology* 41.18 (2011): 1601-1663.
- Aliaskari, Mehran, et al. "A systematic analysis of operating parameters for CO₂ capture from seawater by Bipolar Membrane Electrodialysis (BPMED)." *Separation and Purification Technology* 339 (2024): 126679.
- Archer, David. "Modelling carbon dioxide in the ocean: A review." *Developments in Atmospheric Science* 24 (1999): 171-183.
- Barker, S. & Ridgwell, A. (2012) Ocean Acidification. *Nature Education Knowledge* 3(10):21.
- Eberlee, John C., and Kenneth B. Storey. "Buffering capacities of the tissues of marine molluscs." *Physiological zoology* 57.5 (1984): 567-572.
- Eisaman, Matthew D., et al. "CO₂ extraction from seawater using bipolar membrane electrodialysis." *Energy & Environmental Science* 5.6 (2012): 7346-7352.
- Foteinis, Spyros, and Phil Renforth. "Realistic deployment scenarios/pathways that can be used to constrain Earth System models." (2021).
- Fox, Alex. "NOAA Awards \$2.5 Million to Scripps Researchers Working on Climate Solutions." Scripps Institution of Oceanography, 7 September 2023, <https://scripps.ucsd.edu/news/noaa-awards-25-million-scripps-researchers-working-climate-solutions>.
- Fu, Fei Xue, Avery O. Tatters, and David A. Hutchins. "Global change and the future of harmful algal blooms in the ocean." *Marine Ecology Progress Series* 470 (2012): 207-233.
- Gazeau, Frédéric, et al. "Impacts of ocean acidification on marine shelled molluscs." *Marine biology* 160 (2013): 2207-2245.
- Guinotte, John M., and Victoria J. Fabry. "Ocean acidification and its potential effects on marine ecosystems." *Annals of the New York Academy of Sciences* 1134.1 (2008): 320-342.
- Haoyang, Cai. "Algae-based carbon sequestration." *IOP Conference Series: Earth and Environmental Science*. Vol. 120. No. 1. IOP Publishing, 2018.
- House, Joanna I., I. Colin Prentice, and Corinne Le Quere. "Maximum impacts of future reforestation or deforestation on atmospheric CO₂." *Global Change Biology* 8.11 (2002): 1047-1052.
- Jacobs, Helen L., et al. "Estimates of per capita fish consumption in the US based on the continuing survey of food intake by individuals (CSFII)." *Risk analysis* 18.3 (1998): 283-291.
- Jensen, Anders CS, et al. "Hydrogen bonding in amorphous calcium carbonate and molecular reorientation induced by dehydration." *The Journal of Physical Chemistry C* 122.6 (2018): 3591-3598.
- Khoiruddin, K., I. G. Wenten, and Utjok WR Siagian. "Advancements in Bipolar Membrane Electrodialysis Techniques for Carbon Capture." *Langmuir* 40.18 (2024): 9362-9384.

- Kim, Do-Hyeong, and Moon-Sung Kang. "Cost-effective bipolar membranes for efficient electrochemical water dissociation." *Chemistry Letters* 46.10 (2017): 1459-1462.
- Lackner, Klaus S. "Carbonate chemistry for sequestering fossil carbon." *Annual review of energy and the environment* 27.1 (2002): 193-232.
- Luo, Yu, et al. "Application of bipolar membrane electrodialysis in environmental protection and resource recovery: a review." *Membranes* 12.9 (2022): 829.
- Moon, Alena, et al. "Application and testing of a framework for characterizing the quality of scientific reasoning in chemistry students' writing on ocean acidification." *Chemistry Education Research and Practice* 20.3 (2019): 484-494.
- Moreira, Diana, and José CM Pires. "Atmospheric CO₂ capture by algae: negative carbon dioxide emission path." *Bioresource technology* 215 (2016): 371-379.
- Moroney, James V., and Aravind Somanchi. "How do algae concentrate CO₂ to increase the efficiency of photosynthetic carbon fixation?." *Plant Physiology* 119.1 (1999): 9-16.
- Prentice, Iain Colin, et al. "The carbon cycle and atmospheric carbon dioxide." *Climate change 2001: the scientific basis, Intergovernmental panel on climate change* (2001).
- Renforth, Philip, B. G. Jenkins, and T. Kruger. "Engineering challenges of ocean liming." *Energy* 60 (2013): 442-452.
- Sáenz, María Elena, Walter Darío Di Marzio, and Jose Luis Alberdi. "Assessment of Cyfluthrin commercial formulation on growth, photosynthesis and catalase activity of green algae." *Pesticide biochemistry and physiology* 104.1 (2012): 50-57.
- Series, Marine Foresight. "Ocean Acidification: An Emerging Threat to our Marine Environment." (2010).
- Sokolova, Inna M., et al. "Energy homeostasis as an integrative tool for assessing limits of environmental stress tolerance in aquatic invertebrates." *Marine environmental research* 79 (2012): 1-15.
- Tee, Si Yin, et al. "Recent progress in energy-driven water splitting." *Advanced science* 4.5 (2017): 1600337.
- Thomaidis, Nikolaos S., et al. "Optimal management of wind and solar energy resources." *Computers & Operations Research* 66 (2016): 284-291.
- Thresher, Robert, Michael Robinson, and Paul Veers. "To capture the wind." *IEEE Power and Energy Magazine* 5.6 (2007): 34-46.
- Titorova, Valentina, et al. "Characterization of MK-40 membrane modified by layers of cation exchange and anion exchange polyelectrolytes." *Membranes* 10.2 (2020): 20.
- Varliero, Selene, et al. "Chemical Aspect of Ocean Liming for CO₂ Removal: Dissolution Kinetics of Calcium Hydroxide in Seawater." *ACS Engineering Au* (2024).

PTSD Symptoms Resulting from Incarceration in America by Siena Coburn

By Siena Coburn

Abstract

While research has shown a high correlation between trauma and incarceration, few studies exist that clearly isolate the impact of trauma stemming from incarceration itself from that of pre-existing trauma. Despite this gap in research, among the formerly incarcerated community, talk of "Post-Incarceration Syndrome" or the lasting and damaging effects of incarceration that imitate symptoms of PTSD is widespread. This paper synthesizes existing research that explores those claims with the purpose of identifying the gaps in research and clarifying the nature of the problem so as to examine viable solutions. It finds that the exposure to violence and institutionalization that often occurs in American prisons have a high probability of bringing on symptoms of PTSD after release. Furthermore, these negative effects are often compounded by previously existing trauma. This paper concludes that moderate evidence demonstrates that small policy or management changes alongside non-governmental programs can ease this issue; however, further research is needed.

Introduction

There has long been an established correlation between incarceration and post-traumatic stress disorder (PTSD) (Belet et al). The idea that trauma can contribute to behaviors that lead to incarceration is well-known. Research has concluded that traumatic experiences are highly correlated with a likelihood of being in prison (Jäggi). In fact, data from ten U.S. prisons found that over half of incarcerated men reported traumatic experiences (Wolff and Shi). Furthermore, research confirmed that incarcerated populations have rates of mental illness much higher than the general population (National Health Service). Additionally, incarcerated populations tend to have high rates of childhood trauma (Jäggi, Wolff and Shi). These pre-existing mental health disorders and trauma put incarcerated individuals at heightened risk of developing PTSD from experiences within prison (National Health Service).

However, the prevalence of trauma resulting from incarceration and its reasons have been seldom investigated. Given incarcerated populations' heightened baseline risk of mental health disorders, including PTSD and the fact that mental health issues are known to increase the likelihood of recidivism, this becomes vital to examine (Sadeh and McNiel). Additionally, the U.S. has the world's largest incarcerated population, incarcerating 614 out of 100,000 people (Widra). Considering the size of the population affected by this issue, the issue becomes one of considerable moral weight and an issue of more comprehensive societal safety.

The relationship between trauma and incarceration is two-directional. Many people arrive in prison with trauma already and prisons have the potential to create trauma or re-traumatize. For the purposes of this research paper, the focus will be any trauma clearly resulting from the conditions of incarceration. This type of trauma has gained a reputation in the formerly incarcerated community and has developed a name of its own: Post Incarceration

Syndrome or PICS (National Incarceration Association). While Post Incarceration Syndrome is not a medically diagnosable syndrome, it is defined by the National Incarceration Association (NIA) as a psychiatric disorder affecting formerly incarcerated individuals characterized by a range of psychological, emotional, and social difficulties such as depression, anxiety, flashbacks, hypervigilance, difficulty adjusting to life outside of prison, and difficulty forming and maintaining relationships (National Incarceration Association). The Mayo Clinic notes that symptoms of PTSD include consistently feeling frightened, hypervigilance, self-destructive behavior, irritability, insomnia, and loneliness (Mayo Clinic). The NIA explains that there is no meaningful difference in symptoms between PICS and PTSD because the symptoms are largely the same. However, as PTSD is an officially diagnosable psychiatric disorder and PICS is not, research on PTSD is more prevalent in the literature. As a result, this paper will use the term PTSD in lieu of PICS. However while trauma from incarceration has become large enough of a problem to garner its own unique name, quality research into it has been limited.

In this paper, PTSD symptoms will be examined in the context of incarceration in American prisons, as opposed to alternate forms of detainment such as county jails due to prison conditions being examined more prevalently in existing literature. While, the goal of most U.S. prisons is to act as a “correctional facility,” however, considering that, as of 2023, 44% of released inmates were re-incarcerated within a year, our prisons fall tremendously short (Lahdon). The significance of this literature review lies in its ability to detail one specific impact of American incarceration on the individual that is known to increase future chances of committing a crime. Hence, this research has implications for individuals, as well as recidivism and wider societal safety. Thus, this paper will explore in what ways American incarceration leads to symptoms of PTSD.

PTSD Resulting from Incarceration

The large majority of research done on U.S. incarceration has found strong connections between prison conditions and experiences and symptoms of trauma. First-hand experiences and conversations within the formerly incarcerated community create some of the strongest arguments for incarceration creating trauma. Mika’il DeVeaux, a formerly incarcerated individual, details his experience in New York prisons in the Harvard Law Journal. He plainly describes his time as “traumatic” and “dehumanizing,” explaining that it’s impossible to “leave unscarred” due to the nature of confinement and the conditions inmates are subject to (DeVeaux, 257). Other existing research provides limited support for the idea that many symptoms of PTSD emerge from experiences in prison (Bates; Liem and Kunst). A PhD dissertation from Walden University featured interviews with ten formerly incarcerated individuals who represented the penal demographics of the U.S. in terms of race, gender, age, and time incarcerated. While the sample size was limited, the interviews suggested that many symptoms of PTSD emerge from experiences in prison (Bates). Furthermore, a study published by the *International Journal of Law and Psychiatry* found a noticeable cluster of PTSD symptoms among a group of released inmates who were initially sentenced to life, who served an average

of 19 years, confirming that PTSD symptoms often result from long-term incarceration (Liam and Krust). It's unclear if this correlation would be less strong, if less time is spent in prison, as the study failed to touch on this. It should be noted that this data was collected internationally, from both the United States and Europe; given that overcrowding and violence in prison are known to lead to PTSD (Baggio et al) and those issues are even more prevalent in the U.S. compared to European countries, the correlation between incarceration and PTSD symptoms would likely be just as strong if not stronger had the U.S. been isolated in that study. A separate peer reviewed article published in the International Journal of Criminology and Sociological Theory examined this problem in more depth and confirmed that there is a great deal of reason to believe that many incarcerated individuals' mental health issues relate significantly to the prison atmosphere (Armour). The question then becomes: what are the causes of these issues?

Violence In Prison

One theory for the cause of prison-related trauma is the exposure to consistent violence in prison. The idea that violence can lead to PTSD symptoms isn't new; PTSD is most commonly associated with war veterans who have witnessed violence, however, the effects of witnessing violence specifically within prison has been researched very minimally (Mayo Clinic). Violence is assumed to be highly prevalent in prison (Dholakia). However, conditions vary by location and many incidents go unreported. A study done based on over 7,000 U.S. inmates found a 6-month physical victimization rate of up to 21%, over 10 times the rate in the general community (Wolff et al). Other studies reported findings similarly above the community average (Wooldredge).

It should be noted that individuals who witness traumatic events can also be prone to developing trauma, not just the direct victims, so exposure to violence would be a better metric than the victimization rate (World Health Organization). However, minimal research has assessed inmates' exposure to violence. Furthermore, a perception of a constant possible threat of violence could probably produce symptoms of PTSD similar to as if the events were actually experienced or witnessed, considering that one of the above symptoms of PTSD is sustained hypervigilance or paranoia. For example, a study done in a Tennessee Maximum Security Prison found that 40 percent of inmates consistently avoided areas they thought to be high risk (Haney), demonstrating the widespread hypervigilance and belief in a constant threat. Therefore, while the statistical probability of violence is likely correlated with an inmate's perception of the probability of violence, it is possible for a scenario where there is a low-level statistical threat of violence and minimal violence personally witnessed. Still, due to the very fact of being in prison with violent offenders where one might expect to have a higher chance of danger, one could develop much higher levels of paranoia and fear. However, no research to date has substantiated this claim; this possibility merely explains why the data available about prevalence of violence in prisons is not the only data that would be relevant in addressing the issue of PTSD symptoms developing from the violent conditions of prison, but it is what is available.

Limited research has examined the direct psychological effects of prison violence, given the challenge of collecting data within a prison, however, personal anecdotes among the formerly

incarcerated population draw clear connections between violence exposure and poor mental health (DeVeaux). Furthermore, research has drawn many correlations between community violence and PTSD. A study published in the *Journal of Abnormal Child Psychology* examined 94 children who attended inner-city schools and found that the correlation between violence exposure and PTSD was highly significant, even after controlling for other mental illnesses such as depression and suicidal tendencies (Mazza and Reynolds). A different study also examining the consequences of community violence exposure found that 50% of children in a high-violence inner city neighborhood met the criteria for PTSD (Horowitz et al). While these studies are limited by only having assessed children in a city environment, they support the idea that violence exposure likely is a significant contributing factor to PTSD symptoms, including for those in prison. In fact, other research found that witnessing violence predicted for PTSD symptoms better than all other factors including other traumatic event history (Zinzow et al). Formerly incarcerated individuals recounting their time in prison corroborate this idea of trauma resulting from violence exposure, with many detailing the brutality of day-to-day prison life (DeVeaux). While specific evidence is lacking in assessing the effects of violence exposure, it's highly probable that these carceral conditions lead to PTSD.

Prison Environment and Institutionalization

Others have hypothesized that the prison environment itself, independent of any violence, has the potential to result in trauma. The habits of survival that result from assimilating into the prison environment, known as institutionalization or prisonization, are known to be severely detrimental to mental health. These effects could ultimately result in individuals exhibiting PTSD symptoms. Some research finds that because of prisons' inherently dangerous nature, due to their being no exits, many inmates develop severe hypervigilance, a significant symptom of PTSD (Haney). While in prison, these measures would have rational grounds, if the stress of hypervigilance stays with a person after release, it can develop into a symptom of PTSD. Furthermore, the "rigid and unyielding discipline of prison" that requires the severe adjustment known as institutionalization can cause a great sense of overwhelm in handling typical day to day stressors or unplanned accidents that a non-institutionalized person would otherwise have no problem handling (Haney). The study doesn't pursue the consequences of feeling greatly overwhelmed by day to day activities; however, if one feels overwhelmed so easily by simple stressors, it follows logically they may withdraw themselves from unnecessary events, likely social in nature. A study published in the *Journal of Ethnography* corroborates this idea. Interviewing many former convicts, they found a common theme of institutionalization creating anxiety, resulting in social withdrawal. This was so common that the residence of the halfway house where the researcher stayed already had a name for the practice: "isolating." This social withdrawal is a common symptom of PTSD (Mayo Clinic).

Additionally, following anxiety and hypervigilance often comes irritability and agitation, other PTSD symptoms. Multiple studies, although not with formerly incarcerated populations, have found that anxiety, like that that comes with institutionalization, established above, is

associated with irritability and agitated outbursts (Perlis et al). Hence, while not much research has framed it as such, it appears highly probable that the nature of U.S. prisons and the process of institutionalization may contribute to PTSD symptoms in formerly incarcerated individuals.

It should be noted that with incarcerated populations there is a possibility of retraumatization as a result of prison conditions. Investigating whether it is these conditions of incarceration or already shaky mental states that are producing symptoms of PTSD would provide greater clarity on this issue, however, the research on this is extremely limited and, hence, that discussion is beyond the scope of this paper.

Potential Solutions

Is this the nature of prison? Or is this just a US problem? The inclusive evidence on the subject makes this extremely hard to unravel. Few studies have recently done a comparative analysis of prison conditions (such as violence) across countries, let alone one that also breaks down the causation of various mental illnesses in formerly incarcerated groups across countries. Given the inconclusive research, it's nearly impossible to eliminate this as a problem strictly facing the US, and thus, there may be factors influencing this beyond the control of our policy or prison symptoms. Furthermore, establishing what factors specifically cause institutionalization has yet to be studied on a detailed enough level to be applied to policy changes. However, there is plenty of evidence to show that general prison conditions, such as overcrowding, ventilation, and even food can affect levels of violence in prisons.

Research shows a correlation between rising temperatures and increased prison violence (Sanders). However, many US jails and prisons lack air conditioning, about four in five of the roughly 150,000 people held in Texas prisons in 2017 had no access to air conditioning in their cells (Dholakia). Additionally, research has shown that the food and nutrients provided in prison can reduce antisocial behavior, including violence (Gesch et al.). When prisons are overcrowded or understaffed, violence also becomes more likely, and federal prisons in the United States are chronically about 10 percent over capacity (Baggio et al., United States Justice Department). This issue is often further exacerbated in state prisons. In February of 2024, Alabama's correctional facilities operated at nearly 169% capacity meanwhile their staff has decreased (Burkhalter). Evidence from these prisons backs up the aforementioned research on overcrowding's effect on violence. In 2023, over 300 people died in Alabama state prisons, with only 70 considered deaths of "natural causes" (Burkhalter). Given the relative simplicity of changing temperature, nutrition, and overcrowding, future policy should look to control those factors to prevent violence and PTSD.

Alternatively, a more nuanced approach, other prisons have attempted to address trauma directly. Many prisons have experimented with a different approach called "Trauma Informed" care. The results and efficacy of implementation in these studies varied and often were done on a very small scale with little quantitative data. However, the existence of these programs suggests that those working in prison too have noticed the prevalence of trauma and have a belief that policy and programs can change this. One study conducted in Malaysian Prisons supports this.

The study followed 40 inmates who were given Cognitive Behavioral Therapy and found that compared to the control group, these men's anger and aggression levels significantly dropped (Ayub et al). Furthermore, a meta-analysis of existing research examined trauma-informed care directly with regards to minimizing PTSD and found there was a significant positive influence of trauma-focused interventions (Malik et al). These solutions of course require deeper policy change and adjustment in how society views the nature of prison and therefore may be less attainable than changing simpler factors of temperature and nutrition in the short term.

Similarly, other programs have addressed the culture of prison as a whole. The Vera Institute of Justice has a program called Restoring Promise where they partner with correctional departments and "transform housing units so they are grounded in dignity" (Djokovic and Shanahan). The housing units have many windows to let in light, walls covered in murals, and allow photographs to hang on people's walls. Here, people spend most of the day outside their cell attending classes on financial literacy, life skills, and nurturing healthy relationships. The program also provides incarcerated individuals with a mentor, usually older men serving longer sentences. This has resulted in a safer prison environment. Living in a Restoring Promise unit during incarceration decreased incarcerated individuals' chances of being convicted of a violent infraction by 73 percent (Djokovic and Shanahan). Recidivism rates don't give the exact data for violence inside the unit. However, they do demonstrate that these units lessened individuals' violent tendencies overall, which is very promising as a method for decreasing PTSD outcomes. As many incarcerated individuals when sharing their trauma often reference treatment by the guards and the general prison culture, policy and program efforts like Restoring Promise should be at the forefront of solutions when talking about inmate trauma, mental health, and PTSD. First, however, more in depth research illuminating the extent of trauma incurred in prison would likely be needed to generate momentum for change.

Discussion and Areas of Further Research

There is much research still needed to examine the process of trauma in American prisons. First of all, the prevalence of prisons creating PTSD and other mental health systems, and by extension the prevalence of violence and poor prison culture, needs to be studied on a much larger scale. Most studies are highly local and rely on qualitative data, making the severity of this issue and the urgency of solving it unclear. If the problem can't be quantified, policy advocates' and government officials' attention can't be brought to the problem. Similarly, and as briefly mentioned previously, studies should consider assessing not only the prevalence of violence in prison but the prevalence of feeling threatened by the possibility of violence, as one can likely develop symptoms of PTSD, like hypervigilance, without having directly experienced the violence themselves. Additionally, researchers should work to develop a framework to assess the culture of prisons and their relation to mental health. Limited data is available on this so far despite that interviews with formerly incarcerated individuals have suggested the culture of prison (including treatment by the guards and dehumanization) may contribute to trauma. This

idea of culture has overlapping themes with institutionalization, and hence is important to add to the existing literature.

Furthermore, the existence of previous trauma at high rates in incarcerated populations presents a significant issue for discussing this topic, yet makes it so much more important. Future research should look to examine more critically how early trauma might shape individuals' responses to conditions in prisons, while also delineating more clearly between previous mental health issues and trauma stemming solely from prison. More research needs to evaluate the extent to which outcomes are the same in other countries' carceral systems to provide insight into solvency and to shed light on the extent of the problem in the United States.

Finally, research must be applied directly to policy measures. Certain factors in prisons can affect violence prevalence and by extension, trauma stemming from incarceration, however, much more research is needed to assess policy choices and aspects of prison culture that may shape traumatic outcomes. Research must go beyond efficacy and examine practicality, taking into account public response and attitudes towards incarceration to ensure feasibility.

Works Cited

- Amour, Cherie. "Mental Health in Prison: A Trauma Perspective on Importation and Deprivation." *International Journal of Criminology and Sociological Theory*, vol. 6, no. 3, 2013, pp. 1-12. York University, ijcst.journals.yorku.ca/index.php/ijcst/article/view/35703/32435. Accessed 15 Sept. 2024.
- Baggio, Stéphanie, Nicolas Peigné, Patrick Heller, Laurent Gétaz, Michael Liebreznz, and Hans Wolff. "Do Overcrowding and Turnover Cause Violence in Prison?" *Frontiers in Psychiatry*, vol. 10, 2019, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6992601/>. Accessed 15 Sept. 2024.
- Bates, Wanda. "A Phenomenological Examination of Prisonization and the Psychological Effects of Incarceration." *ProQuest Dissertations Publishing*, 2018.
- Belet B;D'Hondt F;Horn M;Amad A;Carton F;Thomas P;Vaiva G;Fovet T; *Post-Traumatic Stress Disorder in Prison*, U.S. National Library of Medicine, Dec. 2020, pubmed.ncbi.nlm.nih.gov/32921495/.
- Burkhalter, Eddie. "Record Loss of Life in 2023 Pushes ADOC's Death Total Over 1,000 Since DOJ Put State on Notice." *Alabama Appleseed*, 29 Jan. 2024, alabamaappleseed.org/author/eddie-burkhalter/record-loss-of-life-in-2023-pushes-adocs-death-total-over-1000-since-doj-put-state-on-notice/. Accessed 19 Sept. 2024.
- DeVeaux, Mika'il. "The Trauma of the Incarceration Experience." **Harvard Civil Rights-Civil Liberties Law Review**, vol. 48, no. 1, 2013, pp. 257-277. **Harvard Law School**, https://journals.law.harvard.edu/crcl/wp-content/uploads/sites/80/2013/04/DeVeaux_257-277.pdf. Accessed 15 Sept. 2024.
- Dholakia, Nazish. "Prisons and Jails Are Violent; They Don't Have to Be." *Vera Institute of Justice*, 18 Oct. 2023, www.vera.org/news/prisons-and-jails-are-violent-they-dont-have-to-be. Accessed 19 Sept. 2024.
- French, Sheila A., and Paul Gendreau. "Reducing Prison Misconducts: What Works!" *Criminal Justice and Behavior*, vol. 33, no. 2, 2006, pp. 185-218, <https://doi.org/10.1177/0093854805284406>. Accessed 19 Sept. 2024.
- Gesch, C. B., et al. "Influence of Supplementary Vitamins, Minerals and Essential Fatty Acids on the Antisocial Behaviour of Young Adult Prisoners." *The British Journal of Psychiatry*, vol. 181, no. 1, 2002, pp. 22-28, www.cambridge.org/core/journals/the-british-journal-of-psychiatry/article/influence-of-supplementary-vitamins-minerals-and-essential-fatty-acids-on-the-antisocial-behaviour-of-young-adult-prisoners/04CAABE56D2DE74F69460D035764A498. Accessed 19 Sept. 2024.
- Hamm, John. *Monthly Statistical Report: February 2024*. Alabama Department of Corrections, Feb. 2024, doc.alabama.gov/docs/MonthlyRpts/February%202024.pdf. Accessed 19 Sept. 2024.

- Haney, Craig. "The Psychological Impact of Incarceration: Implications for Post-Prison Adjustment." *U.S. Department of Health and Human Services*, 2001, https://aspe.hhs.gov/sites/default/files/migrated_legacy_files/42351/Haney.pdf. Accessed 15 Sept. 2024.
- Haney, Craig. "Prison Effects in the Age of Mass Incarceration." *The Prison Journal*, 25 July 2012, https://www.researchgate.net/profile/Craig-Haney/publication/274998193_Prison_Effects_in_the_Era_of_Mass_Incarceration/links/55d9ef7908ae9d659491e8a0/Prison-Effects-in-the-Era-of-Mass-Incarceration.pdf. Accessed 15 Sept. 2024.
- Her Majesty's Inspectorate of Prisons. *Life in Prison: Food*. Sept. 2016, www.justiceinspectors.gov.uk/hmiprisons/wp-content/uploads/sites/4/2016/09/Life-in-prison-Food-Web-2016.pdf. Accessed 19 Sept. 2024.
- Horowitz, Karyn, et al. "Community Violence and Urban Families: Experiences, Effects, and Directions for Intervention." *Wiley Online Library*, *American Journal of Orthopsychiatry*, 24 Mar. 2010, onlinelibrary.wiley.com/doi/abs/10.1037/0002-9432.75.3.356.
- Jaggi, Lena J., et al. "The Relationship between Trauma, Arrest, and Incarceration History among Black Americans: Findings from the National Survey of American Life." *Society and Mental Health*, vol. 6, no. 3, American Sociological Association, 2016, pp. 352-361. *PMC*, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5079438/>. Accessed 23 Sept. 2024.
- Lahdon, Tenzing. "From the Desk of BJA: November 2023." *Justice Matters*, Bureau of Justice Assistance, 2023, bja.ojp.gov/news/justice-matters/desk-bja-november-2023. Accessed 23 Sept. 2024.
- Liem, Marieke, and Maarten Kunst. "Is There a Recognizable Post-incarceration Syndrome Among Released 'Lifers'?" *International Journal of Law and Psychiatry*, vol. 36, no. 3-4, 30 Apr. 2013, pp. 333-337. <https://pubmed.ncbi.nlm.nih.gov/23642320/>. Accessed October 20, 2024.
- Malik, Nabeela, Emma Facer-Irwin, and Deirdre MacManus. "The Effectiveness of Trauma-Focused Interventions in Prison Settings: A Systematic Review and Meta-Analysis." *Trauma, Violence, & Abuse*, vol. 24, no. 2, 2021, pp. 395-411. Sage Journals, <https://doi.org/10.1177/15248380211043890>. Accessed 28 Oct. 2024.
- Mayo Clinic Staff. "Post-traumatic Stress Disorder (PTSD)." *Mayo Clinic*, Mayo Foundation for Medical Education and Research, 2 Mar. 2023, www.mayoclinic.org/diseases-conditions/post-traumatic-stress-disorder/symptoms-causes/syc-20355967. Accessed 23 Sept. 2024.
- Mazza, James, and William Reynolds. "Exposure to Violence in Inner-City Adolescents: Relationships with Suicidal Ideation, Depression, and PTSD Symptomology." *Journal of Abnormal Child Psychology*, vol. 27, no. 3, 1999, pp. 203-213. *SpringerLink*, <https://doi.org/10.1023/A:1021900423004>. Accessed 15 Sept. 2024.

- Mohamad, Mohd Suhaimi, et al. "Cognitive Behavioural Group Counselling in Reducing Anger and Aggression among Male Prison Inmates in Malaysia." *ResearchGate*, Apr. 2016, www.researchgate.net/profile/Mohd-Suhaimi-Mohamad/publication/301608531_Cognitive_Behavioural_Group_Counselling_in_Reducing_Anger_and_Aggression_among_Male_Prison_Inmates_in_Malaysia/links/571c950308ae7f552a48786f/Cognitive-Behavioural-Group-Counselling-in-Reducing-Anger-and-Aggression-among-Male-Prison-Inmates-in-Malaysia.pdf. Accessed 19 Sept. 2024.
- Perlis, Roy H., et al. "Irritability is Associated with Anxiety and Greater Severity, but Not Bipolar Spectrum Features, in Major Depressive Disorder." *Acta Psychiatrica Scandinavica*, vol. 123, no. 3, 2011, pp. 252-260. *PubMed Central*, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3312008/>. Accessed 20 Sept. 2024.
- Piper, Alicia, and David Berle. "The Association Between Trauma Experienced During Incarceration and PTSD Outcomes: A Systematic Review and Meta-Analysis." *Psychology, Crime & Law*, vol. 26, no. 9, 2019, pp. 945-965. *Taylor & Francis Online*, <https://doi.org/10.1080/14789949.2019.1639788>. Accessed 15 Sept. 2024.
- Sadeh, Naomi, and Dale McNiel. "Posttraumatic Stress Disorder Increases Risk of Criminal Recidivism Among Justice-Involved Persons with Mental Disorders." *Criminal Justice and Behavior*, vol. 42, no. 6, 2015, pp. 573-586. <https://doi.org/10.1177/0093854814556880>. Accessed 19 Sept. 2024.
- Subramanian, Ram. "How Some European Prisons Are Based on Dignity Instead of Dehumanization." *Brennan Center for Justice*, 6 June 2019, <https://www.brennancenter.org/our-work/analysis-opinion/how-some-european-prisons-are-based-dignity-instead-dehumanization>. Accessed 15 Sept. 2024.
- "The Definitive Guide on Post Incarceration Syndrome (PICS)." *National Incarceration Association*, 14 Apr. 2024, [joinnia.com/post-incarceration-syndrome/#:~:text=Post%2Dincarceration%20syndrome%20\(PICS\),a%20result%20of%20being%20imprisoned](http://joinnia.com/post-incarceration-syndrome/#:~:text=Post%2Dincarceration%20syndrome%20(PICS),a%20result%20of%20being%20imprisoned). Accessed 19 Sept. 2024.
- The Marshall Project. "Cooking Them to Death: The Lethal Toll of Hot Prisons." *The Marshall Project*, 11 Oct. 2017, www.themarshallproject.org/2017/10/11/cooking-them-to-death-the-lethal-toll-of-hot-prisons. Accessed 19 Sept. 2024.
- United States Department of Justice. *FY 2024 Performance Budget Congressional Submission*. 23 Mar. 2023, https://www.justice.gov/d9/2023-03/bop_se_fy_2024_pb_narrative_omb_cleared_3.23.2023.pdf. Accessed 15 Sept. 2024.
- Vera Institute of Justice. *Changing Prison Culture Reduces Violence*. Vera Institute of Justice, 2023, www.vera.org/publications/changing-prison-culture-reduces-violence. Accessed 19 Sept. 2024.
- Widra, Emily. "States of Incarceration: The Global Context 2024." *Prison Policy Initiative*, 2024, www.prisonpolicy.org/global/2024.html. Accessed 23 Sept. 2024.

- Wooldredge, John. "Inmate Lifestyles and Opportunities for Victimization." *Journal of Research in Crime and Delinquency*, vol. 35, no. 4, 1998, pp. 480-502. <https://doi.org/10.1177/0022427898035004006>. Accessed 23 Sept. 2024.
- Wolff, Nancy, and Jing Shi. "Childhood and Adult Trauma Experiences of Incarcerated Persons and Their Relationship to Adult Behavioral Health Problems and Treatment." *International Journal of Environmental Research and Public Safety*, vol. 9, no. 5, 2012, pp. 1908–1926. *PMC*, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3386595/>. Accessed 23 Sept. 2024.
- Zinzow, Heidi M., et al. "Prevalence and Mental Health Correlates of Witnessed Parental and Community Violence in a National Sample of Adolescents." *Journal of Child Psychology and Psychiatry*, vol. 50, no. 4, 2009, pp. 441-450. Wiley Online Library, <https://doi.org/10.1111/j.1469-7610.2008.02004>. Accessed 29 Sept. 2024

Polycrystalline Thin Film Technologies for Solar Photovoltaics By Edison Y. Chen

Abstract

This paper shows how polycrystalline thin-film solar cells are a promising alternative to monocrystalline silicon-based photovoltaics. With lower production costs and rapidly improving efficiencies, technologies like cadmium telluride (CdTe) and perovskite solar cells are a promising solution which address the limitations of silicon. The study focuses on advancements in manufacturing costs, improvements in efficiency, and design optimizations that enhance cell performance. Key factors such as grain texture, thickness, and recombination losses are analyzed. The findings suggest that with ongoing improvements, polycrystalline thin-film solar cells could become the dominant technology in solar energy, driving more efficient and sustainable solutions.

Introduction

The ever-increasing global demand for sustainable energy sources has established solar energy as a pivotal component in the transition from fossil fuels to renewable energy. During the past 20 years, solar cells have seen decreases in cost, improvements in efficiency, and an increased awareness of the benefits of renewable energy's benefits. Despite recent improvements, solar cells still accounted for only 4% of total electricity generation in the United States in 2023 (IEA). Drawbacks such as material quality and purity affects efficiency due to their energy intensive and expensive manufacturing processes, thus there have been difficulties in obtaining pure materials in solar cells which have affected efficiency and costs thus preventing solar photovoltaics (PVs) from emerging as a main source of electricity generation. These issues associated with silicon solar cells can be overcome through alternate solar technologies, but there are also alternative solutions. Among various types of solar cells, polycrystalline thin film solar cells have recently emerged as a promising technology due to their low costs and high efficiency. Polycrystalline cells have always had lower efficiency compared to their monocrystalline counterparts, but recent advancements in manufacturing techniques and material science have significantly narrowed this efficiency gap (Fig. 1). Cadmium telluride (CdTe) polycrystalline solar cells show potential for success as they have a theoretical max efficiency of 30% compared to silicon's 29% (NREL). Polycrystalline solar cells also show potential for reductions in costs. The thin film requires less material which reduces costs while maintaining high efficiency. Since 2005, CdTe polycrystalline thin film solar cells have steadily increased in efficiency from 18% to 22.4% showing potential for it to one day replace silicon as the leader.

Polycrystalline silicon features advantages of all thin film technology such as low costs due to low material wastage with up to factor 100 less material compared to wafer-based solar cells. Wafer-based solar cells are a type of photovoltaic (PV) technology that converts sunlight into electricity using tiny slices, or wafers, of semiconductor material, most often silicon. These wafers, which are typically 150 to 200 micrometers thick, are essential components of the most prevalent type of solar panel, crystalline silicon solar panels. CdTe based thin films are being

developed by low cost non-vacuum techniques and high efficiency in small area cells created by vacuum technology have also been reported for CdTe. Perovskite solar, a cheaper alternative, has also had steady increases in efficiency to 25.7% compared to silicon's 26.7%. This paper will give an overview of promising thin film technologies and it will discuss factors like texture, coating, and thickness which affect efficiency in solar cells. compares the efficiencies of the discussed technologies to determine the most promising technology that can be used in the future.

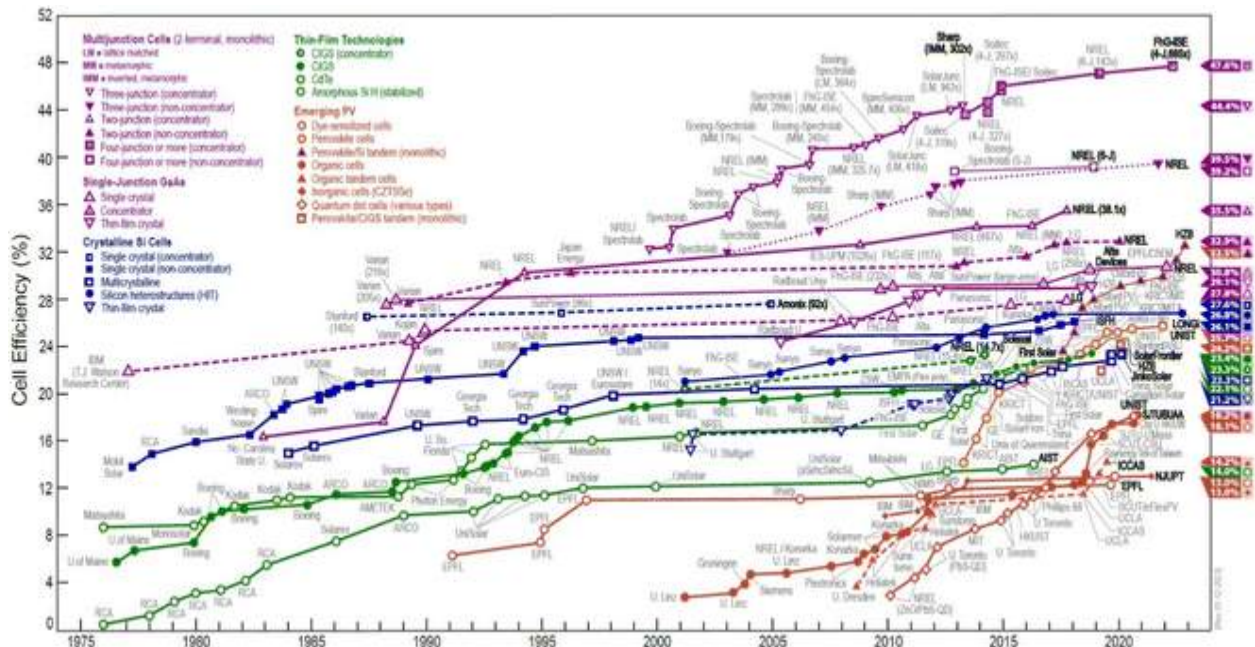


Figure 1. A chart of the highest confirmed conversion efficiencies for PV cells

Overview of Solar Technologies

Currently the most popular solar PV are made from crystalline silicon materials. These cells are either monocrystalline or polycrystalline, all which depend on the presence of grains in the microstructure. Polycrystalline cells have grain boundaries whereas monocrystalline does not which results in greater efficiency and because polycrystalline has grains, it can impede electron flow, affecting efficiency. Monocrystalline silicon cells are fabricated using single silicon crystals through a procedure called the Czochralski process where the entire solid's crystal lattice is continuous and free from grain limits. Crystalline silicon is a non-toxic material with great stability and durability while possessing an energy gap of 1.12 eV which is well-suited for maximum single junction solar cell efficiencies, meaning silicon can effectively convert sunlight into electricity. Since 2017, the record efficiency of these solar cells has reached 26.7% thanks to the heterojunction intrinsic thin layer of technology based on thin a-Si passivating layers and interdigitated back contacts on n-type silicon wafers and cells. Heterojunction Intrinsic Thin-Layer (HIT) photovoltaic cells use crystalline silicon and amorphous silicon (a-Si) to boost solar cell efficiency. The efficiency and stability of traditional silicon solar cells produced from p-type or n-type crystalline silicon wafers are limited by surface recombination and temperature

sensitivity. HIT technology addresses these challenges by employing very thin amorphous silicon (a-Si) layers on both sides of a crystalline silicon wafer. Despite their high efficiency and long lifespan, there are drawbacks to these cells such as production costs. The Czochralski method, used to grow single-crystal silicon ingots, is particularly resource-intensive. This process requires high-purity silicon and precise control over temperature and environment, leading to significant costs (Green et al., 2015).

Polycrystalline thin film solar cells are made from crystals that are grown in large blocks and then sliced into thin wafers. Unlike monocrystalline cells, which are composed of single-crystal silicon, polycrystalline cells consist of multiple crystals. This structural difference leads to distinct grain boundaries, which affect the electrical properties and efficiency of the cells. The efficiency of these cells typically range from 15-20% which is slightly lower than monocrystalline solar cells. The presence of grain boundaries in polycrystalline silicon can impede the flow of electrons, thereby reducing the overall efficiency. However, ongoing research and development have led to improvements in cell design and manufacturing processes, gradually narrowing the efficiency gap between polycrystalline and monocrystalline cells. Polycrystalline thin film solar cells are less expensive than monocrystalline due to their ability to maintain efficiency while requiring less material.

Perovskite solar cells are a type of polycrystalline thin film that have shown promising photovoltaic technology due to their high efficiency and low production costs. Perovskite polycrystalline thin films consist of numerous small crystals or grains and perovskite solar cells are based on materials with a crystal structure similar to the mineral perovskite (ABX_3). In this structure, 'A' is an organic or inorganic cation, 'B' is a metal cation, and 'X' is a halide anion. The most commonly used perovskite material for solar cells is methylammonium lead iodide ($MAPbI_3$). The grain boundaries can influence the electronic properties of the material, affecting the efficiency and stability of the solar cells but despite these challenges, perovskite polycrystalline thin films have shown remarkable performance due to their excellent light absorption and charge transport properties. Perovskite solar cells (PSCs) have reached efficiencies of up to 25.7% which can be compared to the leading silicon solar cells. These cells contain the efficiency of silicon while maintaining the low costs of a polycrystalline thin film.

Cadmium Telluride (CdTe) polycrystalline thin film solar cells show promise in cost, efficiency, and large-scale production. CdTe has a direct bandgap of 1.45 eV and usually consists of a multilayer structure substrate (usually glass), Transparent Conductive Oxide (conducting material), Cadmium Sulfide (CdS) Buffer Layer (forms p-n junction with CdTe layer), and an absorber layer. CdTe has intrinsically better temperature coefficients, energy yield, and lower degradation rates than Si technologies. The temperature coefficient specifies how the performance of a solar cell changes with temperature and the degradation rate indicates how the performance of a solar cell decreases over time. Figure 2 compares different solar PV's efficiency since 2023 and it shows that CdTe's efficiency is near silicon meaning that it can overtake silicon in the very near future. CdTe is also cheaper to manufacture and is well-suited for scaling, thus showing potential to replace silicon solar cells in the future.

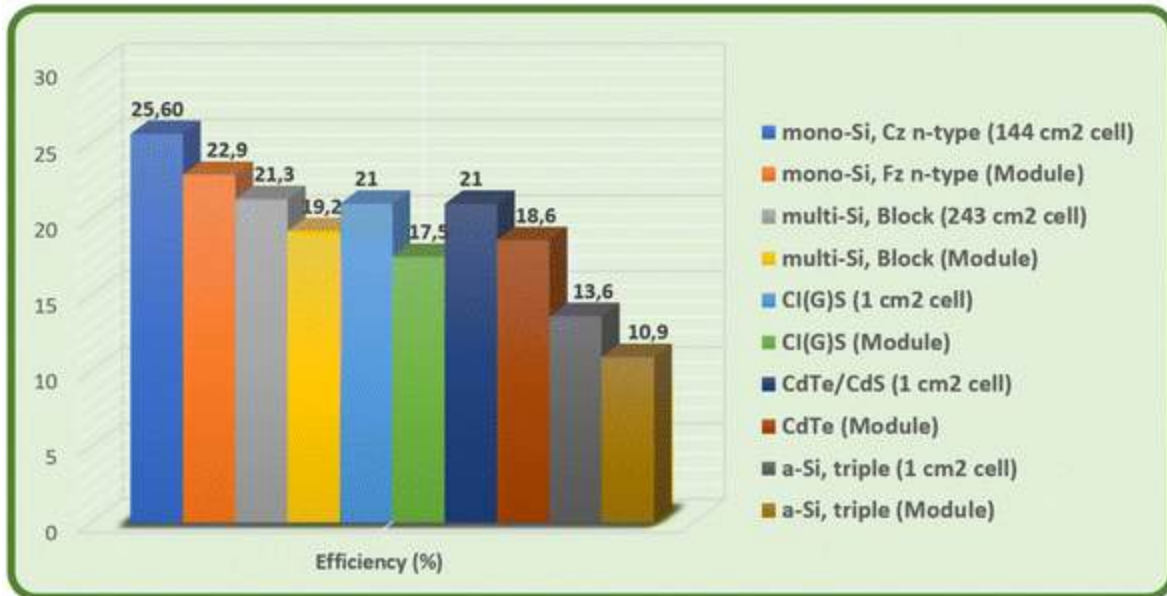


Fig. 2 Efficiency of silicon-based technology in laboratories

Factors Influencing Efficiency

By rearranging grains to produce a randomized texture with better quality, a process known as recrystallization increases the efficiency of solar cells (<https://www.sciencedirect.com/science/article/pii/S0040609024000798>). Recrystallizing the absorber can enhance crystal properties and minimize flaws, both of which are essential for increasing efficiency. Effectiveness of light absorption in polycrystalline solar cells is also influenced by their roughness. Increasing the light route length within the cell can improve absorption and be attributed to textured surfaces. Certain textures like pyramid or dome-shaped grains can improve the formation of electron-hole pairs by improving light absorption and reducing reflection. This results in an increase in the number of photons through the PV cell.

Thickness of a solar cell can impact efficiency but there is not a *set* thickness that is optimal for all. Different types such as silicon, CdTe, and perovskites have different optimal thicknesses that work best in their specific situations. The thickness of the active layer of polycrystalline solar cells affects their ability to absorb light. Thicker layers absorb more light but a certain thickness, the advantages of light absorption lessen because of increased recombination losses and reduced charge carrier collection efficiency.

Recombination occurs when electrons recombine with holes instead of contributing to the electric current. For a thickness to be optimal, it must ensure ample light absorption, enhance

carrier collection, and minimize recombination losses. In CdTe solar cells, their optimal thickness is around 2-10 μm (alluding to ‘thin films’) because of CdTe’s high absorption coefficient. This can explain why CdTe solar cells are inexpensive to build while maintaining high efficiency. The optimal thickness for perovskite cells is 0.3-0.6 μm . Perovskites have high light absorption properties which explain why cells are much thinner. Significant gains in efficiency have been made possible by the exact control of perovskite layer thickness which have been made possible by advancements in fabrication processes including solvent engineering and thermal annealing.

Different types of solar cells are prone to different recombination losses. Cadmium Telluride cells are prone to interface recombination and potential shunt resistance issues. Shunt resistance is a parallel resistance path that occurs within the cell and is usually brought by flaws and contaminants. There is frequently a significant density of defects at the boundaries between the grains of CdTe films, which can serve as sites of recombination where electrons and holes combine again without producing current. Another factor is that a thin coating of CdS (cadmium sulfide) is also commonly utilized as a buffer layer in CdTe solar cells. This can result in higher recombination because of a high density of defect states at the interface between CdTe and CdS.

Perovskite solar cells are prone to grain boundary recombination and shunt resistance from defects. The charge carrier dynamics in these cells create losses because non-radiative recombination can occur near grain boundaries when trapped carriers recombine with carriers of the opposite charge. As a result, the cell's efficiency decreases since there are fewer charge carriers contributing to the electric current.

Silicon cells are more prone to Prone to Shockley-Read-Hall (SHR) and Auger recombination, and require careful management of optical losses. Auger recombination is when an electron recombines with a hole in a semiconductor. Instead of releasing a photon, the energy is transferred to a third carrier—an electron or a hole—which subsequently releases it as heat. Since this process is non-radiative, it doesn't aid in the production of electricity. Silicon solar cells contain imperfections in the crystal structure like defects or impurities. This can result in trap states which is when charge carriers are captured by these traps, they recombine non-radiatively, reducing the number of carriers available to generate current. SHR recombination also affects the efficiency. Commonly used silicon cells have many defects which increases the chance of SHR recombination. To prevent this, high quality silicon is required, but this is impractical as it can be expensive and rare.

Efficiency Comparisons

Polycrystalline Cadmium telluride (CdTe) and Perovskite solar cells for low production costs and high efficiency. But when compared Perovskites have shown superior performance in many important areas. As previously stated, perovskite solar cells have surpassed power conversion efficiencies of 25% while CdTe have only reached a maximum efficiency of about 22% in similar settings.

There are several reasons why perovskite is superior. Polycrystalline perovskites display high light absorption properties because of their straight bandgap and high absorption coefficient. This makes it possible to create absorber layers that are thinner, often less than 1 μm thick—while yet still absorbing a sizable amount of the solar spectrum that is present. While on the other side, CdTe requires thicker layers (usually ranging from 2-3 μm) in order to absorb the same amount of light which increases material consumption and production costs. Perovskites also have long carrier lifetimes and a good defect tolerance. Typically seen in polycrystalline materials, perovskites can retain high efficiency even when flaws and grain boundaries are present. But despite having high defect tolerance as well, CdTe is more likely to experience recombination losses (described earlier) especially at grain boundaries, which can reduce polycrystalline cell efficiency.

Perovskites are scalable and can be manufactured at lower temperatures and use less complicated fabrication methods like spin coating or printing while silicon requires high temperatures and costly materials in order for production. Perovskites also show potential to surpass silicon in efficiency because of its bandgap and configuration. Perovskites can be utilized in tandem arrangements with silicon or other materials because they can be bandgap tuned by changing their composition. These tandem cells have the ability to attain higher efficiency than silicon cells alone because they can surpass the Shockley-Queisser limit for single-junction solar cells. (which is shown in figures above and also mentioned above).

Conclusion

In conclusion, while silicon (Si) solar cells have long dominated the photovoltaic market, developing solar technologies, particularly perovskite solar cells, hold enormous potential for transcending Si's limits. Significant breakthroughs are still needed to fully realize their potential and displace Si, notably in terms of efficiency, stability, and scalability. Increased investment in perovskite research is critical to addressing these problems and allowing these materials to meet the required performance parameters. By addressing these critical difficulties, perovskite solar cells may not only bridge the gap, but also transform the solar energy sector, resulting in more sustainable and efficient energy solutions.

Works Cited

- Andreani, Lucio Claudio, et al. "Silicon Solar Cells: Toward the Efficiency Limits." *Advances in Physics: X*, vol. 4, no. 1, Jan. 2019, p. 1548305. *DOI.org (Crossref)*, <https://doi.org/10.1080/23746149.2018.1548305>.
- . "Silicon Solar Cells: Toward the Efficiency Limits." *Advances in Physics: X*, vol. 4, no. 1, Jan. 2019, p. 1548305. *DOI.org (Crossref)*, <https://doi.org/10.1080/23746149.2018.1548305>.
- Baumann, Andreas, et al. "Identification of Trap States in Perovskite Solar Cells." *The Journal of Physical Chemistry Letters*, vol. 6, no. 12, June 2015, pp. 2350–54. *DOI.org (Crossref)*, <https://doi.org/10.1021/acs.jpcclett.5b00953>.
- Becker, C., et al. "Polycrystalline Silicon Thin-Film Solar Cells: Status and Perspectives." *Solar Energy Materials and Solar Cells*, vol. 119, Dec. 2013, pp. 112–23. *ScienceDirect*, <https://doi.org/10.1016/j.solmat.2013.05.043>.
- Best Research-Cell Efficiency Chart*. <https://www.nrel.gov/pv/cell-efficiency.html>. Accessed 18 Aug. 2024.
- Cheng, Huiyuan, et al. "Understanding and Minimizing Non-Radiative Recombination Losses in Perovskite Light-Emitting Diodes." *Journal of Materials Chemistry C*, vol. 10, no. 37, 2022, pp. 13590–610. *pubs.rsc.org*, <https://doi.org/10.1039/D2TC01869A>.
- Deng, Rong, et al. "Remanufacturing End-of-life Silicon Photovoltaics: Feasibility and Viability Analysis." *Progress in Photovoltaics: Research and Applications*, vol. 29, no. 7, July 2021, pp. 760–74. *DOI.org (Crossref)*, <https://doi.org/10.1002/pip.3376>.
- Di Sabatino, Marisa, et al. "Silicon Solar Cells: Trends, Manufacturing Challenges, and AI Perspectives." *Crystals*, vol. 14, no. 2, Feb. 2024, p. 167. *www.mdpi.com*, <https://doi.org/10.3390/cryst14020167>.
- Dolia, Kshitiz, et al. "Four-Terminal Perovskite–CdSeTe Tandem Solar Cells: From 25% toward 30% Power Conversion Efficiency and Beyond." *Solar RRL*, May 2024. *www.osti.gov*, <https://doi.org/10.1002/solr.202400148>.
- Esmailpour, Hamidreza, et al. "Influence of Auger Heating and Shockley-Read-Hall Recombination on Hot-Carrier Dynamics in InGaAs Nanowires." *Physical Review B*, vol. 109, no. 23, June 2024, p. 235303. *APS*, <https://doi.org/10.1103/PhysRevB.109.235303>.
- Kondolot Solak, Ebru, and Erdal Irmak. "Advances in Organic Photovoltaic Cells: A Comprehensive Review of Materials, Technologies, and Performance." *RSC Advances*, vol. 13, no. 18, 2023, pp. 12244–69. *pubs.rsc.org*, <https://doi.org/10.1039/D3RA01454A>.
- . "Advances in Organic Photovoltaic Cells: A Comprehensive Review of Materials, Technologies, and Performance." *RSC Advances*, vol. 13, no. 18, 2023, pp. 12244–69. *pubs.rsc.org*, <https://doi.org/10.1039/D3RA01454A>.
- Kornienko, Vladislav, et al. "Absorber Texture and the Efficiency of Polycrystalline Thin Film CdTe Solar Cells." *Thin Solid Films*, vol. 793, Mar. 2024, p. 140277. *ScienceDirect*, <https://doi.org/10.1016/j.tsf.2024.140277>.

- Kuciauskas, Darius, et al. "Identification of Recombination Losses in CdSe/CdTe Solar Cells from Spectroscopic and Microscopic Time-Resolved Photoluminescence." *Solar RRL*, vol. 5, no. 4, Feb. 2021. www.osti.gov, <https://doi.org/10.1002/solr.202000775>.
- Lee, Taesoo D., and Abasifreke U. Ebong. "A Review of Thin Film Solar Cell Technologies and Challenges." *Renewable and Sustainable Energy Reviews*, vol. 70, Apr. 2017, pp. 1286–97. *ScienceDirect*, <https://doi.org/10.1016/j.rser.2016.12.028>.
- Liu, Bin, et al. "Novel Broad Spectral Response Perovskite Solar Cells: A Review of the Current Status and Advanced Strategies for Breaking the Theoretical Limit Efficiency." *Journal of Materials Science & Technology*, vol. 140, Mar. 2023, pp. 33–57. *ScienceDirect*, <https://doi.org/10.1016/j.jmst.2022.06.055>.
- Liu, Fengzhen, and Yurong Zhou. "Polycrystalline Silicon Thin Film." *Handbook of Photovoltaic Silicon*, edited by Deren Yang, Springer, 2019, pp. 757–90. *Springer Link*, https://doi.org/10.1007/978-3-662-56472-1_29.
- McGott, Deborah L. *Cadmium Telluride Solar Cells: From Fundamental Science to Commercial Applications*. NREL/PO-5K00-86698, National Renewable Energy Laboratory (NREL), Golden, CO (United States), 8 Aug. 2023. www.osti.gov, <https://www.osti.gov/biblio/1995013>.
- "Solar." *IEA*, <https://www.iea.org/energy-system/renewables/solar-pv>. Accessed 18 Aug. 2024.
- Su, Qiao, et al. "Theoretical Limiting-efficiency Assessment on Advanced Crystalline Silicon Solar Cells with Auger Ideality Factor and Wafer Thickness Modifications." *Progress in Photovoltaics: Research and Applications*, vol. 32, no. 9, Sept. 2024, pp. 587–98. *DOI.org (Crossref)*, <https://doi.org/10.1002/pip.3790>.

The Effect of Gender Stereotypes Found on Zhihu on STEM Career Perceptions for Chinese High School Girls By Zoey Zhi-ning Chen

Abstract

In China, societal norms and the education system are founded upon Confucianism, which highlights the yin-yang work balance. Within this work balance, males are often seen as having more chauvinistic qualities, such as individuality, independence, and competitiveness, traits often tied to STEM fields. On the other hand, females are seen as having soft and conforming characteristics that are related to raising children and being submissive in a familial role. Thus, societal expectations placed upon students in China often highlight female students as inferior to male students in the subjects of science, technology, engineering, and mathematics (STEM). This discourages high school female students from pursuing STEM pathways in their College Entrance Exam, which defines one's future career. A lack of social support for women's STEM career outlook has caused many female students to hesitate in choosing STEM careers over social science careers. This study uses critical discourse analysis to analyze responses on the impact of the social media platform Zhihu, the Chinese version of Quora, on reinforcing gender stereotypes and affecting career perceptions of Chinese female high school students regarding STEM vs. social science pathways. This work is grounded in social cognitive career theory, social role theory, and gender congruity theory to address the portrayal of females in STEM in Zhihu users' postings. The responses are categorized as either reinforcing or challenging these yin-yang gender stereotypes. Finally, this study ties the results found to previous research that highlighted the impact of social media on female high school students' STEM career outlook to conclude that there is a heavy emphasis on environmental stereotype themes and references to male students having more inherent talent suited to STEM than that of female students in Zhihu responses.

Introduction

Gender inequality has existed across diverse cultures for centuries. While there have been improvements in many fields of study and workforces across the globe, gender disparity remains a prominent issue in STEM fields such as the physical sciences (Hazari et al. 96).

In China, this gender disparity is apparent across all levels of education and even the workforce (Guo et al. 225–35). In one of China's leading science institutions, the Chinese Academy of Sciences (CAS), statistics released revealed that in December 2019, only approximately six percent of academicians of CAS were women, while only five percent of the Chinese Academy of Engineering were women (Gu 2). Between the two academies, there were less than 100 female academicians out of 1600 total academicians (Gu 2). Furthermore, data released by the China Association for Science and Technology (CAST) in 2015 demonstrates that while 21 percent of members of all CAST academic societies were women, only 13 percent of the society council and an even smaller eight percent of presidents and vice presidents were

women, underscoring the levels of underrepresentation of women in STEM in the Chinese workforce (Gu 2).

The underrepresentation of women in STEM does not originate in the workforce, but leads back to the Chinese education system. In this system, students across China have their future academic pathways dependent upon the major college entrance exam (deemed the “gaokao”) in order for them to enter higher-ranked universities, which open up better job opportunities (He et al. 1). When high school students begin their high school experience, they initially begin by taking every course subject: math, English, language (Mandarin), physics, chemistry, biology, politics, geography, and history. In their second year of high school, Chinese students must make one of the largest decisions of their lives and choose between the STEM education pathway, focusing on physics, chemistry, and biology, or the social science pathway, focusing on history, politics, and geography. The pathway chosen by students determines the materials they would ultimately be tested on in their college entrance exam and also the majors with which they apply to schools (He et al. 1). As these students are still young, their choices are heavily dependent on and influenced by societal expectations, which can impact their career perception and STEM self-efficacy (He et al. 3-4). STEM self-efficacy, which is the confidence that one has in their ability to pursue and to be able to resolve problems related to STEM subjects, is reliant on the social factors regarding an individual (Wang et al. 2-5) and is a major factor that shapes the decisions of female high school students in China when trying to choose their academic pathway (Chan 11). Negative experiences may deter students—especially women— from STEM pathways where some of these experiences stem from the Confucianist ideology (Doucette and Singh 3-4).

The concept of Confucianism is one of the major factors impacting familial expectations and the education system (Fang 212-14). Confucian teachings of filial piety and ideals of natural orders were passed down for many generations and the education system was remodeled upon it (Fang 213-14). Filial piety follows the values of being obedient and devoted and caring towards one’s parents and elderly family members under the basis of individual moral conduct and social harmony (Fang 212-13). In the Confucian belief system, upholding the natural order was what maintained peace and harmony in society, and this is done by respecting filial piety (Fang 212). To fulfill filial piety, this often meant following the yin-yang gender roles where women focused on the internal aspects of the family such as household roles, while men focused on the external aspects of the family such as earning money for the family (Gu 4). These gender roles from filial piety are also often applied to STEM-related fields. Since the STEM field is associated with being more assertive and competitive, qualities often associated with males under Confucianist ideals, female students are discouraged from pursuing STEM. On the other hand, female students are thought to be gentle, soft, and emotional, which are qualities that are associated with the humanities field, hence reducing female students’ STEM self-efficacy and ability to picture themselves within a STEM career.

In addition to Confucianism’s impact on familial expectations, Confucianism also lays as the foundation for the Gaokao exam, a pinnacle exam in shifting the outlook of a Chinese local

student's life, which has inherent biases against female students in the testing system. The Gaokao exam, which was made based on the old civil service exams, highlights the concept of merit-based success enabling greater social mobility (Fang 213). The old civil service exams tested scholars on their ability to retain knowledge through the several-day exams and only through good performance could poor people be able to join the educated elite. Similarly, the Gaokao exam also tests high school students on their retention of all the knowledge they learned throughout their high school career. Only by attaining a high score on the Gaokao exams can students apply for the highest-tier schools in the nation and eventually have the upper hand in earning higher-paying job opportunities quickly. Since the Gaokao is a high-stakes exam, this makes the exam biased against female students, as previous studies have shown that female students were not able to perform as well in high-stress level environments in comparison to their male counterparts due to factors of lack of confidence and greater societal pressures (He et al. 6).

Additionally, depictions of the role of women in Chinese textbooks often highlight women in domestic roles which are underscored by filial piety. Women are rarely shown taking on diverse management roles in comparison to their male counterparts, reinforcing gender stereotypes in the workforce and limiting career perception for female students (Fang 214). This then extends to the workforce where the stereotypes from the education system and from societal values have limited many women from challenging themselves to take on higher positions, thus limiting their mobility (Gu 2). Since female students are seeing these images over the course of their childhood and young adulthood, this degrades the ideas about themselves and limits the possibilities that they see their future selves being able to take on.

With the rise of social media platforms, adolescents are often caught in the crossfire of adults who try to emphasize different societal expectations, shaping the perspectives of these adolescents in their career perceptions. Zhihu is one of the social media platforms on the rise, where users can ask questions regarding any topic and have responders ranging from other students to experienced doctorates sharing their opinions through responding to the questions. Between 2021 and 2023, the monthly active users on the platform increased from 100 million users to 110.5 million users, highlighting the platform to be frequently used by students who seek guidance in choosing their academic pathways for high school (Statista). In order to investigate the specific ways in which female high school students in China are being pressured against choosing STEM pathways, this work aims to study the impact of social media platforms on a Chinese female high school student's STEM self-efficacy and career perception. By employing critical discourse analysis through analyzing Zhihu responses under the lens of the referential, predicational, and argumentation strategies, this study reveals Zhihu responses having a heavy emphasis on environmental stereotypes - gender expectations created through social media, families, teachers, peers that are depicted as the defined image of what each gender is meant to do- relating to male students having more inherent talent than female students when it comes to STEM.

The research question for this paper is: How does the reinforcement of traditional Confucianist Yin-Yang gender roles through Zhihu, a Chinese social media platform, impact female high school students' STEM self-efficacy and career perception in China?

Literature Review

In recent years, many studies have been conducted to discover the most critical points in a female student's journey in academia to improve female students' STEM identity. In particular, Hazari et al. collected survey data from undergraduate female students in trying to find the time and sources of recognition that enabled persistence in physics interests. The study pinpointed high school as the key time for female students to gain interest in a physics identity (Hazari et al. 98). The study discovered the largest gain of 31.6 percent in interest in a physics career by female students between the beginning of high school and the beginning of college (Hazari et al. 98). The data also reveals how recognition by high school teachers for female students displays the largest increase in self-recognition (Hazari et al. 98-99). This highlights an opportunity for female students to be distinguished from their peers because of physics identities in high school, further underscoring high school as a critical point in time to engage in intervention to increase STEM self-efficacy among female high school students (Hazari et al. 99).

In the case of Chinese female high school students, the need for support from high school teachers to encourage STEM self-efficacy is ever more important considering the choice students need to make early on regarding taking the STEM or the humanities pathway (He et al. 6-7). In a study that collected qualitative data through the form of one-on-one semi-structured interviews with six female high school students enrolled in the social science stream, the six students touched upon the limited guidance from higher figures in choosing a pathway, resulting in reduced STEM career perception due to the stereotypes in job opportunities portrayed across media and reinforced in societal norms which highlighted filial piety (He et al. 7). Additionally, lack of ability to perform as well as their male counterparts in high-stress exams that determine their future also depletes science interest and STEM self-efficacy among these female students as they find it bleak in trying to improve their STEM test scores to be able to stand out when taking the college entrance exam (He et al. 5-6). Overall, there are various factors that push female high school students in China away from the STEM pathway.

Societal norms, specifically Confucianism in the case of Chinese students in particular, have been a major factor in decreasing STEM self-efficacy where a study of various secondary schools across China highlights that while in general, 37.5 percent of female students had STEM self-efficacy in comparison to the 60.2 percent of male students who had STEM self-efficacy, only 19.8 percent of the female students wanted to pursue STEM careers in the future while 49.5 percent of male students wanted to pursue STEM careers (Chan 7). The data was also conclusive that female students who had high levels of endorsed traditional gender roles had lower STEM self-efficacy although conforming to the familial and social norms themselves had no significant impact on gender differences in STEM (Chan 8). Thus, this underscores how because of these

societal norms, this has reduced confidence in enabling female high school students to pursue STEM careers.

In another survey conducted by Wang et al. that collected data from 10th-grade students in five schools in the Hunan province, the study discovered that while both male and female students perceived school education as the environmental factor was the least impact on student interest in STEM careers, the greatest environmental factor that impacted female students' interest in STEM careers was the media while for male students' was social support (13). However, the study does not delve deeper into the aspects of social media itself that impacted female students as it only stopped at the surface level, highlighting the need for further investigation on the specific effect of social media on Chinese female high school students' STEM career perceptions.

The few studies on the impact of social media on reinforcing harmful gender stereotypes in STEM have focused on female PhD students. Like Quora, on the platform Zhihu users can ask questions. One notable study was focused on the LM incident where a male PhD attempts suicide after breaking up with an intimate relationship with his girlfriend, a female PhD student, sparks fire across the internet (Peng et al. 2). By focusing on the stereotyping of accomplished women through responses of Chinese netizens on Zhihu, the study revealed how positions of female professionals were seen as subordinate to men and how the stereotypes undermined female PhD students' accomplishments, highlighting how social media responses were able to reinforce many gender stereotypes (Peng et al. 2-3, 21). Building on these previous studies, this study contributes to the growing literature on the impact of social media on women's participation in STEM fields in China. It does so by investigating the impact of social media platforms under the lens of female high school students in choosing STEM pathways as a starting point to intervene with the lack of gender diversity in STEM careers in China.

Theoretical Framework

The social cognitive career theory model is a model that examines how the dimensions of individual characteristics, psychological factors, educational environment, and family link to an individual's career perceptions (Lent and Brown 11-12). The social cognitive career theory is often used as a model to investigate gender stereotypes in its impact on education such as its usage in investigating "Gender differences in high school students' interest in STEM careers" (Wang et al. 4-5). Due to the complexity of the model, this often results in the different dimensions intertwining with one another, tying together self-efficacy, STEM identity, societal statuses, gender role socialization, and outcome expectations as factors that all impact students' career perceptions (Fouad and Santana 26-27). Since the initial development of the model, many educators have used the model to evaluate how performance and environmental support enable STEM self-efficacy and positive outcome expectations which increased interest in STEM subjects among students (Fouad and Santana 34-35). Having an interest in STEM subjects increases satisfaction among students in learning STEM which predicts higher levels of persistence in students in pursuing STEM fields. However, because of negative environmental

factors such as lack of support from society and gender role socialization, many minority groups such as female students do not feel recognized as capable and lack a STEM identity, inhibiting persistence in STEM careers such as physics or engineering (Lent et al. 28-29). Hence, it is vital to evaluate the individual components that make up the environmental factors in order to investigate their impacts under the social cognitive career theory.

Due to gender role expectation stereotypes under Confucianist Yin-Yang ideals, this results in differences in acceptable behaviors for different gender roles. Since social roles link individuals to social environments, this causes society to shape common patterns of expectations under the social role theory, contributing to the internalization of social expectations and status and impacting gender role socialization. This is because shared gender stereotypes are stimulated through the gender division of labor which creates differentiated skills. Since men are often associated with characteristics of being assertive, competitive, and dominant while women are associated with characteristics of being warm and emotionally expressive (communal), this causes the different genders to have different expectations. In the case of Chinese society, as Confucianism was what shaped the social expectations in societies, this creates set expectations for the different genders and the different characteristics associated with different pathways for jobs and studies.

Moreover, derailing from the gender stereotypical expectations, especially for females, such as choosing to not be a housewife, often results in societal prejudice by peers and family. These prejudices can be taken through reactions of females being perceived as less favorable than their male counterparts when it comes to leadership positions which leads to attitudes less favorable for successful women to grow. This ties to female high school students when they do not follow the social expectation for them to become teachers or follow topics in the humanities. As a female, women often face prejudice in their initial salaries and injustices in the workforce that further discourage their involvement in STEM (Gu 2-3).

Methodology

In this paper, critical discourse analysis was used to conduct a qualitative study of the number of posts on the Zhihu platform. Critical discourse analysis was founded by Fairclough and Wodak, combining a theoretical approach with a textual analysis method that integrated language with contextual usage to critically analyze meaning (Peng et al. 6-7). This theory was developed by assuming that discourse structures that govern and shape human interaction (KhosraviNik and Sarkhoh 7), are the foundation for people to exercise power in our present-day society. The widely accepted critical discourse analysis method utilizes a three-dimensional framework that looks at language from a textual, discursive, and socio-cultural level. In this study, this method will be used to analyze the content of 21 responses to Zhihu questions through only the textual level.

Since the social media platform used in the study was a Chinese social media platform, all responses were translated into English to be analyzed. For the original Mandarin statement, those interested can reach out to the corresponding author listed on the paper. In this study, the

data was first chosen through selecting questions that came up that asked about choosing between social sciences or STEM for their high school pathway on Zhihu by high school students. Questions were then categorized according to whether they were written in a negative or neutral tone where the questions either reinforced or showed no bias toward gender stereotypes in choosing the pathway. Examples of negative questions included ones that had explicit bias of gender stereotypes such as “Why are girls not as good at science? Is it because we are stupid?” On the other hand, examples of neutral questions included ones that did not have any implicit bias towards any gender stereotypes, asking “Should girls study social sciences or STEM?” As a majority of the questions showed a neutral bias towards gender stereotypes, 76 percent of the responses were collected from neutral questions to create a representative sample.

As there were between 20-60 responses under each question, a representative sample size of five to seven responses was chosen from each question to show the overall reinforcement or challenging of stereotypes by Zhihu commenters for a total of 21 responses. One group of commenters that were excluded from the study were responders who simply copied academic papers word-for-word in a translated format and only included conclusions provided by the paper rather than personal opinions. As research papers contain a more objective tone due to formal language usage in comparison to social media language which includes more slang, therefore the responses that copied academic papers would not be able to accurately highlight the lexical choices of regular social media commentators since the posts did not originate from everyday language. Hence, the responses were excluded from the study.

For each response in the sample, the researcher recorded the word count for terms on the vocabulary list that related to gender, career choices, bias, or STEM words in accordance with the referential strategy. Then, each response was analyzed through the lens of the argumentation strategy to uncover the themes applied to each response. The themes stemmed from groups of words in the vocabulary list that related to common social issues with themes of “inherent vs. acquired talent”, “environmental stereotypes”, “effort vs. reward”, “equality”, “interest rather than expectation”, “career pathways”, and “family background.” While the theme of “inherent vs. acquired talent” involved phrases such as “logical thinking”, “spatial imagination”, “inquired”, and “inherent talent”, the theme of “family background” included the terms “employment/job”, “social environment”, and “tomboy/masculine.” The last level of analysis involved analyzing the responses through the predicational strategy to understand the type of impact that these Zhihu responses may have. This study focused on two types of impacts: The challenging or reinforcement of the stereotype of female students being inherently weaker at STEM than male students.

Results and Analysis

Referential Strategy

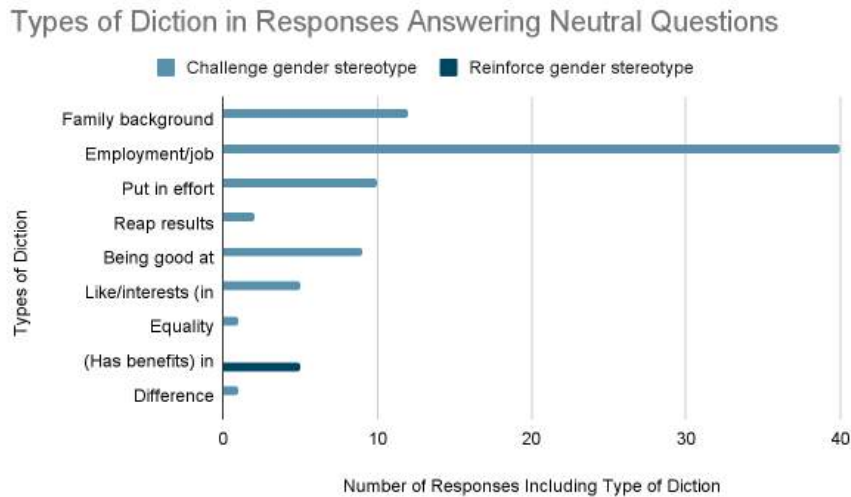


Figure 1: Word count of lexical choices of Zhihu responses for Neutral Questions that challenged gender stereotypes vs. reinforced gender stereotypes

To highlight the gendered references in the text in order to uncover the stereotypes that are being reinforced or challenged, the referential strategy was applied to the Zhihu responses. As seen in Figure 1, the lexical choices were split based on whether or not those lexical choices were included in posts that reinforced or challenged gender stereotypes of female students being inherently weaker in STEM than male students, which will be analyzed in a further section. As the responses analyzed were chosen from both questions that were more neutral and also questions that had negative connotations towards gender stereotypes, the analysis in this study was also conducted based on the two different types of questions: For neutral questions, the responses that answered neutral questions talked about inherent qualities in relation to family background rather than gender and its impact on employment as seen by the high word counts in employment (40) and family background (12). Moreover, responses that answered neutral questions also highlighted lexical choices of “putting in effort” (10) in order to “reap results” (2), “being good at something” (9), and STEM being an individual’s “likes/interests”(5). These responses to neutral questions highlighted the concept of merit-based and interest-based decisions for choosing STEM, challenging the gender stereotypes and tying the decision between STEM and social science to non-gender related factors. Despite the responses mainly focusing on lexical choices that challenged the gender stereotype, some of the responses still tied choosing the STEM pathway to lexical choices often associated with the reinforcement of gender stereotypes: Many of the responses discussed female students having to consider a lower-intensity but high salary job to have kids or to have benefits when going on blind dates (5),

though the responses that mentioned this were of seemingly less quantity than that of responses regarding family background or employment.

Types of Diction in Responses Answering Negative Questions

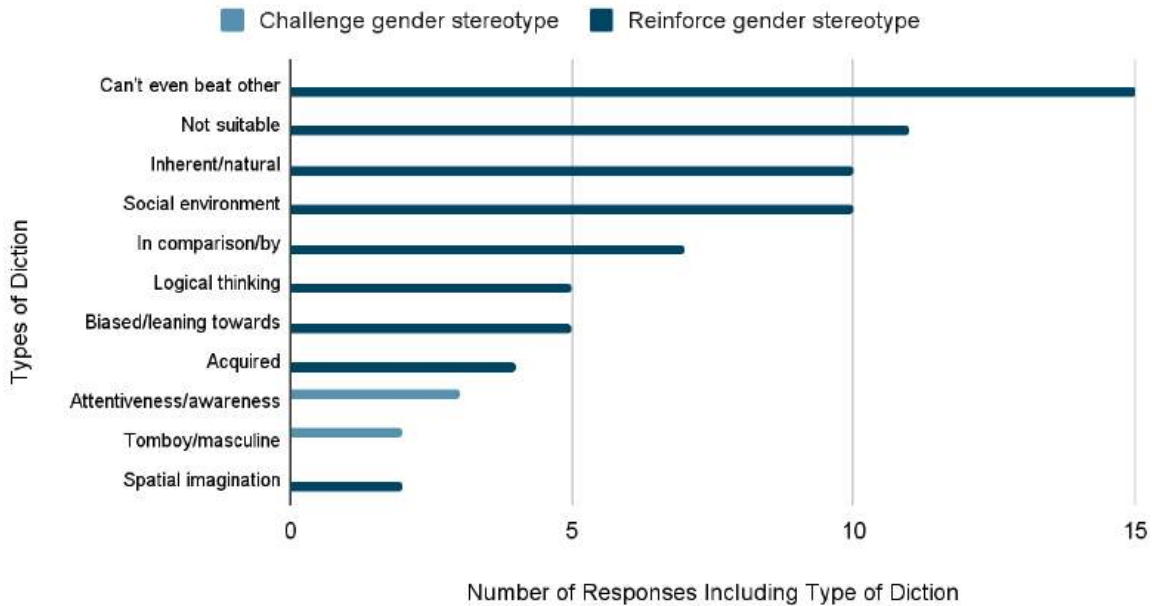


Figure 2: Word count of lexical choices of Zhihu responses for Negative Questions that challenged gender stereotypes vs. reinforced gender stereotypes

For negative questions, 80% of the responses that answered those questions reinforced the stereotypes. All of the responses to negative questions touched upon the idea of inherent qualities where the responses tied strong logical thinking (5) and spatial imagination (2) as inherent qualities (10) that male students had which were stronger than female students, quoting many girls to be seen as not suitable (11) for STEM due to their lack of inherent qualities. There was a large repetition of negative diction which are words with negative connotation that emphasize the stereotype of male students being more suited to choosing STEM than female students due to the possession of inherent talent that enforces bias and gender stereotypes. Since the majority of the responses replying to the negative questions reinforce the gender stereotype, thus the overtly negative diction has negative impacts on the decisions of female students when having to choose between STEM and social sciences (Felmlee et al. 25).

Predicational strategy

By analyzing the general stance of each post with stating the stance as reinforcing or challenging the gender stereotype of female students being inherently weaker than male students, the data highlights that 38.1 percent of the responses reinforced the gender stereotype whereas 61.9 percent challenged the gender stereotype. This means that overall, the responses that

Chinese female high school students come across mainly challenge the gender stereotype. However, there is still a large portion of responses that may negatively impact their decisions in choosing STEM or social sciences. Through previous research conducted by other researchers on the impact of the frequency of key derogatory terms used against women through social media platforms like Twitter on forming cyber aggression, the data highlighted that higher frequency usage of these terms enhances feminine norms and stereotypes (Felmlee et al. 17-18). Hence, on Zhihu, a similar case of repetition of responses that reinforce gender stereotypes may end up reinforcing the gender stereotypes that men are inherently better than women in STEM despite a majority of responses that challenged the gender stereotype. This is because there was enough repetition of responses within both neutral and negative questions that had inherently reinforced the gender stereotype which does not account for the questions themselves that include biases in the question which reinforces gender stereotypes, especially for negative questions.

Argumentation strategy

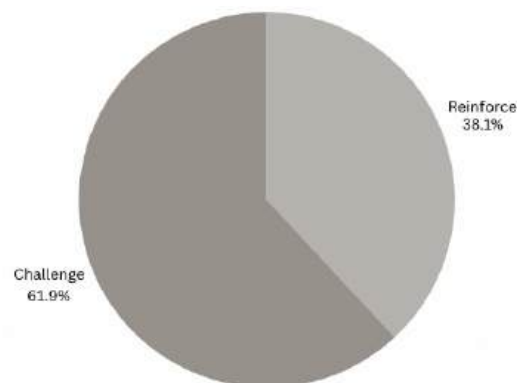


Figure 3: Percentage of Responses Challenging vs. Reinforcing Gender Stereotype of Female Students Being Inherently Weaker than Male Students at STEM

In all 21 of the responses, each of the posts touched upon at least one of the seven themes: inherent vs. acquired talent, environmental stereotypes, effort vs. reward, equality, interest over expectation, career pathways, or family background. Despite the lower percentage of responses “challenging” stereotypes in the predicational section of the study, due to the high level of references to the theme of environmental stereotypes, this study can conclude that the biases prominent in daily life are further accentuated through the usage of social media (Felmlee et al. 16). By repeatedly referencing environmental stereotypes, this suggests that the environmental stereotypes related to gender differences in STEM are heavily prompted and suggested. As the majority of the users who propose the questions have backgrounds that highlight the users as high school students, this means that high school students are engaged with these gender stereotypes often which further enforces the stereotypes the students heard in a school background and makes students hesitant to choose STEM.

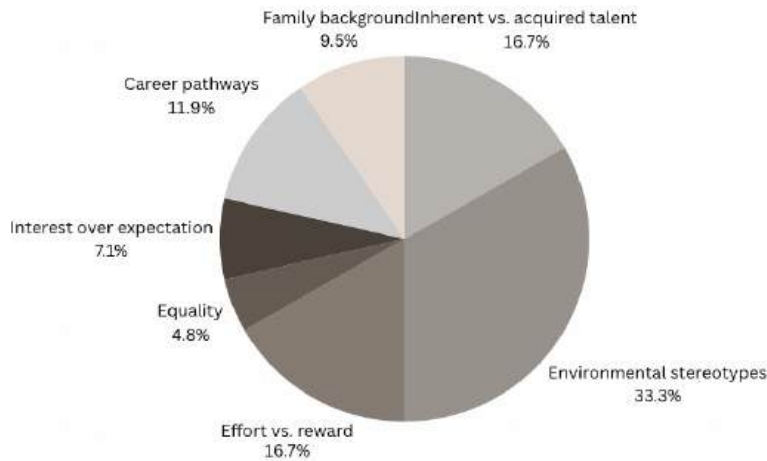


Figure 4: Percentages of Each Theme Mentioned in All Responses

Conclusion

Under the social cognitive career theory, in order to increase the involvement and persistence of females in STEM careers, especially with physical sciences, it is pertinent to start early in engaging female students to become interested and develop self-efficacy in their math and other STEM abilities. As social media is highlighted as a significant factor in previous research papers on impacting the STEM self-efficacy of female high school students, the responses of users to the questions posted on Zhihu are an important area to study. Although the majority of the responses challenging gender stereotypes in STEM fields, there is a heavy emphasis on environmental stereotype themes and references to male students having more inherent talent suited to STEM than that of female students. Not only does this further enforce the gender stereotype for female students, but it also deters them away from choosing STEM pathways. This study highlights the need to further investigate a direct correlation between the comments on Zhihu with the eventual pathways of female high school students in their choice of majors in college.

Works Cited

- Chan, Randolph C. H. "A Social Cognitive Perspective on Gender Disparities in Self-Efficacy, Interest, and Aspirations in Science, Technology, Engineering, and Mathematics (STEM): The Influence of Cultural and Gender Norms." *International Journal of STEM Education*, vol. 9, no. 1, Dec. 2022, p. 1-10. <https://doi.org/10.1186/s40594-022-00352-0>
- Doucette, Danny, and Chandralekha Singh. "Why Are There So Few Women in Physics? Reflections on the Experiences of Two Women." *The Physics Teacher*, vol. 58, no. 5, May 2020, pp. 1-5. <https://doi.org/10.1119/1.5145518>
- Fang, Xiran. *Influence of Confucianism on Gender Inequality in Chinese Education and Employment: 2021*. pp. 212-15. <https://doi.org/10.2991/assehr.k.210617.066>
- Felmlee, Diane, et al. "Sexist Slurs: Reinforcing Feminine Stereotypes Online." *Sex Roles*, vol. 83, no. 1–2, July 2020, pp. 16–28. <https://doi.org/10.1007/s11199-019-01095-z>
- Fouad, Nadya A., and Mercedes C. Santana. "SCCT and Underrepresented Populations in STEM Fields: Moving the Needle." *Journal of Career Assessment*, vol. 25, no. 1, Feb. 2017, pp. 24–39. <https://doi.org/10.1177/1069072716658324>
- Gu, Chao. "Women Scientists in China: Current Status and Aspirations." *National Science Review*, 9 June 2021, pp. 1-5. <https://doi.org/10.1093/nsr/nwab101>
- Guo, Congbin, et al. "Gender Disparities in Science and Engineering in Chinese Universities." *Economics of Education Review*, vol. 29, no. 2, Apr. 2010, pp. 225–35. <https://doi.org/10.1016/j.econedurev.2009.06.005>
- Hazari, Zahra, et al. "The Importance of High School Physics Teachers for Female Students' Physics Identity and Persistence." *The Physics Teacher*, vol. 55, no. 2, Feb. 2017, pp. 96–99. <https://doi.org/10.1119/1.4974122>
- He, Lizhi, et al. "Female Underrepresentation in STEM Subjects: An Exploratory Study of Female High School Students in China." *EURASIA Journal of Mathematics, Science and Technology Education*, vol. 16, no. 1, Sept. 2019. pp. 1-7. <https://doi.org/10.29333/ejmste/109657>
- KhosraviNik, Majid, and Nadia Sarkhoh. "Arabism and Anti-Persian Sentiments on Participatory Web Platforms: A Social Media Critical Discourse Study." *International Journal of Communication*, vol. 11, no. 0, Sept. 2017, pp. 7. <https://ijoc.org/index.php/ijoc/article/view/6062>
- Lai Lin, Thomala. "Main Usage Figures of the Chinese Quora-like Social Media Site Zhihu between 2021 and 2023(in Millions)." *Statista*, 21 Feb. 2024, <https://www.statista.com/statistics/1223829/china-key-usage-figures-of-zhihu/>
- Lent, Robert W., et al. "Social Cognitive Predictors of Adjustment to Engineering Majors across Gender and Race/Ethnicity." *Journal of Vocational Behavior*, vol. 83, no. 1, Aug. 2013, pp. 22–30. <https://doi.org/10.1016/j.jvb.2013.02.006>
- Lent, Robert W., and Steven D. Brown. "Social Cognitive Career Theory at 25: Empirical Status of the Interest, Choice, and Performance Models." *Journal of Vocational Behavior*, vol. 115, Dec. 2019, pp. 10-12. <https://doi.org/10.1016/j.jvb.2019.06.004>

- Peng, Altman Yuzhu, et al. “‘She Uses Men to Boost Her Career’: Chinese Digital Cultures and Gender Stereotypes of Female Academics in Zhihu Discourses.” *Social Semiotics*, vol. 33, no. 4, Aug. 2023, pp. 2-21. <https://doi.org/10.1080/10350330.2021.1940920>
- Wang, Ning, et al. “Gender Differences in High School Students’ Interest in STEM Careers: A Multi-Group Comparison Based on Structural Equation Model.” *International Journal of STEM Education*, vol. 10, no. 1, Oct. 2023, p. 1-20. <https://doi.org/10.1186/s40594-023-00443-6>
- Wodak, Ruth. “Argumentation, Political.” *The International Encyclopedia of Political Communication*, edited by Gianpietro Mazzoleni, 1st ed., Wiley, 2016, pp. 1–9. <https://doi.org/10.1002/9781118541555.wbiepc080>

Imperfect Yet Impactful Reforms: How Educational Reformers Successfully Transformed the US Public Education System in the Progressive Era By Jake Kangas

I. Introduction

In the late nineteenth century, US cities grew at an extraordinary rate as industrial expansion attracted millions of immigrants and rural Americans to urban centers. Many found themselves living in slums characterized by substandard housing, unsanitary conditions, and social disorganization. Residing in dire conditions without an education or a path toward a promising future, many immigrants and lower-class families felt discouraged and powerless. For them, the promise of an education system that could supply new opportunities for their children provided hope.

During the Progressive Era, which lasted from 1880 and 1920, education reformers sought to make this educational promise a reality. Driven to correct the moral failings and inequality of the preceding Gilded Age, and responding to needs created by immigration, industrialization, and the growth of a corporate economy, the reformers believed the public education system could be a vehicle for reducing inequality and helping the US maintain social cohesion during a time of great change. Through their efforts, the reformers aspired to create a public school system that would preserve America's democratic ideals, address the welfare of children and other marginalized groups, and provide opportunities for all young people to build productive futures.

This paper discusses several significant changes education reformers made to the US public education system during the Progressive Era. It argues that education reformers were successful in making these changes, despite being constrained by oppressive race, gender, and class systems, for two key reasons. First, their reforms addressed important societal needs at the time. Second, the reformers were able to garner widespread support from a diverse group of Americans. Although the reforms were imperfect, they transformed the US education system from narrowly tailored schools for the elite to a multifaceted system for the masses, providing important insights about how our modern school system developed and uncovering reform strategies that current education reformers can learn from.

II. Educational Reformers Significantly Changed the US Public School System

During the Progressive Era, education reformers successfully enacted three significant changes to the US public school system, which fundamentally altered the landscape of American education. First, the reformers dramatically expanded public secondary education. Before the progressive movement, high schools catered to college-bound students from wealthy families and only a small portion of Americans were educated beyond the eighth grade (Beston). Progressive leaders believed the limited secondary school system was outmoded and that the modern world necessitated universal education. Thus, they fought for expansion and between 1900 and 1920, the number of public high schools doubled, public school enrollment quadrupled, and public high schools shifted from institutions for upper-class white men to schools for all

Americans (Beston). By creating a system that enrolled students on a mass scale, progressive education reformers forever changed American notions of whom school should be for.

Additionally, education reformers implemented a new pedagogic approach developed by philosopher John Dewey that moved away from “outmoded ideas of discipline and didactic exercises” and instead focused on preparing students for democratic participation (“Progressive Education”). Embedded within the larger progressive movement that was invested in creating “a new moral basis for American society,” the reformers transformed the philosophical foundation of American education at all levels to focus on cultural transmission, socialization, and personal and community growth, which they believed were essential for active civic engagement (Emirbayer, 623). With schools organized as “embryonic societ[ies],” students learned to interact and collectively solve problems, creating “the best guarantee of a larger society which is worthy and harmonious” (Dewey, 32).

Lastly, reformers secured vocational education within public high schools, expanding the curriculum to include skills-based courses such as typing, metal shop, and home economics (Beston; Kreysing, 29). The reformers fought ardently against the classical academic curriculum, which they considered elitist because it focused solely on the needs of those preparing for college. Believing their mission was to give all children the training they needed for their diverse futures, the reformers prevailed over traditionalists, instituting diverse curricula that resembled the offerings in today’s US high schools.

While each of these changes was significant in itself, together, they revolutionized the US public school system and set the stage for the modern school system in America. Making sweeping changes to the existing education system was not easy. However, as detailed below, the reformers developed a strategy that enabled them to succeed.

III. Educational Reforms Addressed Important Societal Needs At the Time

One key reason Progressive Era education reformers succeeded was that they specifically tackled pressing needs that stemmed from changes to the US economic and social systems. To illustrate, during the late nineteenth and early twentieth centuries, industrialization and corporate growth created a need for more educated workers, both in rapidly growing professional jobs, and in factories, where employers were increasingly seeking high school graduates (Beston). According to Harvard economist Claudia Golden:

Most Americans had traditionally worked in jobs that required little schooling... [but] in 1920, more than 25 percent of American workers held jobs that required a high school diploma (Beston).

By expanding secondary schools, the reformers filled this need.

The reformers also addressed societal needs stemming from immigration. Between 1880 and 1919, more than 23 million immigrants arrived in America, coming mainly from Southern and Eastern European countries that typically had different cultures, religions, and languages

than the US (Edson, 66). Blinded by the prejudices of the day, Americans wanted immigrants to assimilate into American society, largely because they believed assimilation was essential to protect US democratic values and maintain a culturally cohesive society (Dewey, 22; Thornton; Zervas, 79). Thus, “Americanizing” immigrants became a top priority, and expanding public high schools that provided mandatory English language and civics classes, patriotic exercises, and pedagogy focused on democratic participation filled a perceived need.

Child labor, which escalated with industrialization, was also a significant problem in the Progressive Era. High school expansion and vocational curriculum that reduced dropout rates aligned with society’s need to protect children from dangerous, exploitative jobs by keeping them in school. While poverty made it impossible for some children to forgo their jobs, tempering the reformers' success in this area, the reforms nonetheless represented a layer of protection for some vulnerable children before Congress passed child labor laws (“In the Playtime”).

Finally, as high schools started including more diverse students, which decreased the proportion of high school students attending college, there was a “crying need” for a practical curriculum to supplement the exclusively liberal curriculum that prevailed before the Progressive Era (Graubard, 11; Ogden, 248). Incorporating vocational education into the curriculum met this need and helped keep non-college-bound students interested and enrolled in school. Together, these examples show how societal needs arising from specific social and economic conditions created an opening for educational reformers to successfully transform US public education in ways that expanded opportunities, protected marginalized children, and promoted national unity.

IV. Educational Reforms Garnered Widespread Support

Another reason the educational reformers succeeded was that they built a “coalition of disparate Americans” who supported the reforms, based on the promise of a better future, a cohesive democratic society, and a more educated workforce (Emirbayer, 621-664). Importantly, the reformers strategically labeled school reform as a “humanitarian initiative” and framed public schools as institutions that could transform economic opportunities and improve life for all Americans (Ewert, 66). This rhetoric tapped into Americans’ frustrations over the social problems and extreme inequality that had plagued America since the Gilded Age.

Native-born working-class Americans in particular came to believe the public school system was their best hope for securing new opportunities and leading a more prosperous life. Although many workers depended on the wages of their children, they supported expanding the public education system because public schools were a way of integrating children of the working class with wealthier children. Some working class Americans suggested that schools allow students to study part of the day and work part of the day: “In this way, working-class children could still contribute to the family income, but also receive an education which would allow them to reach a higher station in life” (Brackemyre).

For immigrants, support for education reforms was more complicated as some immigrants saw public schools as “a device for assimilation” that would extinguish their native

traditions (Zervas, 80). These immigrants resisted public schools and sometimes started separate institutions, like private Catholic schools, that better fit their cultural and religious needs (Zervas, 80). Still, many immigrants wanted to integrate their children into American life through the public schools, even as they hoped to preserve their native identity. Because the public schools represented hope for increased opportunities and inclusion in American life, these immigrants rallied around efforts to expand and reform public schools and, in the process, helped advance innovations like vocational curriculum (Celis; Udea, 666-667).

Support for educational reform also came from Americans who feared that immigrants would weaken American cohesion and democratic ideals. Seizing on this concern, the reformers used patriotic rhetoric and framed public schools not only as sources of new opportunities but also as institutions for “preserving America’s democratic heritage” (Ewert, 66; Cabán, 27). Through the reformers’ patriotic appeals, fearful Americans came to support public school expansion and pedagogy that emphasized democratic participation because they believed the reforms would protect America’s democratic national identity.

Finally, industry and business leaders also supported educational reform, especially school expansion and vocational education. As reformers emphasized the “efficiency of school-based job training,” business and industry executives, seeking a consistent influx of trained and educated workers, threw their support behind the reforms to meet their own interests (Page). Thus, groups like the National Association of Manufacturers and the American Federation of Labor joined educational reformers to promote change, even helping to pass the 1917 Smith-Hughes Act, which provided federal funding for vocational education in public secondary schools (Graubard, 11).

The reformers’ ability to utilize rhetorical strategies to rally such a disparate group of Americans around public school reform is notable. Together, the support of these disparate groups, who had diverse reasons for supporting the reforms, helped the reformers go beyond dreaming to successfully transform American public education in ways they believed would create a more equitable and cohesive democratic society.

V. Discriminatory Race, Class and Gender Systems Restrained, but Didn’t Stop, the Reformers

Despite support from this broad array of Americans and the fact the reforms filled critical societal needs, the reformers had to work within discriminatory race, class, and gender systems, which tempered, but did not erase, their success. Because education is a gateway to higher-paying jobs, education reformers hoped their reforms would markedly reduce inequality. In the end, the reforms produced a sharp rise in human capital; however, they could not overcome society’s deeply rooted oppressive social systems, which relegated women to homemaking roles, segregated schools by race in the South, and perpetuated discrimination in employment, making anything close to economic parity for women or underrepresented groups impossible (Goldin; “Learning From”; Odum, 185-186). That economic injustice and inequality remained prevalent underscores limits to the reformers’ success. Still, the education reforms took

necessary first steps in improving economic conditions by creating new opportunities. Not only did high school expansion increase access to higher education, especially for women, but it also created new job options, and helped decrease poverty as earnings increased with each year of high school, even in blue-collar jobs (Wells, 731-732; Chamberlain; Goldin and Katz, 783). Thus, while the education reformers' lofty vision of dramatically decreasing inequality was unattainable, their vision ignited changes that moved the country in the right direction.

VI. Conclusion

Bounded by oppressive societal systems that hindered social mobility and perpetuated inequality, Progressive Era educational reformers successfully implemented imperfect but significant reforms to the public school system by addressing important societal needs and garnering support from a diverse array of Americans. During the Progressive Era, America faced weighty challenges related to a changing labor market, immigration, child labor, and an increasingly diverse high school population. By specifically aligning their reforms to address these issues, the reformers helped ensure their reforms would succeed. The reformers also prevailed because they amassed widespread support from diverse groups through rhetoric that tapped into concerns about future prosperity, a cohesive democratic society, and a properly educated labor supply.

Today, education reformers are again arguing that the US public education system needs sweeping reform. Issues like common core curriculum, standardized testing, and achievement gaps are hotly debated, and US education rankings are falling globally (Amadeo). If twenty-first-century reformers hope to substantially change the existing system, they would be wise to learn from their Progressive Era predecessors, aligning their reforms with current societal needs and rhetorically framing their reforms to elicit broad support. This will be challenging given today's sharp political divisions and reduced confidence in US public institutions. Still, because education is linked to better health, political and civic engagement, and long-term prosperity, the challenge is worth taking (Nazzaro).

Works Cited

- Amadeo, Kimberly. "U.S. Education Rankings Are Falling Behind the Rest of the World." *The Balance*. Last modified March 26, 2023. <https://www.thebalancemoney.com/the-u-s-is-losing-its-competitive-advantage-3306225>.
- Beston, Paul. "When High Schools Shaped America's Destiny." *City Journal*. August 1, 2017. <https://www.city-journal.org/article/when-high-schools-shaped-americas-destiny#:~:text=High%20school%20>.
- Brackemyre, Ted. "Education to the Masses: The Rise of Public Education in Early America." *U.S. History Scene* (2012, August 9). <https://ushistoryscene.com/article/rise-of-public-education/>.
- Cabán, Pedro. "Subjects and Immigrants during the Progressive Era." *Discourse* 23, no. 3 (2001): 24-51. <https://www.jstor.org/stable/41389614>.
- Celis, William. "The Forgotten History of Immigration." *Education Week*. October 3, 2006. <https://www.edweek.org/leadership/opinion-the-forgotten-history-of-immigration/2006/10>.
- Chamberlain, Mariam K. *Women in Academe*. New York, Russell Sage Foundation, 1988.
- Dewey, John. "The School and Social Progress." In *The School and Society: Being Three Lectures*, (Chicago, The University of Chicago Press, 1899), 19-44. https://www.gutenberg.org/cache/epub/53910/pg53910-images.html#Page_19.
- Edson, C. H. "Curriculum Change during the Progressive Era." *Educational Leadership* 36, no. 1 (1978): 64-69. https://files.ascd.org/staticfiles/ascd/pdf/journals/ed_lead/el_197810_edson.pdf.
- Emirbayer, Mustafa. "Beyond Structuralism and Voluntarism: The Politics of Discourse of Progressive School Reform 1890-1930." *Theory and Society* 21, no. 5 (October 1992): 621-664. <https://www.jstor.org/stable/657804>.
- Ewert, Cody Dodge. "Schools on Parade: Patriotism and the Transformation of Urban Education at the Dawn of the Progressive Era." *The Journal of the Gilded Age and Progressive Era* 16, no. 1 (2017): 65-81. <http://www.jstor.org/stable/26347243>.
- Goldin, Claudia and Lawrence F. Katz. "Education and Income in the Early 20th Century: Evidence From the Prairies." *The Journal of Economic History* 60, no. 3 (2000): 782-818. <https://www.jstor.org/stable/2566438>.
- Goldin, Claudia. "How America Graduated From High School: 1910 to 1960." Working Paper No. 4762. Department of Economics, Harvard University, June, 1994. <http://nrs.harvard.edu/urn-3:HUL.InstRepos:32785054>.
- Graubard, Allen. "Could Vocational Education Be Progressive?" *The Radical Teacher*, 69 (2004): 10-17. <https://www.jstor.org/stable/20710240>.
- "In the Playtime of Others: Child Labor in the Early 20th Century." *Smithsonian Learning Lab*. December, 1988. <https://learninglab.si.edu/collections/in-the-playtime-of-others-child-labor-in-the-early-20th-century/vCUnDaYDT4UCjrh3>.
- Kreysing, Matthias. "Vocational Education in the United States: Reforms and Results."

- Vocational Training: European Journal* 23 (May-August 2001): 27-35. https://www.cedefop.europa.eu/files/etv/Upload/Information_resources/Bookshop/232/23_en_kreysing.pdf.
- “Learning From the Progressive Era.” *Harvard University Press* (June 2, 2014). https://harvardpress.typepad.com/hup_publicity/2014/06/learning-from-the-progressive-era-david-huyssen.html.
- Nazzaro, Miranda. "Higher Education Levels Linked With Civic Engagement, Better Health: Gallup." *The Hill*. (August 30, 2023). <https://thehill.com/changing-america/enrichment/education/4179203-higher-education-levels-linked-with-civic-engagement-better-health-gallup/#:~:text=Those%20with%20higher%20education%20levels,of%20educational%20attainment%2C%20Gallup%20found>.
- Odum, Howard D. “Some Studies in the Negro Problems of the Southern States.” *The Journal of Race Development* 6, no. 2 (October, 1915): 185-191. <https://www.jstor.org/stable/29738122>.
- Ogden, William R. "Vocational Education: A Historical Perspective." *The High School Journal* 73, no. 4 (1990): 245-51. <https://www.jstor.org/stable/40364875>.
- Page, Hon. Carroll S. "Vocational Education." Speech, The Senate Of the United States, June 5, 1912.
- "Progressive Education." *Encyclopedia.com*. Last modified June 8, 2018. <https://www.encyclopedia.com/social-sciences-and-law/education/education-terms-and-concepts/progressive-education>.
- Thornton, Bruce. "America's Problem of Assimilation." *Hoover Institution*. May 24, 2012. <https://www.hoover.org/research/americas-problem-assimilation>.
- Ueda, Reed. "Second-Generation Civic America: Education, Citizenship, and the Children of Immigrants." *The Journal of Interdisciplinary History* 29, no. 4 (1999): 661-81. <https://www.jstor.org/stable/206978>.
- Wells, D. Collin. “Some Questions Concerning the Higher Education of Women.” *American Journal of Sociology* 14, no 6 (May, 1909): 731-739. <https://www.jstor.org/stable/2762548>.
- Zervas, Theodore G. "Finding a Balance in Education: Immigration, Diversity, and Schooling in Urban America, 1880-1900." *Athens Journal of Education* 4, no.1 (February 2017) 77-84. <https://files.eric.ed.gov/fulltext/EJ1219878.pdf>.

The Influence of Emotional Knowledge on Visual Perception: A Comprehensive Review of Facial Expression Interpretation By Chloe Lai

Abstract

Vision and memory are fundamental to cognitive function, operating synergistically to shape our perception and interaction with the world. The intricate visual process enables us to discern shapes, colors, depth, and motion. Concurrently, memory processes, orchestrated by brain regions including the hippocampus, encode and retrieve visual experiences. The integration of prior knowledge profoundly influences these processes, modulating how individuals interpret visual complexity and emotional expressions. Cultural backgrounds, learned social cues, and personal experiences collectively inform perceptual frameworks, shaping cognitive interpretations and responses. Moreover, emotional knowledge (EK) plays a pivotal role in understanding facial expressions and emotional cues, guiding interpersonal interactions and decision-making. EK, developed through experiences and cultural contexts, modulates the interpretation of emotional content conveyed through facial expressions, thereby enriching social dynamics and fostering empathetic responses. Aesthetic perception extends this framework by evaluating qualitative and affective aspects of stimuli beyond basic recognition. Influenced by cultural norms and multisensory integration, aesthetic experiences involve cognitive appraisal and emotional valuation, enriching our understanding and appreciation of sensory input.

Introduction

Vision and memory are the dream team of brain functions—inseparable partners in the intricate dance of cognition. While vision scouts out the terrain, memory holds the map (and occasionally misplaces it). Together these systems collaborate to accomplish such tasks as recognizing a friend in a crowd or recalling where you left your phone (again). From the vibrant hues of a sunset to the familiar contours of a loved one's face, these sensory impressions etch themselves into our consciousness, shaping how we perceive and interact with the world. This symbiotic relationship not only enriches our experiences but also underscores the profound role of visual perception in defining our memories and understanding of reality.

Vision is a complex process orchestrated by the brain, where light entering the eye is converted into electrical signals that travel through the optic nerve to the visual cortex at the back of the brain.¹ This journey involves the eye's lens focusing light onto the retina, which contains photoreceptor cells called rods and cones. These cells detect different wavelengths of light and send corresponding signals to the brain, where they are processed and interpreted.² The visual cortex integrates these signals, allowing us to perceive shapes, colors, depths, and motion, forming our visual perception of the world around us³ (Figure 1). This intricate system highlights the brain's remarkable ability to transform light into meaningful images and underscores the importance of vision in our daily experiences and interactions. Visual input travels through the eyes, to the optic nerve, through the optic tract, and makes its way through the optic radiation of

the brain before orienting and perceiving the two seen objects in the brain (Figure 1). This intricate process is what allows us to experience vision.

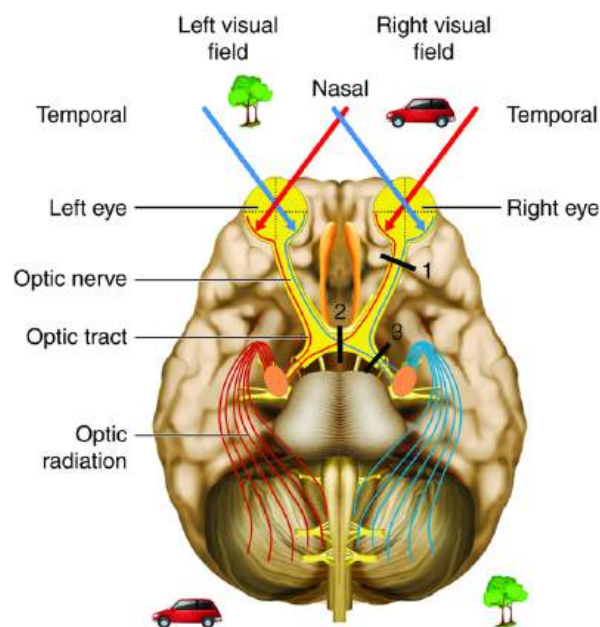


Figure 1 Visual Path From Eyes Through Brain. How vision is orchestrated and travels from our eyes and through our brains, accounting for the orientation of the objects we see.⁴

The visual process plays a pivotal role in cognition by providing essential sensory input that informs decision-making, learning, and understanding of the environment. As the brain processes visual information, it not only identifies objects and their spatial relationships but also integrates this data with memories, emotions, and learned associations.⁵ For instance, recognizing a familiar face triggers a cascade of cognitive processes involving memory retrieval and emotional response. Moreover, visual cues guide actions and facilitate complex tasks such as reading, driving, and spatial navigation.⁶ This interaction between vision and cognition underscores their interconnectedness, illustrating how the brain utilizes visual input to construct our understanding of the world and shape higher-level cognitive functions.

Vision and memory are intricately linked in the brain, with various regions working together to process and store visual information. The temporal lobes, particularly the hippocampus and surrounding structures, play a crucial role in forming new memories associated with visual experiences.⁷ When we perceive something visually, such as landmarks in a cityscape or faces of loved ones, the visual cortex processes these images and sends signals to the hippocampus for encoding into memory. These memories can be episodic (events) or semantic (facts), both of which heavily rely on visual details for retrieval and recall. Moreover, visual cues often trigger memory retrieval, where seeing an object or place can evoke associated memories and emotions.⁸ This dynamic relationship between vision and memory highlights the brain's ability to integrate sensory input with memory formation and retrieval, essential for navigating daily life and constructing personal narratives.

Perception is the brain's interpretation of sensory information received from our environment.⁹ It involves not only the physical detection of stimuli such as light, sound, and touch but also the cognitive processing that assigns meaning to these sensations. It is influenced by factors such as past experiences, expectations, and context, shaping how we understand and interact with the world. This complex process allows us to recognize patterns, form judgments, and make decisions based on the information our senses provide. Perception is essential for everyday tasks, from recognizing faces to navigating through space, illustrating its fundamental role in our ability to comprehend and engage with our surroundings.

Understanding human perception involves unraveling the intricate ways in which prior knowledge shapes our interpretations of visual and emotional cues. Prior knowledge, accumulated through diverse experiences and cultural backgrounds, plays a pivotal role in modulating how individuals perceive and respond to the world around them. Here, we explore the multifaceted influence of prior knowledge on the perception of visual complexity and emotional expressions, highlighting its profound implications across various contexts.

Emotional Knowledge (EK)

EK, defined as the ability to recognize, distinguish between, express, and monitor emotions, encompasses our understanding and awareness of emotions, both within ourselves and in others.¹⁰ It involves recognizing, interpreting, and effectively responding to emotional and social cues and situations. This knowledge is developed through experiences, upbringing, cultural influences, and personal interactions, shaping our ability to navigate social dynamics and relationships. Emotions are processed in the brain's limbic system, particularly the amygdala and prefrontal cortex, where they are integrated with cognitive functions¹¹ (Figure 2). This integration allows us to perceive emotions in ourselves and others, empathize with different perspectives, and regulate our own emotional responses. EK enables us to make informed decisions, build meaningful connections, and foster a deeper understanding of ourselves and the world around us, highlighting its pivotal role in our daily lives and interactions.

EK is intricately linked to the limbic system of the brain (Figure 2), a crucial network involved in processing and regulating emotions. The limbic system, which includes structures such as the amygdala and the hippocampus, plays a pivotal role in emotional experience and memory formation. This neural network helps us recognize, interpret, and respond to emotional cues, thereby shaping our emotional knowledge. For instance, the amygdala is essential for detecting and responding to emotional stimuli, while the hippocampus integrates these emotional experiences into long-term memory, influencing how we understand and interact with emotional contexts. This connection underscores the fundamental role of the limbic system in the development and application of emotional knowledge in our daily lives.

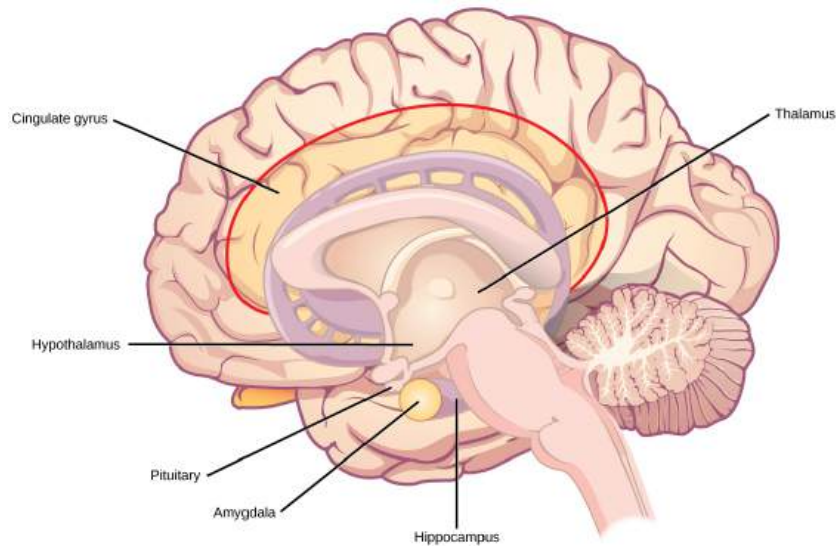


Figure 2 Limbic System of the Brain. Regions of the brain that are responsible for controlling emotions, behaviors, and motivations; these areas are also involved in memory and smell.¹²

Aesthetics alter visual perception

Aesthetic perception involves cognitive appraisal and emotional valuation, extending beyond basic perceptual recognition to evaluate qualitative aspects of stimuli. Cultural influences and multisensory integration further shape aesthetic perception, highlighting the complex interplay between emotional understanding, cognitive processing, and neural mechanisms involved in aesthetic perception.¹³

Aesthetic perception involves more than just recognizing basic features; it includes cognitive appraisal and emotional valuation. Cosmetics and certain facial features can influence aesthetics and visual perception by altering how we interpret and respond to visual stimuli (Figure 3). Cosmetics, through its use of color, contour, and enhancement, can dramatically change the perception of facial features, affecting both the aesthetic appeal and emotional expression conveyed. For example, the application of specific colors or shading techniques can enhance or diminish certain facial characteristics, influencing how observers perceive attractiveness, mood, and even personality traits. This interplay between makeup and visual perception highlights the role of aesthetic choices in shaping our cognitive and emotional responses to visual stimuli. By modifying facial appearance, makeup not only affects personal self-expression but also impacts how others interpret and interact with the individual, demonstrating the profound connection between aesthetics, visual perception, and social interaction.

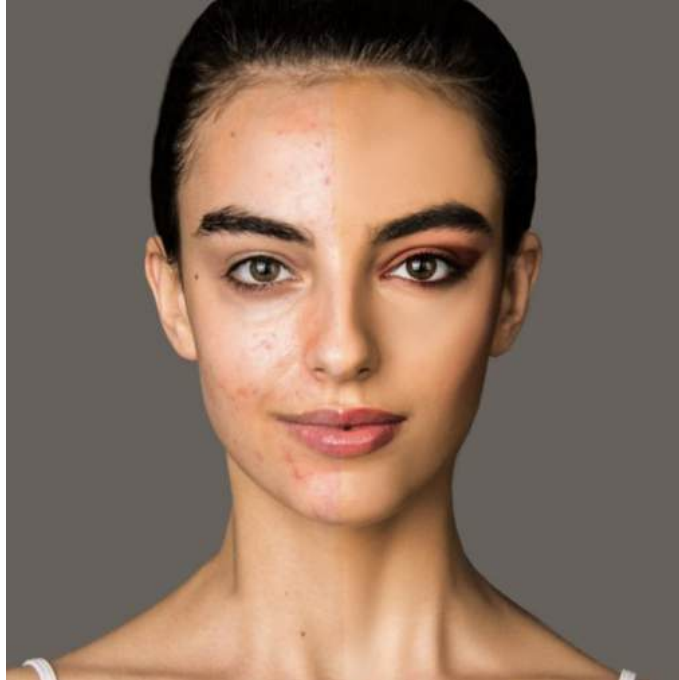


Figure 3 Altered Perception Based on Aesthetics. Half of the woman's face has applied makeup while the other does not, which influences thoughts and perspectives about this woman to outside viewers.¹⁴

Aesthetic perception, as distinguished from basic perceptual recognition, involves a cognitive process that evaluates qualitative and affective aspects of stimuli rather than their basic physical properties¹³. Aesthetic experiences can vary across sensory modalities, involving multisensory cues that influence our perceptions and behaviors. A novel local-global integrative model was also introduced, elucidating how aesthetic processing operates independently of basic perceptual processing but not independently of cognitive processing.¹³ EK guides how facial expressions are perceived aesthetically, enhancing the appreciation of genuine emotions like happiness compared to neutral or insincere expressions. This knowledge shapes the viewer's interpretation and evaluation of facial expressions, enriching the aesthetic experience through multisensory integration and cultural influences. EK is variable across cultures and individuals, affecting how people aesthetically perceive and emotionally respond to facial expressions.¹⁵ This variability underscores the complex interplay between emotional understanding, cognitive processing, and neural mechanisms involved in aesthetic perception.

Visual information may not have a correlation to emotion

Emotional perception involves multisensory integration, where auditory, visual, and contextual cues collectively contribute to understanding emotional meaning.¹⁶ This finding aligns with theories of embodied cognition, suggesting that emotional understanding requires simulating corresponding emotional states through sensorimotor processes.

Facial expressions, often considered universal indicators of emotions like happiness or sadness, exhibit considerable variability across contexts and individuals. Cultural norms and individual differences influence how facial expressions are perceived, challenging the simplistic notion of universal emotional displays.¹⁷ This variability underscores the dynamic and context-dependent nature of facial expressions in emotional communication.

While facial movements like smiles or frowns are commonly thought to universally convey emotions such as happiness, sadness, or anger, this belief lacks robust empirical support (Figure 4).

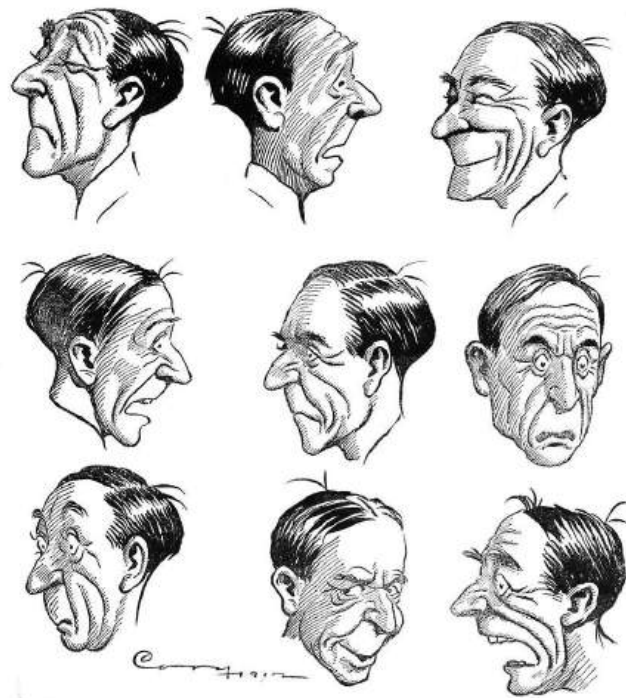


Figure 4 Same Face, Different Expressions. The same face is drawn with different expressions, aimed to induce assumptions about what emotions the man was experiencing.¹⁸

While the man is clearly expressed through the different features and angles of his face, there are no indications as to what the man is feeling. With no explicit mentions of an emotion, all perceptions of the man's feelings are mere assumptions, maintaining the ambiguity of his true emotions. This is supported by a comprehensive analysis spanning studies of varying individuals that determined that while certain facial configurations may occasionally align with specific emotions, the variability across contexts, individuals, and cultures undermines their reliability as diagnostic displays of emotional states. This highlights that facial movements serve a broader communicative role influenced by situational factors, cultural norms, and individual differences, emphasizing that while faces are indeed powerful tools for social interaction and convey rich

information, including emotional cues, the simplistic link between specific facial configurations and discrete emotional states does not hold under rigorous scientific scrutiny.¹⁹ This critical analysis underscores that facial expressions are highly context-dependent and subject to cultural norms and individual differences. Rather than serving as universal “emotional fingerprints”, facial movements play a broader communicative role influenced by situational factors. The variability observed challenges the simplistic assumption that specific facial configurations directly correspond to discrete emotional states across all contexts and individuals. It is recognized that there is a dynamic nature of facial expressions, which aids in avoiding the reinforcement of oversimplified models or stereotypes.

The interplay of EK and visual perception

Prior knowledge significantly influences how individuals perceive visual complexity in different environments. Exposure and familiarity with specific visual textures can alter perceived complexity.²⁰ This interaction underscores the nuanced relationship between prior knowledge and visual perception, where familiarity can either amplify or diminish the perceived complexity of visual stimuli.

Prior knowledge is known to influence the perceived complexity of visual scenes. For example when participants viewed texture images with or without prior knowledge, an interaction effect between prior knowledge and texture type was observed, emphasizing a nuanced relationship where the influence of prior knowledge on complexity perception varied depending on the specific texture being observed.²⁰ Outside of the controlled laboratory environment, prior knowledge of visual information comes in many forms. Individuals bring various forms of prior knowledge and experience that can significantly shape how they interpret and perceive emotions conveyed through facial expressions. For instance, cultural backgrounds, personal experiences, and learned social cues all contribute to an individual's EK.²¹ Just as verbal cues altered perceived complexity in the texture images experiment, EK can amplify or attenuate the perceived emotional intensity or sincerity of facial expressions.

The connection between EK and socialization

Beyond controlled laboratory settings, individuals bring a diverse array of prior knowledge to everyday visual encounters. This knowledge encompasses cultural norms, personal experiences, and learned social cues, collectively shaping interpretations of emotional content conveyed through facial expressions.

Cultural backgrounds profoundly influence EK, which in turn modulates how individuals interpret facial expressions.¹⁵ Cultural norms dictate the appropriateness and interpretation of emotional displays, guiding individuals in understanding the emotional states of others within their cultural context. This variability underscores the dynamic nature of emotional perception, where cultural influences shape the EK necessary for accurate interpretation of facial expressions. Cultural backgrounds, personal experiences, and social cues collectively shape an individual's EK and decision-making processes. For instance, when participants engaged in a

task where they chose between safe and harmful options, participants' decisions were influenced significantly by partners who gave predictive gaze cues, particularly when coupled with fearful facial expressions. This highlights that humans can effectively use social cues, such as gaze direction and emotional expressions, to enhance their learning and avoid potential harm, which are crucial abilities for adapting to changing social dynamics and environments where partners may vary in reliability.²¹ This integration of social cues into decision-making processes highlights how individuals leverage shared social information to navigate complex environments. By understanding these dynamics, researchers gain insights into how cultural backgrounds, personal experiences, and social cues collectively contribute to the development of EK and influence decision-making across diverse contexts.

Facial perception is altered by EK

Social cues, such as gaze direction and facial expressions, play a crucial role in decision-making and emotional perception.²² This adaptive ability highlights the integration of social information into decision-making processes, where individuals leverage shared social cues to navigate complex social dynamics effectively.

EK plays a crucial role in modulating how individuals interpret and understand facial expressions. In a study in which individuals were given varying levels of contextual information before detecting words and faces, participants showed higher accuracy in identifying the emotional content of written word. In contrast, the study found no corresponding improvement in emotion detection accuracy for facial expressions when contextual auditory support was provided, suggesting that while auditory context can enhance understanding of emotional meaning in linguistic contexts, its impact on facial expression perception may be more nuanced or less direct.²³ This finding aligns with the theory of embodied cognition, which posits that understanding emotional meanings involves simulating the corresponding emotional states through sensorimotor processes.²⁴ This emphasizes that while verbal cues can enhance emotion detection in linguistic contexts, their impact on facial expression perception may be nuanced or indirect, underscoring the complexity of emotional communication, where individuals integrate multiple sources of information to interpret emotional signals accurately.

However, it was found that facial mimicry was an indicator of emotional resonance.²³ Both instances underscore the complex interplay between different modalities of emotional communication in social interactions, where verbal and non-verbal cues interact in varying ways to shape emotional understanding. While there is no correlation between the direct enhancement of emotion detection accuracy for facial expressions with auditory contextual support, facial mimicry, a process linked to emotional resonance, responded to facial expressions. This suggests that EK, acquired through experiences and cultural backgrounds, influences how individuals interpret and resonate with facial expressions by facilitating the simulation of corresponding emotional states.

Conclusion

The results of this investigation of the interplay between vision and memory reveal a complex and synergistic relationship that profoundly affects how we perceive and interact with our environment. Here, we demonstrate that vision and memory are not isolated functions but work together to enhance cognitive processes. Vision provides the sensory detail necessary for object recognition, scene interpretation, and action guidance, while memory integrates these visual inputs with prior experiences and emotional contexts. This dynamic integration is crucial for tasks such as identifying familiar faces and navigating complex environments. This highlights the significant role of prior knowledge in shaping visual perception and emotional interpretation. Cultural backgrounds and personal experiences play a critical role in how we interpret visual complexity and emotional cues, influencing everything from facial expression recognition to aesthetic appreciation. EK is particularly important in modulating these interpretations, guiding responses to emotional cues and enhancing social interactions.

Furthermore, it was found that aesthetic perception involves evaluating qualitative aspects of stimuli beyond mere recognition, intertwining deeply with cognitive and emotional processes. Notably, while facial expressions are often viewed as universal indicators of emotions, this research shows considerable variability influenced by cultural norms and individual differences. This variability challenges simplistic models of emotional communication and underscores the need to account for broader contextual factors in interpreting facial expressions. Additionally, the study found that although verbal cues can improve emotion detection in linguistic contexts, their effect on interpreting facial expressions is more nuanced.²³ This suggests that a combination of sensory inputs, contextual knowledge, and cultural factors interact to shape our understanding of emotions. Overall, these insights underscore the intricate, interconnected nature of vision, memory, and emotional processing, highlighting the necessity of integrating sensory input with prior knowledge and cultural context to fully appreciate the richness of human perception and cognition.

Recent results have demonstrated the complex nature of human cognition. They reveal that perception is an active, interpretive process influenced by prior experiences, cultural contexts, and emotional states. This dynamic interaction challenges the notion of universal emotional expressions, highlighting the need for more nuanced models of emotional recognition and interpersonal communication. The variability observed in facial expression interpretation based on cultural and individual differences underscores the importance of considering contextual factors when studying emotional communication. However, several areas remain unexplored. The exact mechanisms by which prior knowledge and emotional states shape sensory input and perception are not fully understood. Further research using advanced neuroimaging techniques could elucidate the neural networks involved in these processes. Additionally, expanding cross-cultural studies and longitudinal research could enhance our understanding of how cultural norms and personal experiences influence emotional recognition and aesthetic perception. Interdisciplinary approaches combining cognitive psychology, neuroscience, and cultural studies will be crucial for addressing these gaps and advancing our knowledge of human cognition and its applications.

Works Cited

- How the Eyes Work. *Natl. Eye Inst.* (2022).
- Dragoi, V. Chapter 14: Visual Processing: Eye and Retina. *Neurosci. Online* (2020).
- Shapley, R. & Rubin, N. Vision: Brain Mechanisms. *ScienceDirect* (2002).
- Anatomy & Physiology*. (Pressbooks, Colorado Community College System).
- Visual Perception. *Interaction Design Foundation*
<https://www.interaction-design.org/literature/topics/visual-perception>.
- Vijayabaskaran, S. & Cheng, S. Navigation task and action space drive the emergence of egocentric and allocentric spatial representations. *PLoS Comput. Biol.* **18**, e1010320 (2022).
- Preston, A. R. & Eichenbaum, H. Interplay of hippocampus and prefrontal cortex in memory. *Curr. Biol. CB* **23**, R764-773 (2013).
- Frankland, P. W., Josselyn, S. A. & Köhler, S. The neurobiological foundation of memory retrieval. *Nat. Neurosci.* **22**, 1576–1585 (2019).
- Cherry, K. What Is Perception? *verywellmind*
<https://www.verywellmind.com/perception-and-the-perceptual-process-2795839> (2024).
- Colman, A. M. *A Dictionary of Psychology*. (Oxford University Press, Oxford, 2015).
- Torabi, N. Emotional Intelligence: Navigating the Power of Emotions in Human Nature and Actions. *Medium*
<https://neemz.medium.com/emotional-intelligence-navigating-the-power-of-emotions-in-human-nature-and-actions-c68049b23c2f> (2023).
- CNX OpenStax. *Biology*. (2016).
- Karim, A. K. M. R., Proulx, M. J., De Sousa, A. A. & Likova, L. T. Do we enjoy what we sense and perceive? A dissociation between aesthetic appreciation and basic perception of environmental objects or events. *Cogn. Affect. Behav. Neurosci.* **22**, 904–951 (2022).
- Phillips, M. *Faces Aesthetics*. (2024).
- Engelmann, J. & Pogosyan, M. Emotion perception across cultures: the role of cognitive mechanisms. *Frontiers in Psychology*. **4**, 118 (2023). doi:10.3389/fpsyg.2013.00118.
- Lange, E., Funderich, J. & Grimm, H. Multisensory integration of musical emotion perception in singing. *Psychological Research*. **86**, 2099-2114 (2022).
doi:10.1007/s00426-021-01637-9.
- Perception of Facial Expressions Differs Across Cultures. *Am. Psychol. Assoc.* (2011).
- Facial Expressions - The Cartoonist's Art*. (1912).
- Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. & Pollak, S. D. Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. *Psychol Sci Public Interest* **20**, 1–68 (2019).
- Li, L. & Choi, W. Does prior knowledge increase or decrease perceived visual complexity of texture images? *Heliyon*. **9**, e15559 (2023). doi:10.1016/j.heliyon.2023.e15559.
- Pärnamets, P. & Olsson, A. Integration of social cues and individual experiences during instrumental avoidance learning. *PLoS computational biology*. **16**, e1008163 (2020).

doi:10.1371/journal.pcbi.1008163.

Frith, C. Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*. **364**, 3453-3458 (2009).

doi:10.1098/rstb.2009.0142.

Blom, S. S. A. H., Aarts, H. & Semin, G. R. Perceiving emotions in visual stimuli: social verbal context facilitates emotion detection of words but not of faces. *Experimental Brain Research*. **239**, 413-423 (2021). doi:10.1007/s00221-020-05975-9.

Shapiro, L. & Spaulding, S. Embodied Cognition.

The Inequalities of Green Policy By Armaan Sarkhel & Ishan Parikh

Abstract

Environmental concerns are becoming a hugely prevalent issue in today's society. Businesses and people are increasingly promoting behaviors that are supposed to aid the battle for protecting our environment. Large companies have a tremendous impact on the environment through their production processes, distribution methods, and most other aspects of their business. As companies aim to meet the demands of an ever-growing consumer base they are ramping up production. The resulting increased extraction of materials leads to increased carbon output, deforestation, and the general destruction of many habitats while releasing harmful chemicals into our atmosphere. As more and more people are being made aware of the harmful effects of environmental-related issues, companies have been establishing positive stances on the matter to secure consumer trust and align with regulations. These positive stances are convoluted with clever marketing strategies and corporate jargon that favors the big business. A company's number one priority is to make an increased profit, and thus there is an inherent bias towards company benefits in their "positive" stances. For example, a popular phrase being thrown around in the corporate community is "carbon neutrality". This is when a company's carbon output is offset by its carbon absorption. Essentially, if a company cuts down ten trees, it will plant another ten to restore its neutrality. While this sounds nice on paper, there is increased complexity within this matter that favors a company's profitability. This review examines the complex relationship between carbon neutrality of big businesses, and the inequalities that lay hidden behind the surface.

Introduction

Companies across the globe continue to contribute to sustainable initiatives to cut down on greenhouse gas emissions. Since 2016, Apple has assigned a total of \$4.7 billion in Green Bonds used to fund projects that have positive environmental benefits (Apple).

This paper introduces a critical examination of the economic, social, and environmental inequalities associated with sustainable initiatives. Many companies partake in carbon-offsetting practices, the act of investing in projects that absorb the carbon that companies have produced, such as planting trees. However, due to the de-regulation of this process, carbon offsetting tends to benefit the wealthy for business profit while undermining the rights of underrepresented populations. To understand this duality between a company's promise of sustainable development and its inherent inequalities, this paper examines 4 main attributes of big business: Consumer, Supply, Sustainable Development, and Employer. This paper will mainly focus on big business since as time has gone on big business has become a more prevalent part of our lives than ever before. For example, services such as Amazon Prime have become more central within our lives due to our increasing dependence on packages arriving faster than they ever have before. More and more people's needs are being served by big businesses, leading to their greater influence throughout society.

Climate change is not an issue tied to the boundaries of national borders, regional jurisdiction, or private enterprises. Climate change is a global issue that is spreading throughout the air and natural environments irrespective of artificial boundaries. For this reason, the nations and companies of the world must work together to effectively combat climate change. However, each nation has different interests and priorities, as well as different financial and economic situations. The fight against climate change is bound to harm one country's ambitions while benefiting another's. As a result, sustainable goals have been hotly contested and have been slow to enact.

Climate change also continues to be a major problem due to the lack of understanding of the true size and scope of the problem within the general population. Since not nearly enough people know the severity of the issue, there is not enough pressure being placed on world leaders to take action regarding the crisis and there is a lack of realization among the general public on the extent of the effects of climate change.

A large contributor to the execution of certain goals and policies is profitability and clear advancements. However, climate change is a deeply rooted issue that will take much more than a single company or nation to change. For this reason, it is hard to find instant gratification through sustainable advancements and hence people tend not to favor them. In essence, we are turning a blind eye to climate change because it is not a tangible statistic or object we can change in one go.

Employer

The first aspect this research paper will examine is the employer. When discussing inequality, there is generally a group of people who have the chance to better the world, but choose to fill their own pockets instead. There are many practices carried out by employers that directly and indirectly harm groups of people while benefiting the elite. These include corruption, avoidance of taxes, implementation of cheap labor, diversity, and an overemphasis on investors. All these problems arise from the end goal of increasing profit for the rich. This section will generally follow paragraph eight from the book *Net Positive: "Embrace the Elephants"*, while emphasizing the effects of inequality. This section will outline some of the inequalities that will be expanded upon further throughout the rest of the paper.

Taxes play an instrumental role within certain societies in facilitating the improvement and well-being of communities. Many large companies do all they can in order to avoid paying a large share of their taxes. This not only places a large burden on everyday people who have to make up for those taxes, but it also takes away from the company's own resources. For example, Amazon (Paid \$3.4 billion in taxes on \$960 billion in revenue and \$26 billion in profits) Some years amazon paid 0 (Polman and Winston 194-195). The money that Amazon is keeping for itself could have been used for bettering community roads and signage. The same roads and signage that Amazon relies on to deliver packages and carry out its business.

It is a company's responsibility to look out for the well-being of its employees as it pertains to the workplace. Labor standards must aid the employee rather than the employer in

order to achieve true equality. If a company is producing sustainable products, but using cheap labor in unsafe working conditions, those products being made are not benefiting everybody; there is inequality present.

When a company is purely after profit, it tends to only pay attention to the investors and their requirements while turning a blind eye to the others who are affected by the company such as consumers or laborers. When the sole commitment is to shareholders and their profit, companies are planning for short-term profit by raising their stock price rather than long-term growth. This, for example, can lead to the exploitation of laborers for higher profit margins in order to attract more investors. Conforming to the wants of the wealthy is ignoring the needs of the less fortunate.

One of the biggest challenges that businesses face is how to maximize their profit to be as large as possible while still maintaining ethical standards that do not disadvantage any particular people, groups, or areas. In fact, the issue is so prevalent in large corporations that corruption, bribery, and tax evasion all together take away \$1.26 trillion dollars from developing countries each year. This amount would be enough to help the 1.4 billion poorest people in the world and bring them above the poverty line, however, it is unfortunately wasted in these malicious practices amongst those in power at these large corporations (Polman and Winston 197). In addition to the corruption going on amongst higher-ups at these companies, there is also the simple fact that many of the executives in these big businesses receive much more money compared to the average employees at the company. In the United States for example, in 2019, the average CEO made 320 times more than the average employee. This gap in pay is following an exponential trend as this is higher from 61 times more in 1989 and 21 times more in 1965 (Polman and Winston 199-200). Lastly, trade association lobbying puts big corporations in difficult and peculiar situations. The biggest issue regarding lobbying is when the promises and goals a company makes are lobbied against by fellow members of the trade associations they are a part of. This gap between company promises and trade association lobbying is most commonly seen in the fossil fuel industry. Many companies in the industry often make statements solidifying their support for a price on carbon, however, when a policy actually arises to make the statement a reality, representatives from these companies are quick to mobilize against it. For example, one such policy appeared on the state voting ballot in Washington state in 2018, but the Western States Petroleum Association decided to invest \$30 million towards strategies to convince people that the policy would be bad for them. (Interestingly, \$13 million alone was from petrol giant BP) (Polman and Winston 208-209). The lack of transparency and the commitment discrepancies between company statements and the actions taken during lobbying shows how once again profit is the number one interest for these big corporations even at the cost of employer trust and making false promises.

Diversity can help represent more viewpoints within a company. The needs of minorities and oppressed groups can be represented in a boardroom filled with more people of different backgrounds. This could lead to less exploitation of minority groups and greater equality through company policies.

Sustainable Development

The second attribute we will be talking about is sustainable development. As this research paper has established, it is vital that all people within society work together to combat climate change. However, a lot of money and systemic change is required in order for all corners of society to become sustainable. Within the present attempts to undertake this huge task and the monetary and systemic challenges involved, it has become apparent that some level of inequality is present.

One roadblock in transitioning to a completely sustainable future is different governments and societal systems. For the United States of America, sustainable development is met with the prospect of profitability. In contrast, China may, hypothetically, have different ambitions and values regarding the implementation of sustainability. In this case, the ways in which these two nations undergo the implementation of eco-friendly projects will be inherently different and thus there is no one solution that fits all. However, this brings about the idea that some nations are not interested in sustainable development and as a result, they are being harmed by certain policies and projects.

One potential solution to increased oil and gas usage is increasing the cost of these substances in hopes that people use less oil and gas due to their higher prices. This is known as conservation: “the reduction of the amount of energy consumed” (Laxer). On paper, this seems like a great idea, however, this is another example of a policy that would harm the lower class while leaving the upper class unscathed. “While the affluent may continue to drive SUVs in times of high prices, poor rural residents often have no access to public transit and are forced to keep driving and to make other painful lifestyle choices such as skimping on high quality food or turning down the thermostats in their homes during the winter” (Laxer). People in the middle to upper class have enough disposable income to be able to afford the higher gas prices called for, while people who are not as financially secure will suffer, as a major necessity in their lives will be almost unattainable. In Europe, there are increased gas taxes which see fewer people on the roads, however, this is a perfect example of different societies needing different solutions (Laxer). Many European countries have the infrastructure in place to facilitate the movement of people without the need for cars while many parts of the US are wholly dependent on cars. While a fairly simple proposition such as increased gas prices and taxes seems feasible at first glance, it is clear that this overlooks differences between societies while harming the lower class. For sustainable development to be viable, projects and policies must not be one size fits all and they must not serve as a mere inconvenience for the more wealthy, one which they can easily steer around.

Unequal policies such as this one are apparent all throughout the world of sustainable development. For example, many developed countries are calling for increased sustainable policy, disregarding the economies of developing countries that rely on certain emissions. “For over a century, the world has run on oil. Its availability and relatively low price allowed for the industrial growth and development of the world’s leading economies. Today, the new rapidly expanding giants, India and China, want access to the same possibilities...” (Laxer). Elites

within developing countries are also being pressured into supporting green policies that favor the West, indirectly harming their nation's development. When examining the relationship between developed and developing nations, we see an abundance of hypocrisy. For one, developed nations are trying to prevent developing nations from accessing as much oil, the same oil that made developed nations develop. Secondly, developed nations outsource their factory jobs to developing nations for cheaper labor. If developed nations limit these country's access to oil then developing nations will struggle to produce, hurting not only the respective country's economy but also the economies of developed nations (Laxer).

Another commonly thought of approach in regards to creating a more sustainable society has been to try to influence people's choices when it comes to their food habits; specifically for meat eaters to transition to more of a plant-based diet. The reasoning behind this is that animals such as cows emit a large amount of greenhouse gasses, such as methane, as a result of burping and passing gas. In fact, some estimates approximate that livestock raised for food makes up around 15 percent of the total amount of greenhouse gasses that are produced every year (Nardo). Additionally, the feed that is produced for these livestock results in the creation of carbon dioxide emissions along with the animal manure which emits nitrous oxide. While switching diets seems like an easy-to-accomplish first step, it does not account for the various different cultures and diets around the world for whom meat is regarded as an essential item. For example, in China, pork has been consumed for thousands of years and it serves as a staple ingredient in many dishes such as Char Siu and Baozis. Meanwhile, in countries in South Asia, meat is implemented in various traditional dishes such as kebabs and curries.

As a response to this suggestion, many grocery brands such as Stop & Shop Nature's Promise and fast food chains such as McDonalds and Burger King are coming out with plant-based meats and burgers which are meant to incentivize customers to help support the Earth. However, a lot of the time these items are not available in all countries, and even if they are, the products seem to be priced at a higher price than the regular meat option. For instance, Burger King's Whopper is averagely priced at around \$5.39 as opposed to their Impossible Whopper which is plant-based which is priced at \$6.39 (Grubhub). At the end of the day, this supposedly sustainable development goal by large businesses in the food industry turns out to be another vessel for which these companies can build profit while harming some cultures. Once again, this shows the inequalities associated with current techniques of shifting society to a more sustainable future.

This research paper is not implying that the imposition of green policy is bad and should not be done, rather it is simply pointing out the inequalities in today's green policy and how they favor certain groups of people.

Another example of policy favoring the Western world is in California. California is implementing ambitious goals for vehicle sustainability such as through the state's Advanced Clean Cars II regulations which state that all consumer cars (cars, trucks, & SUVs) will have to be zero-emission vehicles by 2035 (California Air Resources Board). California however, is one of the costliest states to live in within the United States and has the 5th highest per capita income

in the country (Data Pandas). Therefore, the state has enough income to implement EV infrastructure such as charging stations. Companies will be pressured to develop and manufacture more EVs due to California’s policies and the rest of the world will feel the effects of this shift. A region like rural India may not have the funds to set up charging stations for electric cars and large manufacturers would not be able to cater to this market as it is too small. As a result, the cost and efficiency of owning a car would deteriorate.

Through the different case examples discussed throughout this section, this research paper has established the downsides of current and proposed developments. As a result, the obvious question arises: how do we shift to sustainable societies in a more equal manner?



Fig 1. Courtesy of DoughnutEconomics.org

A sustainable development model that can be helpful in achieving this is the Doughnut Model, created by Kate Raworth in her book, *Doughnut Economics: Seven Ways to Think Like a 21st-Century Economist*. The model essentially serves as a guide or “compass” to attaining the basic needs of all humans in the 21st century while also ensuring the providing capacity of the planet is not breached. As the name suggests, the model is in the shape of a Doughnut with two rings: the inner one called “Social Foundation” and the outer one called the “Ecological Ceiling.” The inner ring consists of basic necessities for making up a prosperous society such as water, food, health, and housing. On the other hand, the outer ring is surrounded by nine ecological ceilings ranging from Climate Change to Ocean Acidification, which are some of the planetary boundaries that society has to stay within. The only way to attain a sustainable and just society is by staying within “the safe and just space for humanity”, not exceeding our resources while maintaining basic needs for all (Doughnut Economics Action Lab).

This has proved difficult for nations to do since providing for better social needs increases carbon output while remaining within our ecological boundaries decreases the access to resources for many people (Saito). However, if the profits of big business are set aside, nations

can remain very close to being within these boundaries, as the prospect of profit and increased economic growth of developed nations shift these priorities (Saito). The doughnut model displays the importance of maintaining a just society in the quest for sustainability while also showing that both sustainability and justice are possible to achieve together.

Supply

An important aspect of green development is on the supply side. A large component of a company's carbon output is the manufacturing of a product. For this reason, many of a company's environmental goals and projects relate to their production systems and how they account for their emissions.

In order to fully explain the inequalities associated with a company's green goals on the supply side, it is important to understand the idea of carbon offsetting. To put it simply, carbon offsetting is when a company or individual attempts to compensate for their carbon production by investing/trading in commercial programs that aim to balance the number of metric tons of carbon produced by large corporations. These programs and organizations achieve this by taking part in initiatives such as reforestation and methane capture projects which sequester carbon and other greenhouse gasses. The way big businesses manage to partake in Carbon Offsetting is by purchasing many Carbon Offsets from a market known as the Voluntary Carbon Market, with 1 Carbon Offset accounting for 1 metric ton of CO₂. In fact, $\frac{2}{3}$ of the world's biggest corporations actually purchase Carbon Offsets to meet their ever-increasing pledges towards sustainable development and climate goals (Gabbatiss). These large corporations are so dependent on the practice of Carbon Offsetting that it was found that only 34 companies were found to have purchased enough Offsets to offset 38 Million tons of Carbon Dioxide produced during the years 2020-2022, which is equivalent to the annual emissions of Ethiopia and Kenya combined (Gabbatiss).

At first glance, this seems like a sensible idea. Major businesses need to meet the demands of the ever-growing market and so to do so large carbon outputs by these corporations are bound to happen. With these companies investing in projects that aim to nullify the damage done, it appears as if accountability is being held and the planet is being sustained responsibly and fairly. However, this initial viewpoint can be regarded as surface-level naivete, as these carbon offsetting projects introduce new levels of inequality not only in the environment but also in societies around the world.

In a bid for cheaper labor, many companies are outsourcing their production to developing countries like India, China, and Bangladesh. For example, China produces 27% of the global manufacturing output and is 5% cheaper on average compared to production in the US (West and Lansang). More recently, companies have been shifting their production to India, with Apple producing 14% of their global iPhones in India (aiming for 25%), where the average cost of labor is \$0.92 per hour which is even cheaper than China's \$3.52 per hour (Zhu). As a result, this means that these nations' greenhouse gas emissions are rising more and more. The inequalities of this are apparent, with companies "cutting pollution in one location, but leaving

another factory belching asthma-causing particulates near low-income communities” (Polman and Winston). Fixing this in the long run is not exactly simple. For example, companies such as Unilever have many factories in developing countries that are contributing to the development of their respective economies, but a lot of the time these local factories do not have access to the cleanest methods of technology. Thus, many of these factories have to use fuels such as coal which does push back on their renewable goals but ultimately contributes to the overall profit and success of the corporation.

One particular effect of a company’s flawed production methods is deforestation. “The deforestation taking place in Brazil and in other parts of the world, including the Congo basin in central Africa and in the jungles of Southeast Asia, represents a major threat to the planet” (Nardo). Companies rely on cutting down trees for the extraction of wood while also clearing land to build factories and office spaces. While carbon offsetting may seem as though it counteracts these problems with the planting of new trees, it is clear that this is very ineffective. First of all, it is important to address that the whole carbon offsetting process/system is very unregulated. There is no inherent governing body that manages the efficiency or legitimacy of these carbon-offsetting projects and companies are not required to invest in projects that replant in the same country that the initial trees have been cut down in. Additionally, the goal of reforestation is one that has many difficulties in the process. For one, there is a large financial burden associated with reforestation. Governments must dedicate large sums of money to a goal that does not have immediate benefits. Additionally, planting new trees and protecting forests detracts from land that could be used for farming or other needs that are centric to a country’s economy (Nardo). For the upper class, this may not seem like a beneficial compromise to a goal that won’t show instant gratification. It is not just factories and product manufacturing that directly harm these low-income communities but also whole new industries as well that are aimed towards pleasing the wealthy. One great example of this is the whole concept of Space Tourism and Travel done by companies such as Jeff Bezos’ Blue Origin. Blue Origin’s New Shepard Spacecraft aims to bring “space tourists” about 100 kilometers into the atmosphere and then return in 10 minutes. Doing so however ends up releasing a footprint of 330 Carbon Tons for the pilot and the three passengers onboard which equates to an average of an annual 5-ton living lifestyle per minute (Berners-Lee). As space tourism grows and companies such as Virgin and SpaceX join in the mix, carbon outputs are going to continue to increase exponentially as these flights get more frequent and ultimately it will be the lower class who will feel the effects of this increased footprint first through environmental disasters and living conditions.

Many carbon offsetting initiatives, even if they are conducted in exploited areas that would, in theory, benefit from the plantation of trees and other sources of carbon absorption, still undermine the rights and wants of groups of people. “Investigations into individual carbon-offset projects by journalists and non-governmental organisations have revealed that many of these schemes can come with devastating impacts for Indigenous peoples and local communities” (Dunne and Quiroz). There are generally four ways in which this happens: Indigenous people

being affected, an overestimation of offsetting goals, illegal land use, and an effect on food production.

CarbonBrief has mapped many instances of harmful carbon offsetting, however, this paper will focus on two cases which together cover all four harmful effects. In Liberia, the government made a deal with a UAE-based firm, signing away almost 10% of its total land mass. This deal could violate many laws aimed at protecting communities and the land rights of the local people in Liberia. This comes as the UAE is ramping up preparations for the COP28 Climate Conference on climate action and rule-making. This is a clear case in which a nation or company aims to use climate policy as a marketing tool. When we uncover the details of this deal with Liberia, it is clear that Indigenous people will be affected through illegal land use, covering two of the four harmful effects of supposedly beneficial carbon offsetting. The UAE-based firm aims to generate protected forests for carbon credits, affecting the local community's rights to govern how their land is used.

Liberia's deforestation rate is the 10th highest rate of increase of any country between 2020 and 2022 (Mukpo). Deals like this one are taking advantage of communities falling into poverty rather than forming well-built plans for tackling deforestation. As this research paper has examined, carbon credits are not an effective strategy to combat deforestation. "Kate Dooley, a research fellow at the University of Melbourne's Climate and Energy College, told Mongabay she worries credits sold by [the UAE-based firm] directly to governments would have weak monitoring and verification requirements" (Mukpo). Additionally, the UAE-based firm behind these plans, Blue Carbon, serves as a middleman in transactions between companies seeking carbon credits and the Liberian government. This means that Blue Carbon is able to claim a profit through these transactions. Once again, the wealthy are using climate goals as a marketing strategy to generate more profit while exploiting less fortunate people in the process.

Another example of a carbon offset project that had done more harm than good would be the Northern Kenya Grassland Carbon Project. The massive project covers around 2 million hectares of land and was started by the Northern Rangelands Trust (NRT) and was labeled as the world's largest soil carbon removal project to date when it was first launched in 2012. At the start of the project, NRT claimed that by using "planned rotational grazing methods" they would be able to foster more vegetation more efficiently compared to the indigenous Samburu, Maasai, Borana, and Rendille Tribes of the area leading to a higher carbon collection. NRT's initial projections estimated that 3 quarters of a ton more carbon per hectare per year could be achieved leading to an overall increase of 41 million net tons worth of available carbon credits leading to a \$300 to \$500 million profit (Rosane).

To put it simply, the aim of the Northern Kenya Grassland Carbon Project was to increase vegetation in the area, however, according to an analysis done on the project, NRT's own maps from 2012 to 2020 show that vegetation in 48.5 percent of the project area has been declining. And since vegetation cover is correlated with soil carbon sequestration this blatantly shows how the soil carbon pool is reducing and that the aims of the project are going in the backward direction (Rosane).

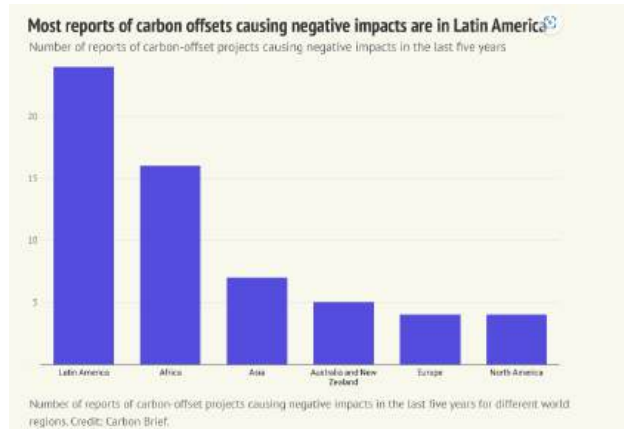


Fig 2. Courtesy of Carbon Brief

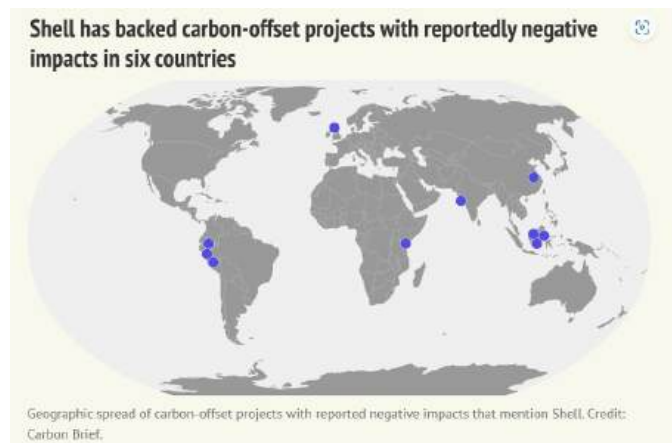


Fig 3. Courtesy of Carbon Brief

Unfortunately cases such as this have occurred many times before and conservationists state that results such as these are caused by predating colonial and racist attitudes which have long dominated the conservation industry. In fact, many national parks and reserves have been established in locations without the consent of the indigenous people and local communities in the area which has often resulted in eviction and the separation of these groups from their livelihoods. Turning back to NRT specifically, the organization has been involved in many different conservation efforts as NRT manages 43 community conservancies all of which cover a total of about 63,000 square kilometers (Rosane). In regards to their local community treatment over time, an Oakland Institute report from 2021 observed that NRT’s project had “allegedly dispossessed pastoralist communities of their ancestral lands, through corruption, cooptation, and sometimes through intimidation and violence, to create wildlife conservancies” (Rosane). Not only this but NRT conservations have been found to be patrolled by rangers that the Trust

specifically hires from neighboring tribes that have conflict with the tribe displaced by the organization itself. As NRT’s carbon credits prices have increased over time, they use the funds obtained by these increased prices to pay for more rangers and thus continue tribal violence. To top it all off, NRT did not indeed gain consent from the local tribes before starting this commotion.

We not only see inequality through economic status and representation, but we also see inequalities regarding labor standards within these green initiatives. As this paper has established, many green initiatives primarily benefit the wealthy and the people of Western nations. This is all possible due to the means of production used in developing countries. For example, electric cars, which are expensive and require infrastructure mainly available in Western countries, also require the mining of cobalt. The largest exporter of cobalt is the Democratic Republic of Congo, where “freelance workers who do extremely dangerous labor for the equivalent of just a few dollars a day” are used (Gross).

Taking a closer look at Tesla, a large electric car manufacturer, we can more clearly see this duality between climate-friendly production and the systems behind that production. In 2019, Tesla was 5th on Influence Map’s report on climate policy. Tesla was increasing their support for the Paris Aligned Climate Policy, aimed at increasing worldwide efforts in combating climate change. Despite this, Tesla was one of the lowest-ranked automobile manufacturers on the 2020 Corporate Human Rights Benchmark. This displays how even though Tesla claims to be eco-friendly, their production methods and systems still exploit cheap labor with “Forced labour, lack of living wages and restrictions to freedom of association” being many of the issues contributing to Tesla’s low score (World Benchmarking Alliance).



Fig 4. Courtesy of Corporate Human Rights Benchmark; World Benchmarking Alliance

This is a common theme across the corporate world, with green policy and production being used as a buzzword to drive up profitability. Companies are using a policy that harms

many people, however, they are effectively turning a blind eye toward the issue whether that be labor standards, the harming of Indigenous people, or the false positivity of carbon offsetting.

Consumer

The last attribute of big business this research paper examines is the consumer side. This includes marketing strategies aimed at consumers as well as sustainable products that are intently made less available to certain groups of people. It is first important to establish the proper definition of a commonly used phrase throughout the corporate world: carbon footprint. When companies say they want to minimize their carbon footprint, what does this actually refer to? Consumers are made to believe that a reduction in one's carbon footprint means a reduction in carbon offset and increased usage of sustainable products. However, the phrase "carbon footprint" is one that is majorly abused (Berners-Lee 6). A carbon footprint may not be as trustworthy of a statistic as one thinks due to its high individuality. One can minimize their carbon footprint through carbon offsetting means (refer to supply) even while emitting huge amounts of greenhouse gasses. It is not difficult for groups of people to shape the definition of carbon footprint to fit their own needs. "In fact, the term was popularized by oil giant British Petroleum (BP) in the early aughts as part of a marketing campaign crafted to redirect attention for solving climate change away from fossil fuel corporations; in 2004, the company worked with a public relations firm to launch a carbon footprint calculator to help the individual assess their own role in global warming (Carleton and Gordon).

When examining these various marketing techniques, we can see a clear pattern of disregarding the wealthy and placing all the emphasis on middle to lower-class individuals as well as individuals in eastern regions. For example, the average US person's annual carbon footprint is about 21 tons, however, due to the fact that the U.S. is an extremely unequal society this number does not represent the majority of the country's population. This is as the extremely wealthy people of the nation drive this number up and inflate it beyond the range of the average middle-class American with a high carbon lifestyle. Aside from domestic differences in carbon output, when looking at the world as a whole it can be clearly seen how inequality runs rampant. Specifically, it would only take the average US person a couple of days to achieve the same footprint as the average Nigerian or Malian's annual footprint. (The global average is about 7 tons). In fact, China, a country that is notorious to the West for environmental concerns, has a much lower annual carbon footprint per person at 8 tons of CO₂ per year (Berners-Lee).

These marketing strategies and their preference for the wealthy are even more apparent when analyzing simple everyday habits that are supposed to be sustainable. Electric cars, reusable water bottles, and reusable bags; how are these shifts to eco-friendly products helping the environment? Well, that's simple, an electric car is releasing less greenhouse gasses while reusable water bottles and bags are helping lessen waste. However, "these good deeds are meaningless in the end. They can even cause more harm than good" (Saito). These various countermeasures against climate change merely serve as an alibi. They take our attention away from the true atrocities. How much of a difference is a reusable bag really making? In fact,

reusable bags just serve as “Catholic indulgences,” “[preventing] us from taking part in the larger actions truly necessary to combat climate change” (Saito). They are an outlet for the upper and middle class to take their minds off of climate change. This is largely because they cannot see (or choose not to see) the true impact of climate change largely felt in lower-income areas due to a lack of available sustainable systems (mainly due to cost) and the exploitation of those areas and populations (see Supply for more). When one buys an electric car, they feel as though they are largely reducing their carbon emissions. However, this does not take into consideration the emissions used to produce the cars, ship the cars, and charge the cars. Additionally, electric cars are expensive, and generally available to those who have the money to afford them. These cars make the wealthy who buy them feel good by blocking their view of the far-reaching dangers of climate change. This brings up another important point: sustainable products are not available to everyone. Specifically, sustainable products are priced as luxury items when they should be standard for all. In fact, “the prices of sustainable products are higher by 75-80% on average (Misona). Sustainability should not be bought in a way that only the wealthy can benefit from. Sustainability should be an inherent part of our society so that everyone benefits. However, capitalism proves to have inherent hurdles associated with it. Companies are obsessed with profits even if those profits come at the cost of the lower class and sustainability. When a company markets a product as being sustainable, they can charge higher for those products and people will still buy them. This cuts off the lower class from having availability to these supposedly sustainable products while paying the wealthy more and more.

A case study that this paper has specifically looked into would be concerning Richard Branson and the Virgin Group. The Virgin Group is a prime example of what would be considered a “big business” given that it is a multinational conglomerate company that is involved in the travel industry, telecommunications, media, financial services, etc. Richard Branson, the co-founder of the Virgin Group, wanted to implement a “road to damascus” initiative in order to fight climate change primarily by taking the dividends, share sales, and profits gained by his airlines and trains and then funnel them back towards finding new clean fuels and sources of energy. As a result, Branson launched the Virgin Earth Challenge which would award a \$25 million prize to the first person to come up with an innovative way to remove 1 billion tons of carbon (Klein). Ultimately, the winner of the 11 finalists of the Virgin Earth Challenge was never selected and no team received the promised prize even after 13 years of ongoing assessment of the projects. This was just one of the many projects started by Virgin, whose goals were never met and the reasoning for these failures came back to the simple fact of how building a bigger business took precedence over Virgin’s climate initiative goals time and time again. That’s not to say that necessarily all their plans were failures but at the end of the day it was running more planes and using oil that benefitted Virgin the most by putting more money into the hands of the wealthy and so that is what they prioritized.

A case study such as this highlights the issues that society faces due to a profit-driven economy. The move to a sustainable society requires the overhauling of current infrastructure and production methods which in turn requires a lot of money (Saito). This is money that does

not meet the instant gratification that big businesses are looking for profit. As a result, sustainability is being put on hold while the world's wealthy fill their pockets. Greenhouse gas emissions are soaring and all big business is doing is putting out clever marketing that drives up their products and service's prices (if they even go through with bringing these products and services to fruition) while raking in a huge chunk of money for the global elite.

Conclusion

The general consensus among consumers is that green development is not only beneficial, it is a necessity. However, big business is taking advantage of this mindset by using sustainability as a marketing strategy to earn more profit. When these sustainability claims are examined more closely, there is apparent inequality in who benefits from these companies' policies with some of their policies harming the environment more than they help. By blindly believing large businesses' claims of helping the world, consumers are indirectly supporting the harming of populations and environments while the companies rake in profit. Undoubtedly, a greener world is a necessity, however, the proper steps must be taken to prevent the harming of ecosystems, the harming of populations, and the false claims of big businesses, while transitioning to a greener world. Many of these companies' harmful effects come from their focus on profits. However, profit can remain alongside the benefit of the people. "When we believe in something bigger than ourselves and conduct ourselves accordingly--in all aspects of our lives, meaning in our private lives, in our business activities, and our philanthropy--we reach our highest potential" (Arison). If false and unequal claims of green development are detrimental to certain people and places, they are not helping the environment at all and will instead lead to a more efficient deterioration of our world, harming big business in the process.

Works Cited

- Apple. "Apple's \$4.7B in Green Bonds Support Innovative Green Technology." *Apple Newsroom*, 24 Mar. 2022, www.apple.com/newsroom/2022/03/apples-four-point-seven-billion-in-green-bonds-support-innovative-green-technology/. Accessed 28 July 2024.
- Arison, Shari. *The Doing Good Model*. BenBella Books, Inc., 19 May 2015.
- Berners-Lee, Mike. *HOW BAD ARE BANANAS? : The Carbon Footprint of Everything*. S.L., Greystone Books, 2022.
- California Air Resources Board. "Cars and Light-Trucks Are Going Zero - Frequently Asked Questions | California Air Resources Board." *Ww2.Arb.ca.gov*, ww2.arb.ca.gov/resources/documents/cars-and-light-trucks-are-going-zero-frequently-asked-questions#:~:text=As%20part%20of%20the%20Advanced. Accessed 28 July 2024.
- Carleton, Audrey, and Aaron Gordon. "'Carbon Footprint' and 'Net Zero' Don't Mean What You Think." *VICE*, VICE, 20 Sept. 2021, www.vice.com/en/article/88np3v/what-is-carbon-footprint-net-zero-greenwashing-emissions-guide. Accessed 28 July 2024.
- Data Pandas. "Per Capita Income by State 2024." *Datapandas.org*, 2024, www.datapandas.org/ranking/per-capita-income-by-state. Accessed 28 July 2024.
- Doughnut Economics Action Lab. "About Doughnut Economics | DEAL." *Doughnuteconomics.org*, 2024, doughnuteconomics.org/about-doughnut-economics#what-is-the-doughnut. Accessed 28 July 2024.
- Dunne, Daisy, and Yanine Quiroz. "Mapped: The Impacts of Carbon-Offset Projects around the World." *Carbon Brief*, 26 Sept. 2023, interactive.carbonbrief.org/carbon-offsets-2023/mapped.html. Accessed 28 July 2024.
- Gabbatiss, Josh . "Analysis: How Some of the World's Largest Companies Rely on Carbon Offsets to 'Reach Net-Zero.'" *Carbon Brief*, 28 Sept. 2023, interactive.carbonbrief.org/carbon-offsets-2023/companies.html. Accessed 28 July 2024.
- Gross, Terry. "How 'Modern-Day Slavery' in the Congo Powers the Rechargeable Battery Economy." *NPR*, Feb. 2023, www.npr.org/sections/goatsandsoda/2023/02/01/1152893248/red-cobalt-congo-drc-mining-siddharth-kara#:~:text=In%20his%20new%20book%2C%20Cobalt. Accessed 28 July 2024.
- Grubhub. "Burger King Menu." *Grubhub.com*, 2024, www.grubhub.com/restaurant/burger-king-963-livingston-ave-north-brunswick/1017475. Accessed 28 July 2024.
- Klein, Naomi. *This Changes Everything: Capitalism vs. the Climate*. London, Penguin Books, 2015.
- Laxer, James. *Oil*. Toronto, Groundwood Books/House of Anansi Press, 2008.

- Misona. "Are Sustainable Products Actually More Expensive?" *Misona*, 15 Feb. 2024, misona.co.uk/blogs/news/myth-or-fact-are-sustainable-products-actually-more-expensive. Accessed July 2024.
- Mukpo, Ashoka. "Massive Carbon Offset Deal with Dubai-Based Firm Draws Fire in Liberia." *Mongabay Environmental News*, 4 Aug. 2023, news.mongabay.com/2023/08/massive-carbon-offset-deal-with-dubai-based-firm-draws-fire-in-liberia/. Accessed 28 July 2024.
- Nardo, Don. *Climate Change*. Referencepoint Press, 19 June 2023.
- Polman, Paul, and Andrew S Winston. *Net Positive: How Courageous Companies Thrive by Giving More than They Take*. Boston, Ma Harvard Business Review Press, 2021.
- Rosane, Olivia. "'Blood Carbon': How a Kenyan Carbon Offset Project Harms Indigenous Pastoralists and Does Nothing about Climate Change." *EcoWatch*, 12 Apr. 2023, www.ecowatch.com/carbon-offset-kenya-greenwashing-indigenous-rights.html. Accessed 28 July 2024.
- Saito, Kohei. *Slow Down*. Astra Publishing House, 9 Jan. 2024.
- West, Darrell, and Christian Lansang. "Global Manufacturing Scorecard: How the US Compares to 18 Other Nations." *Brookings*, 10 July 2018, www.brookings.edu/articles/global-manufacturing-scorecard-how-the-us-compares-to-18-other-nations/. Accessed 28 July 2024.
- World Benchmarking Alliance. *Corporate Human Rights Benchmark across Sectors: Agricultural Products, Apparel, Automotive Manufacturing, Extractives & ICT Manufacturing*. 2020.
- Zhu, Raymond. "Have You Ever Considered Relocating Your China Manufacturing Operations to Another Country? When Buyers Talk about Sourcing from China, One of the Most Commonly Mentioned Challenges Is the Rising Cost of Labor. Double-Digit Increases in China's Minimum Wages Year-After-Year Have Driven Some Manufact." *Linkedin.com*, 9 Sept. 2016, www.linkedin.com/pulse/comparison-manufacturing-between-india-vs-china-william-zhu#:~:text=When%20it%20comes%20to%20manufacturing. Accessed 27 July 2024.

Examining Japanese Perceptions of Immigrants By Kaho Yamamoto

Abstract

As Japan faces an increasing labor shortage due to decreasing birthrate and an aging population, the idea of increasing immigration is often regarded to be one of the only solutions. Those opposing this increase in immigration in Japan cite the homogeneous population and subsequent xenophobic ideals of the country to depict Japan to be a country unfit for immigration. This paper conducts a literary review on what little research exists on attitudes toward immigration in Japan, making comparisons to societal attitudes and ideas of national identity in traditional immigrant-receiving countries and suggesting an alternative way to view Japanese national identity distinct from the commonly held views in these nations. By suggesting that the exclusionary views toward outsiders only applies for the purposes of considering a person to be “Japanese”, the paper comes to the conclusion that immigrants in itself, when considered to be a distinct but accepted group, do have a place in Japanese society. With this, the commonly-held belief of a xenophobic and anti-immigration Japanese population is reframed to reflect an accepting one, potentially changing outside views on Japanese society in itself and allowing immigrants to better integrate into Japanese society.

Introduction

Japanese views on immigration are commonly regarded to be negative, with the outsider perception being that they, as a largely monoethnic state, perceive immigrants as a cultural threat. The need for immigrants has become increasingly urgent as the aging population produces an unsustainable old age dependency ratio: the OECD reports that over the population of citizens aged 65 and over is over 50% comparative to that of citizens 20 to 64, and that this number is expected to rise to 79% in 2050 with current projections (2018). Many nations with similar issues of declining population have resorted to immigration to boost the laboring population, however Japan’s image as a xenophobic nation and the government’s passive efforts to encourage the practice has prevented Japan from adopting this practice on a larger scale.

Despite the growing need for foreign labor and focus on an international society, immigration policies on the unitary level are passive and scarce, with very little intended influence. However, Burgess (2015) argues that ideas such as 内向き志向 (inward-looking perspectives), which refers to the bleak future that awaits Japan due to the inward-looking youth, have merely been an excuse made by policymakers to justify their lack of effort to attract and accept more immigrants. This raises the question of how the Japanese population truly perceive immigrants. While studies are limited, many argue that in fact, the Japanese public are generally supportive of foreign laborers, albeit with a few restrictions (Simon and Lynch 1999; Takahashi 2015; Seebruck 2013). With this, there arises a need to examine where exactly this negative perception comes from, and study the variables that factor into Japanese perception of immigrants.

This paper seeks to examine existing studies on Japanese perception of immigrants, and understand the nuances behind the inconsistencies between negative common perception of Japanese attitudes, studies of positive perception, and passive policies on this matter. I argue that, while Japanese society views immigrants positively and recognizes their potential and existing contributions to Japanese society, they also see immigrants as a distinct group and value that distinction. With this, the Japanese public seeks to accept more immigrants, however they do not intend to fully integrate them into Japanese society.

Commonly Regarded Views on Immigrant Perception in Immigrant-Receiving Countries

Regardless of population patterns, there exists a pattern of a shift in status as an emigrant nation to an immigrant nation with economic and societal developments in various countries (Seol, Skrentny 2009). This is most commonly seen within the Western world, and in examining the causes for this pattern, it becomes clear why Japan, alongside other developed Eastern Asian countries, face barriers in this transition. Further, this allows for a more detailed view on immigrant perceptions in both traditional immigrant nations and in Japan.

Most research on immigration and the views concerning immigrants are centered around Western nations. In looking at the reasons for immigration, it is clear that Japan does not fit into the patterns typically seen in immigrant countries. Traditional immigrant countries such as the United States and Australia have been made up of immigrants since their start, and the recent influx of immigration in those nations have been a continuation of their origins, albeit accelerated due to globalization efforts. For these countries, the global image as an “immigrant country” and their relatively high acceptance of immigrants have played a large part in attracting immigrants. In regards to European nations, while they do not have a long history of immigrant acceptance, immigration has become more common with the presence of regional supranational organizations, most notably the European Union, pushing for immigration reforms for its member nations (Zhang 2018). In both cases, the receiving nations have policies and regulations in place for the practice, and whatever barriers there may exist, the process of immigration in itself and the path that exists upon doing so is generally well-documented and available. In comparison, Japan’s immigration policies remain relatively unknown to outsiders, and the path forward seems unstable (Seol, Skrentny 2009).

Of course, clear legal frameworks for the immigration process does not guarantee a smooth transition within the host nation. Group threat theory, originally by Bulmer and Blalock (as cited in Green 2017), and group conflict theory by Quillian 1995, Scheve and Slaughter 2001, and Mayda 2006 (as cited in Kobayashi et al. 2015), both state that tensions between majority and minority groups stem from the competition over resources, such as job opportunities and social welfare. This idea of resource scarcity and competition has generally been studied in Western nations as a source for immigrant rejection and apprehensive public attitudes. Group threat theory in particular is attributed to the racial tensions that came about in the Civil Rights Era in the United States (Green and Kadoya 2015). However, this can scarcely be attributed to Japan, especially in regards to job opportunities: job opportunities are not a scarcity for which

competition is rampant, especially among blue-collar industries (Kobayashi et al. 2015). While perceptions of racial hierarchy do seem to exist, there does not exist sufficient racially distinct groups to fully confirm the group threat theory in Japan (Green and Kadoya 2015, 68). Thus, the group conflict theory and group threat theory, which are often cited as a source for immigrant rejection in other immigrant-receiving nations, do not apply to Japan.

Additionally, the rapid influx of immigrants that is often attributed to anti-immigrant sentiment is also limited to western states. The salience-of-change hypothesis is a view arising from looking at Western immigrant-receiving nations that states a negative relationship between immigrant growth rate and immigrant tolerance due to the increased focus on issues that arise out of increased immigration and resources — while a general pattern similar to this has been confirmed in Japan, much like the group threat theory, the results were mixed due to the limited number of foreign groups within Japan (Green 2017). Sandovici, Jakobsen, and Strabac (2012) note a similar pattern in European countries, wherein increased immigration leads to more media coverage, political focus on immigration, and the issues that come about from it, resulting in more anti-immigrant sentiment. As immigration policy in Europe is largely encouraged by the European Union, they note that political leaders within the individual countries often make immigration a large campaign focus, making it a central issue in considering national identity and policy.

The differences between views on immigration also apply to pro-immigrant stances. For this, while the general theories seem to be applicable to Japan, the scarcity of immigrants in Japan at this time minimizes their impact. Person-positivity bias states that on an individual level, increased contact with foreigners has proved to have a positive effect on a Japanese person's perception of immigrants (Kobayashi et al. 2015). This is supported by the Contact Theory by Allport (1954), which states that contact between different groups has a negative correlation with prejudice. Under circumstances where this interaction is voluntary, in a collaborative environment, and among equals, Green (2015) has shown that this theory applies to Japanese immigrant communities. However, it is important to consider that these circumstances are rare compared to other immigrant nations.

While studies are scarce, there also seems to be a difference in immigrant cultural identity. While limited to nations built off of immigration, there seems to be a general attitude that a person is regarded to be of the host nation's nationality just by immigrating. This is especially the case in the United States, in which modern-day immigrants are compared to those who founded the nation – Gutierrez, the former US Secretary of Commerce, makes the case that “[t]oday's immigrants, just like our forefathers, arrive here seeking the ability to freely worship, to express themselves without fear of government redistribution, and to chart their own economic destiny” (2017). In comparison, similar attitudes of immigrant societal acceptance do not exist within Japan. Takahashi notes the use of “Zainichi (Resident Japanese) - [nationality]” rather than the hyphenated identities to articulate those who do not fit into this mold, indicating this separation even linguistically (2018, 564).

Economic Incentive for Japanese Immigration

So what, instead, incentivizes Japan towards becoming an immigrant nation? The need for immigrants comes primarily from the economic perspective – an ever-declining birthrate and the subsequent aging of the population has left Japan with an insufficient labor pool to support the economic development needed to stay competitive within the global market. The decrease in tax revenues and the increased need for social welfare for the elderly has made for an unsustainable system to support the needs of its citizens. As the small labor pool is forced to take on more labor, problems concerning overworking and “workaholic culture” lead to difficulties in childbearing and raising. Currently, Japan faces a bleak future as problems with declining birth rate, aging population, and economy feed into one another.

Various measures have been proposed to combat this issue, however the various factors that go into it makes it difficult for them to succeed. Pro-natalist policies such as promoting marriage and government payments after childbirth have widely been criticized as both insufficient and inefficient. As wages stagnate and costs of childcare continue to rise, the amount currently proposed of 500,000 yen, let alone the current amount of 420,000 yen, will only last so long in the decades-long process of caring for a child. Additionally, the core issue of social infrastructure such as nursery shortages and rising costs of after-school facilities is glossed over with the government’s current approach (Ryall 2022). However, fixing the structural issues are no easy task either, requiring a shift in societal views and overall agreements in the redistribution of government funds.

With this, a commonly proposed solution has become allowing in more immigrants to boost the working population and remedy the labor shortage. Early efforts such as by “foreign technical interns” under the Technical Intern Training Program, made up mostly by South Americans with Japanese ancestry in the 1990s, dramatically aided in the survival of companies requiring unskilled labor (Yokohama 2024). This practice, albeit on a minor scale, has continued and increased in popularity since then. In late 2018, there was an expansion of this program to accept increased numbers of both low-skill and high-skill workers. The focus on this program was solely labor, with Prime Minister Abe stressing the labor shortage and denying any correlation to relaxed immigration policy (Yamaguchi 2018). In the current day, it is common to see foreign immigrants working in convenience stores, construction, and recently, with increased focus on language ability, as English teachers.

Expanding on the simple advantage of increased labor with immigrant acceptance, most arguments made in favor of immigration in Japan boast the addition of skilled labor as a means to boost Japanese participation in global affairs. The “New Growth Strategy” (新成長戦略) proposed by the Japan Business Federation cites the acceptance of “highly skilled foreign professionals” to be one of the crucial steps towards a sustainable future (Japan Business Federation 2020). The federation also notes the importance of collaborating with the international world to re-establish an international economic order, with Japan taking an active role in the process. Historically, one of the biggest growths and modernizations of the Japanese economy happened in the Meiji Restoration, which came with the opening of borders and

citizenship rules. It sets the precedent for the successful use of immigrant labor to benefit the Japanese economy. (Lu, Menju, and Williams 2005). Thus, both Japanese policymakers and the population recognize the economic incentives for the acceptance of immigrants.

Cultural Incentive for Japanese Immigration

Additionally, the Japanese public recognizes that immigrant acceptance comes with the diversification of Japanese society that is needed especially as globalization progresses and internationalization efforts become more of a priority. Efforts toward “internationalization” have been present, facing a rapid increase especially in the 1980s with the bubble economy and continuing to this day (Lu, Menju, and Williams 2005). The idea that multicultural perspectives and ideals will add valuable contributions in the workplace and increase productivity and creativity has been integrated into Japanese company ideals. Adding onto the benefit of gaining more labor, the idea of Global Human Resources gives Japanese companies incentive to better acclimatize into the internationalizing landscape to take part in global business (Burgess 2015). By employing the 新成長戦略 (“New Growth Strategy”), businesses do see value in employing foreign labor, especially on a professional level.

On both unitary government and local community levels, the ideal of 他文化共生 (multicultural coexistence) has been promoted. The efforts promote a goal of a society in which “those of various nationalities and ethnicities recognize their differences, make efforts to maintain an equal-standing relationship, and coexist as essential members of the community” (Ministry of Internal Affairs and Communications 2006). This ideal is generally supported among the Japanese public, with 90% of those surveyed in the International Social Survey Programme 2003 favoring the idea of ethnic minorities maintaining their culture rather than adapting to the majority (Takahashi 2015). The only conditions under which these cultural differences were seen as an issue were when they happened to be visible and not in line with Japanese social norms, as studied in school settings by Takahashi (2015). In this study, Takahashi states that “Japan is exceptional in that it has outstandingly low (high) percentage of the people favouring the idea of cultural adaptation (maintenance of the ethnic minority’s culture), regardless of one’s conceptualization of national identity” (2015, 192). While there exists the barrier which stipulates that these minority cultural practices must be done in private and must adhere to Japanese norms, there exists strong evidence that the Japanese people welcome those with varying cultural backgrounds and practices.

Additionally, the Japanese public recognizes that immigrant acceptance comes with the maintenance of a positive image as a member of international society. Especially following the end of World War II, Japan has recognized the importance of its international public perception. International human rights laws have allowed for the expansion of Japanese citizenship laws and immigrant legal protections, and Japan’s cooperation within supranational organizations has allowed for a more open view towards immigration in general (Lu, Menju, and Williams 2005). Japanese citizens have been encouraged to take the lead globally in with movements toward “fostering of global citizens” (国際人育成) (Burgess 2015), and efforts such as the Sister City

Movement have allowed for a “more global conception of citizenship” (Lu, Menju, and Williams 2005). These efforts go beyond the borders of Japan, with Japanese emigrant communities (predominantly in South America) seen as bridges between the two cultures (Minamikawa 2017). With this, Japan’s desire to expand beyond citizens of Japan, and into citizens of international society, becomes clear.

However, the homogenous population and the xenophobic attitudes that come with it is a popular conception of the Japanese people that hinders this effort to accept more immigrants. While an economic threat has never been considered to be a factor, it is widely recognized that “the politics of immigration and belonging at the mass level in Japan are at least as much about economic concerns as the alleged threat that immigrants may pose to ethno-national identity” (Kobayashi et al. 2015, 18). However, the idea of the ethnic homogeneity has been argued to be a lesser factor than generally regarded, such as by Kashiwazaki (2009), who claims that the permanent structural inequalities are what truly keeps more immigrants from entering.

Exclusivity of Japanese National Identity – Ethnicity

While the negative perception of Japanese immigrant inclusion has little evidence, the exclusivity of Japanese national identity remains. Historically, its inaccessibility as an island country has allowed for limited outside communication. The Sakoku policy in the Tokugawa period, which allowed only a select few ports to be open to limited traders, further made Japan into a nation of a culturally homogeneous population. As Japan opened its borders and started to imperialize, with it came a rise in ethno-nationalism, which was promoted as a means of defense against western powers. Thus, Japanese identity came to be a “demonstration of state power and cultural superiority” (Lu, Menju, and Williams 2005). This makes many traditionally accepted theories of immigrant perception commonly accepted in traditionally immigrant-receiving nations, inapplicable to Japan.

In the current day, many studies show that ethnic and civic identities are crucial to the Japanese people’s perception of foreigners. Takahashi (2015) shows that between ethnic (born in Japan, Japanese nationality, Japanese ancestry) and civic (respecting the law, feeling oneself as Japanese), the Japanese people place more importance on the former in regards to nationality. In a 2022 study, Woo surveyed the importances placed in various factors for whether a person should be granted citizenship when offered a combination of: being born in Japan; being raised in Japan; having a Japanese father; having a Japanese mother. Woo found that the most emphasis was placed on the citizenship of the parents, especially along the paternal lines. With this, when it comes to perceiving foreign immigrants as “Japanese”, Japanese ancestry is highly valued as a factor.

Exclusivity of Japanese National Identity – Culture

However, Japanese ancestry is not the sole determinant in whether or not a person is considered to be “Japanese”. Nikkei people, the term for which describes ethnic Japanese that reside overseas (predominantly in Brazil), have been allowed easier immigration access to Japan

with the amendment to the Immigration Law in 1990. While the assimilation of these Nikkei immigrants were expected to be easy due to their Japanese ancestry, they were seen to be culturally different and therefore struggled to adjust to Japanese society. Ishii (2005) brings up the views of two scholars on the subject: Tsuda saw Brazilians to be “strangers in the ethnic homeland”, describing how they were othered and stereotyped to be impoverished, and Brody saw the people as a cause for an acceleration of a “crisis” in Japanese society. Thus, cultural and societal differences also make up the criteria of what it means to be “Japanese”.

This rigid definition is also evident in the study conducted by Lu, Menju, and Williams (2005) on the diversity within the domestic borders of Japan. The indigenous people of the Ainu (now referred to as “Hokkaido”), coming under Japanese control 1869 under the Meiji government, are a distinct cultural group that were not formerly belonging to the Japanese people. The Meiji government employed a policy of assimilation under measures such as the Hokkaido Former Aboriginal Protection Act in 1898 — because they were now legally a part of the Japanese people, everything that made them distinct were deemed necessary to strip away. This is comparable to the treatment of immigrants, in the way that immigrants are not considered “Japanese” due to their distinct cultural practices. A similar process occurred for the Okinawans when they were integrated into Meiji Japan in 1879, and like the Ainu, faced various policies aimed at assimilation. While historically a quasi-independent territory, this territorial integration called for strict assimilation policies such as religious monuments and language education. The indigenous groups of Japan, much like immigrants in the current day, faced much difficulty for their distinct cultural identities. However, while immigrants are made into a distinct group within Japanese borders, because they were forcibly taken into Japanese control and were considered to be Japanese citizens, the government felt it necessary to employ these assimilation measures. In other words, there is a choice that “outsiders” such as immigrants and indigenous people are forced to make: be viewed as an “Other”, or strip away their cultural identities to fully be accepted into the society. With this, it is clear that cultural and social differences make it difficult to fit into this rigid definition of what is considered to be “Japanese”, regardless of ethnicity or bloodline.

Japanese View of Immigrants as a Distinct Group

Thus, both ethnic and civic senses of identity may be needed for immigrants to be fully accepted into Japanese society. Takahashi (2018) has shown that a combination of ethnic and civic senses of identity are valued. However, this offers a very narrow definition of Japanese national identity that, in essence, allows only for ethnic Japanese that grow up in the country to be considered within it. Despite general Japanese acceptance of immigrants, this exclusive definition of Japanese national identity indicates that these immigrants will hardly ever be accepted as a “true” Japanese, instead being seen in their own category as immigrants.

This reluctance to consider immigrants “Japanese” has been studied as an alternative way to nationalism as opposed to American and European ideas of nationalism. Nihonjinron (日本人論), first proposed by Yoshino, highlights this strict definition as a “nationalism concerned with

the preservation and maintenance of national identity, against a ‘primary nationalism’ that tries to effect the creation of a nation that does not yet exist” (Takahashi 2015, 72). While any person can *present* Japanese traits – such as being able to communicate in the Japanese language and having the collectivist ideals commonly associated with Japanese culture (Takahashi 2018) – having those traits does not necessarily *make* that person Japanese. Yet the lack of social integration policy in Japan signifies that the Japanese public does not seek to assimilate any perceived outsiders into the community (Takahashi 2015). There remains the idea that foreign exchange, whether it be for schooling, labor, or otherwise, is a temporary, one-time event that does not require full integration (Lu, Menju, and Williams 2005). This ideal is almost unique to Japan, producing the idea of “tertiary nationalism”, wherein the preservation of rule-based culture is valued above incorporation. The maintenance of immigrant ethnic identity and traditional culture is encouraged to be done in private, however the social contract of Japanese society must be followed (Takahashi 2015).

The legal definition of being “Japanese”, or on citizenship, also exemplifies these restrictive ideals of Japanese nationalism. The Japanese public places importance on citizenship as one of the greatest factors in considering a person “Japanese”, with it coming first in importance followed by self-identification, Japanese language abilities, being born in Japan, respect for Japanese laws, residence in Japan, Japanese ancestry, and adoption of Japanese customs in Takahashi’s analysis of the International Social Survey Programme’s 2013 data (2018). This may correlate to the fact that the naturalization process in Japan requires extreme commitment, and obtaining it is seen as evidence enough that a foreign immigrant could be considered “Japanese”. A notable part of the Japanese citizenship law is its ban on multiple nationalities – those who hold multiple citizenships by birth must choose one by the age of 20, and those who get naturalized must renounce the one they previously held (The Ministry of Justice n.d.). This stipulation especially is regarded to deter many from the Japanese citizenship itself, extending beyond the naturalization process with Takeuchi (2024) arguing that it also leads to the loss of Japanese citizens who are successful abroad.

Efforts by the national government remain passive. Local efforts such as multicultural education initiatives and ethnic classes are present in areas with comparatively high immigrant populations, even leading to the formation of “Alliance of Cities with a high Density of Foreign Citizens” (Takahashi 2020, 2015). Exchange programs have also been encouraged as English education continues to rise in importance, and efforts to ensure a more global sense of citizenship among the residents of those municipalities has been valued (Lu, Menju, and Williams 2005). However, with this comes the passiveness of the Ministry of Education, which handles such matters, as it comes to rely on local governments (Green 2014).

Conclusion

As the aging population and declining birth rate creates an increasingly unsustainable old age dependency ratio in Japan, immigration remains one of the most viable solutions. However, the homogenous Japanese population and the commonly-held belief of a xenophobic Japanese

public leads to much of the international public remaining skeptical of the Japanese people's attitudes toward immigrants. The focus on Japanese exclusivist image has allowed this largely negative view to persist. Further, much of the existing research tries to apply western perceptions of immigrants and national identity onto a nation in which many historic and political backgrounds vary greatly to traditional immigrant-receiving nations such as the United States and nations within the European Union.

The Japanese do recognize the economic and societal benefits of opening their borders. While there is some truth to this common perception in that the Japanese people hold very exclusivity ideas of Japanese national identity and what it means to be "Japanese", this paper, through a literary analysis of existing works, find that this exclusivity only applies in regards to fully assimilating its immigrants to become "Japanese". In fact, as long as the immigrants contribute positively to the economy, Japan participates in international settings, or both, the Japanese public is generally supportive of accepting immigrants when viewing them as a distinct group. This distinction brings both positives and negatives – it allows immigrants to be supported in maintaining their own cultures and indicates better treatment of immigrants than traditionally regarded, however social interactions between majority (native Japanese) and minority (immigrants) are scarce. Regardless, this idea suggests the possibility for a change in general attitudes of people seeking immigration to Japan and subsequently the potential rise in immigration in Japan.

Works Cited

- Allport, Gordon W. *The Nature of Prejudice*. Addison-Wesley Publishing Company, 1954.
- Burgess, Chris. "To Globalise or Not to Globalise? 'Inward-Looking Youth' as Scapegoats for Japan's Failure to Secure and Cultivate 'Global Human Resources.'" *Globalisation, Societies and Education*, vol. 13, no. 4, 2015.
<https://research.ebsco.com/linkprocessor/plink?id=deeecca90-ea42-308b-9d22-eb07e4158de7>.
- Green, David. "Education of Foreign Children in Japan: Local Versus National Initiatives" *Journal of International Migration and Integration*, vol 15, 2015, pp. 387-410.
<https://link.springer.com/article/10.1007/s12134-013-0299-z>.
- Green, David. "Immigrant Perception in Japan: A Multilevel Analysis of Public Opinion." *Asian Survey*, vol. 57, no. 2, 2017, pp. 368-394. <https://www.jstor.org/stable/26367754>.
- Green, David, and Yoshihiko Kadoya. "Contact and Threat: Factors Affecting Views on Increasing Immigration in Japan." *Politics & Policy*, vol. 43, no. 1, 2015, pp. 59-93.
<https://doi.org/10.1111/polp.12109>.
- Gutierrez, Carlos. "Immigrants Put America First: In Coming Here, They Affirm Our Values" *The Catalyst*, Issue 08 , 2017.
<https://www.bushcenter.org/catalyst/democracy/gutierrez-immigrants-validate-values>.
- Ishii, Yuka. "Beyond the Illusion of Homogeneity: Inward-looking 'Japanese' in the Age of Globalization" *Social Science Japan Journal*, vol. 8, no. 2, 2005, pp. 267–271.
<https://doi.org/10.1093/ssjj/jyi035>.
- Japan Business Federation. "新成長戦略" ["New Growth Strategy"]. 日本経済団体連合会 [Japan Business Federation] 2020. <https://www.keidanren.or.jp/policy/2020/108.html>.
- Kashiwazaki, Chikako. "Gurōbaru-jidai no Nihonshakai to Kokuseki" ["Japanese Society and Nationality in the Age of Globalization"] *Social Science Japan Journal*, vol. 12, no. 1, 2009, pp. 179-182. <https://doi.org/10.1093/ssjj/jyn059>.
- Kobayashi, Tetsuro, Christian Collet, Shanto Iyengar, and Kyu S. Hahn. "Who Deserves Citizenship? An Experimental Study of Japanese Attitudes Toward Immigrant Workers." *Social Science Japan Journal*, vol. 18, no. 1, 2015, pp. 3-22.
<https://doi.org/10.1093/ssjj/jyu035>.
- Lu, Catherine, et al. "Japan and 'the Other': reconceiving Japanese Citizenship in the Era of Globalization." *Asian Perspective*, vol. 29, no. 1, 2005, pp. 99–134.
<http://www.jstor.org/stable/42704492>.
- Minamikawa, Fuminori. "戦後期における出入国管理体制の成立と「非移民国」日本" ["The Making of Postwar Immigration Control Regime and an Origin of Japan as a 'County of Non-immigration'"]. 立命館言語文化研究 [Ritsumeikan Studies in Language and Culture], vol. 29, 2017, pp. 137-144.
https://www.ritsumei.ac.jp/acd/re/k-rsc/lcs/kiyou/pdf_29-1/lcs_29_1_minamikawa.pdf.
- Ministry of Justice (Japan). "The Choice of Nationality." *Ministry of Justice (Japan)*.
<https://www.moj.go.jp/content/001343325.pdf>.

- Ministry of Internal Affairs and Communications (Japan). “多文化共生の推進に関する研究会報告書—地域における多文化共生の推進に向けて” [A Study Group Report on the Promotion of Multicultural Symbiosis: Towards the Promotion of Multicultural Symbiosis in the Region]. *Ministry of Internal Affairs and Communications*, 2006. https://www.soumu.go.jp/kokusai/pdf/sonota_b5.pdf.
- Organisation for Economic Co-operation and Development. “Working Better with Age: Japan.” *Ageing and Employment Policies*. Paris, OECD Publishing, 2018. <https://doi.org/10.1787/9789264201996-en>.
- Ryall, Julian. “Will Japan’s New Plan to Boost Birth Rates Work?” *DW News*, 14 Dec. 2024, <https://www.dw.com/en/will-japans-new-plan-to-boost-birth-rates-work/a-64091588>.
- Sandovici, Maria Elena, et al. “Political Nationalism and Attitudes towards Immigration: The Interaction of Knowledge and Policy.” *International Journal on Minority and Group Rights*, vol. 19, no. 2, 2012, pp. 115-127. https://brill.com/view/journals/ijgr/19/2/article-p115_1.xml?language=en&ebody.
- Seebruck, Ryan. “Technology and Tolerance in Japan: Internet Use and Positive Attitudes and Behaviors Toward Foreigners.” *Social Science Japan Journal*, vol. 16, no. 2, 2013, pp. 279-300. <https://doi.org/10.1093/ssjj/jyt017>
- Seol, Dong-Hoon, and John D. Skrentny. “Why Is There So Little Migrant Settlement in East Asia?” *The International Migration Review*, vol. 43, no. 3, 2009, pp. 578–620. <http://www.jstor.org/stable/20681722>.
- Simon, Rita J., and James P. Lynch. “A Comparative Assessment of Public Opinion toward Immigrants and Immigration Policies.” *The International Migration Review*, vol. 33, no. 2, 1999, pp. 455–467. <https://doi.org/10.2307/2547704>.
- Takahashi, Fumiko. “誰を「日本人」らしいと見なすのか：多文化社会におけるナショナルアイデンティティと教員” [“Who is regarded as ‘Japanese’?: National Identity and Teachers in Multicultural Society”]. *UTokyo Repository*, vol. 58, 2017, pp. 563-582. https://repository.dl.itc.u-tokyo.ac.jp/record/52047/files/edu_58_47.pdf.
- Takahashi, Fumiko. *Integration and Separation of Immigrants in Japan*. 2015. University of Oxford, PhD thesis. <https://ora.ox.ac.uk/objects/uuid:53b34de4-6d8c-4079-82ba-857bdaf0e6f8>.
- Takeuchi, Maiko. “Japan to Cut Off Japanese? Protecting the Ties between Japan and the Japanese Citizens Who are Successful Overseas.” *Research Institute of Economy, Trade, & Industry*. 12 Jan, 2024. https://www.rieti.go.jp/en/columns/s24_0008.html.
- Yamaguchi, Mari. “Japan OKs divisive bill allowing more foreign workers.” *Associated Press*. 7 Dec, 2018. <https://apnews.com/general-news-f312b64eeaf2402d8f85733c35ce41bc>
- Yokohama, Erika. “Record 2 Million Workers are Changing the Face of Japan.” *The Japan Times*, 7 Feb. 2024. <https://www.japantimes.co.jp/news/2024/02/07/japan/society/japan-mass-foreign-immigration/>.
- Zhang, Jie. “移民受入の国際社会学—選別メカニズムの比較分析” [“International Sociology of Immigration Acceptance: Comparative Analysis of Selection Mechanism”]. *Social Science Japan Journal*, vol. 21, no. 2, 2018, pp. 351–354. <https://doi.org/10.1093/ssjj/jyy009>

Exploring the Absence and Double Peaks of Alpha Activity in Resting-State EEG

By George Rohan

0 | Abstract

This research investigates the dynamics of resting-state alpha activity in surface electroencephalography (EEG) and the implications of its role as a biomarker. Specifically, it addresses the prevalence of absent alpha activity and the occurrence of double peaks within the alpha frequency range (8-12 Hz). Utilizing a dataset from OpenNeuro, which includes EEG recordings from 110 healthy participants, the distribution of resting-state alpha activity is quantified, and the characteristics of double peaks are analyzed.

The analysis reveals that 20.9% of participants exhibited no detectable resting-state alpha activity, while 79.1% displayed measurable alpha oscillations. Among those with detectable alpha activity, 39.5% exhibited double peaks, suggesting complex neural oscillation patterns. These findings highlight the necessity of considering individual differences in neurophysiological assessments, as the absence of alpha activity may have significant implications for the detection of various conditions.

Moreover, this study discusses the potential role of double peaks in reflecting distinct cognitive states or strategies, emphasizing their relevance in understanding cognitive efficiency. The mechanisms underlying these phenomena remain poorly understood, indicating a critical need for further research to elucidate their functional significance.

In conclusion, this study contributes to a deeper understanding of resting-state alpha activity dynamics and its implications for cognitive function and clinical practice. By systematically investigating both absent alpha activity and double peaks across diverse populations, we aim to enhance diagnostic practices and inform targeted interventions.

1 | Introduction

The exploration of resting-state alpha activity in electroencephalography (EEG) has garnered significant interest due to its implications for cognitive function and neural efficiency. Alpha activity, primarily oscillating within the 8-12 Hz frequency range, serves as a crucial marker for various cognitive processes, including attention, memory, and self-regulation (Bazanov & Vernon, 2013). Its absence or atypical patterns have been associated with a range of neurological and psychiatric disorders, highlighting the need for a deeper understanding of its role in brain health and functionality. This research aims to address the gaps in the existing literature regarding the prevalence and implications of absent alpha activity and the phenomenon of double peaks in alpha oscillation.

1.1 | The Importance of Alpha Activity and Implications of Absent Alpha Activity

Importance of Alpha Activity in Cognitive Function

Research has shown that increased alpha amplitude is associated with optimal cognitive performance, reflecting a state of mental readiness and focus. For instance, studies have

demonstrated that individuals with higher individual alpha peak frequency (IAPF) tend to perform better on cognitive tasks compared to those with lower IAPF levels (Bazanova & Aftanas, 2008). This shows that alpha activity actively contributes to the facilitation of cognitive processes rather than being a passive result of brain function. Moreover, the modulation of alpha activity through neurofeedback training has been linked to improvements in cognitive performance, indicating its potential as a target for enhancing cognitive abilities (Angelakis et al., 2007). Understanding the dynamics of alpha activity can therefore provide valuable insights into the neurophysiological mechanisms underlying cognitive function and may inform interventions aimed at optimizing mental performance.

Alpha Activity as a Biomarker for Mental Health

The absence or abnormal patterns of resting-state alpha activity have been implicated in various neurological and psychiatric disorders, making it a potential biomarker for mental health conditions. For example, individuals diagnosed with Attention Deficit Hyperactivity Disorder (ADHD) often exhibit reduced alpha amplitude and altered alpha frequency patterns compared to healthy controls (Barry et al., 2003). Similarly, research has linked low resting-state alpha activity to depression and anxiety disorders, suggesting that deficits in alpha oscillations may reflect underlying cognitive dysfunctions or neural inefficiencies (Loo et al., 2010). The identification of these patterns can aid in early diagnosis and monitoring of treatment efficacy. Furthermore, biofeedback interventions aimed at enhancing alpha activity have shown promise in improving symptoms associated with these disorders, highlighting the therapeutic potential of targeting alpha oscillations in clinical settings (Bazanova et al., 2009). Thus, exploring the relationship between alpha activity and mental health can contribute to a better understanding of the neural correlates of psychological conditions and inform therapeutic approaches.

The Role of Alpha Activity in Self-Regulation

Alpha activity is closely associated with self-regulation processes, which are essential for effective cognitive control and emotional regulation. Research indicates that increased alpha power is linked to enhanced self-control capabilities, allowing individuals to manage their thoughts and emotions more effectively (Bazanova & Aftanas, 2010). This relationship is particularly evident during tasks requiring sustained attention or inhibition of impulsive responses. For instance, higher levels of alpha activity during resting states have been correlated with improved performance on tasks requiring self-control (Klimesch et al., 2007). Neurofeedback training targeting alpha enhancement has also been shown to facilitate better self-regulation skills by promoting an optimal state of arousal and attentional focus (Angelakis et al., 2007). Understanding the interplay between alpha activity and self-regulation can provide insights into how individuals can harness their neural resources to enhance cognitive performance and emotional well-being.

Individual Variability in Alpha Activity

Individual differences in alpha activity are significant factors influencing cognitive performance and emotional regulation. Variability in IAPF has been shown to correlate with cognitive abilities such as working memory and processing speed (Clark et al., 2004). Moreover, factors such as age, hormonal status, and genetic predispositions can influence IAPF and overall alpha power (Bazanov & Aftanas, 2008). For instance, studies indicate that IAPF tends to increase during childhood until puberty and then gradually declines after age 40 (Bazanov et al., 2011). This variability highlights the importance of considering individual differences when interpreting EEG data related to alpha activity. Tailoring interventions based on an individual's unique neurophysiological profile could enhance the effectiveness of neurofeedback training or other cognitive enhancement strategies. Therefore, understanding the sources of variability in alpha activity is crucial for developing personalized approaches to optimizing cognitive function.

1.2 | Double Peaks in Alpha Activity

The phenomenon of double peaks in alpha activity, characterized by the presence of two distinct peaks within the alpha frequency range, has emerged as an intriguing area of study in EEG research. These double peaks may reflect underlying cognitive processes or brain states that differ from typical single-peaked alpha activity. The presence of double peaks can indicate a more complex neural oscillation pattern, potentially associated with enhanced cognitive flexibility or dual-task processing capabilities (Bazanov & Aftanas, 2010). This suggests that individuals exhibiting double peaks may possess unique cognitive strategies or neural efficiencies that differentiate them from those with standard alpha activity.

Despite the potential significance of double peaks in understanding cognitive function, there remains a notable gap in the literature regarding their prevalence and implications. While some studies have reported the occurrence of double peaks in specific populations, such as musicians or individuals engaged in complex cognitive tasks, comprehensive investigations into the prevalence of this phenomenon across diverse populations are limited (Bazanov et al., 2008). Furthermore, existing research primarily focuses on the amplitude and frequency characteristics of alpha activity without adequately addressing the functional relevance of double peaks. This oversight hinders our understanding of how these oscillatory patterns relate to cognitive performance and mental health.

Moreover, the mechanisms underlying the emergence of double peaks in alpha activity are not well understood. Researchers have proposed various hypotheses, but empirical evidence supporting these hypotheses is scarce. The lack of clarity surrounding the conditions that facilitate double peak formation further complicates efforts to elucidate their functional significance.

To address these gaps, future research should aim to systematically investigate the prevalence and characteristics of double peaks in various populations while exploring their relationship with cognitive performance and mental health outcomes. Additionally, studies should focus on elucidating the neurophysiological mechanisms that give rise to these oscillatory patterns. By advancing our understanding of double peaks in alpha activity, we can enhance our

comprehension of brain function and its implications for cognitive processing and clinical practice.

1.3 | Addressing the Gap

The aims of this research are to address the existing gaps in the literature regarding the prevalence and implications of absent alpha activity and double peaks in alpha oscillation. By analyzing EEG data from an OpenNeuro dataset, I will quantify the percentage of individuals exhibiting resting-state alpha activity and those displaying double peaks within the alpha frequency range. This analysis will involve plotting the EEG data to visualize patterns and identify significant trends. Following this, I will comment on the findings, discussing their relevance to cognitive function and mental health, thus contributing to a deeper understanding of alpha activity dynamics in diverse populations.

2 | Dataset Description

The dataset in question, taken from OpenNeuro, comprises resting-state EEG recordings collected from 111 healthy subjects, with a mean age of 37.6 years (range: 17-71 years, SD: 14.0). Each participant underwent EEG recording using a 64-channel BioSemi ActiveTwo system, adhering to the extended 10-20 electrode placement system. The data collection involved a segment of four minutes of resting-state EEG with participants instructed to keep their eyes closed, ensuring a controlled environment for capturing spontaneous brain activity. The dataset is structured according to the Brain Imaging Data Structure (BIDS) specification, which facilitates easy access and interoperability for researchers in various fields, including cognitive neuroscience and clinical neurophysiology.

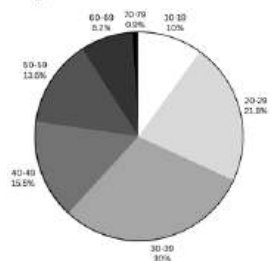
2.1 | Dataset's Original Purpose

The EEG data were collected as part of an experimental paradigm investigating long-term potentiation-like modulation of sensory-evoked potentials. This involved a series of tasks that included visual and auditory stimuli, with the resting-state EEG recorded approximately 45 minutes into the session. The dataset also includes neuropsychological test scores for each subject, allowing for a comprehensive analysis of the relationship between brain activity and cognitive performance.

Sex Distribution



Age Distribution



2.2 | Dataset's Suitability for Study

This dataset is particularly suitable for research focused on resting-state EEG due to its ample sample size and the high quality of the recordings.

2.3 | Modifications to Dataset

For the purpose of this paper, only the resting-state EEG data from the first 110 subjects has been considered, ignoring the 111th recording. No relevance has been given to the second EEG reading for the participants. The participants in the modified dataset have an average age of 37.35, with a 37.3 : 62.7 ratio of males to females.

3 | Methodology

3.1 | Data Conversion and Preparation

Upon downloading the dataset, data was first processed and converted the .set files into the MATLAB-compatible .mat files for subsequent analysis. Then, each subject's EEG recording was loaded into the MATLAB environment, with a dynamic file path generated based on a user prompts input of the subject identifier. The EEG data was stored on a multi-dimensional matrix, where each dimension represented electrode channels, epochs, and samples respectively.

The data structure was then reshaped to a two-dimensional matrix where:

- The first dimension corresponded to the electrode channels
- The second dimension combined the samples across epochs into a continuous time series for each electrode channel, converting each into a vector

3.2 | Frequency-Domain Analysis Through Fast Fourier Transform

The Fast Fourier Transform (FFT) is an efficient algorithm for computing the discrete Fourier transform and its inverse. The primary function of the FFT is to convert time-domain signals, which represent how a signal varies over time, into their frequency-domain equivalents, allowing for the identification of power in the alpha frequency range.

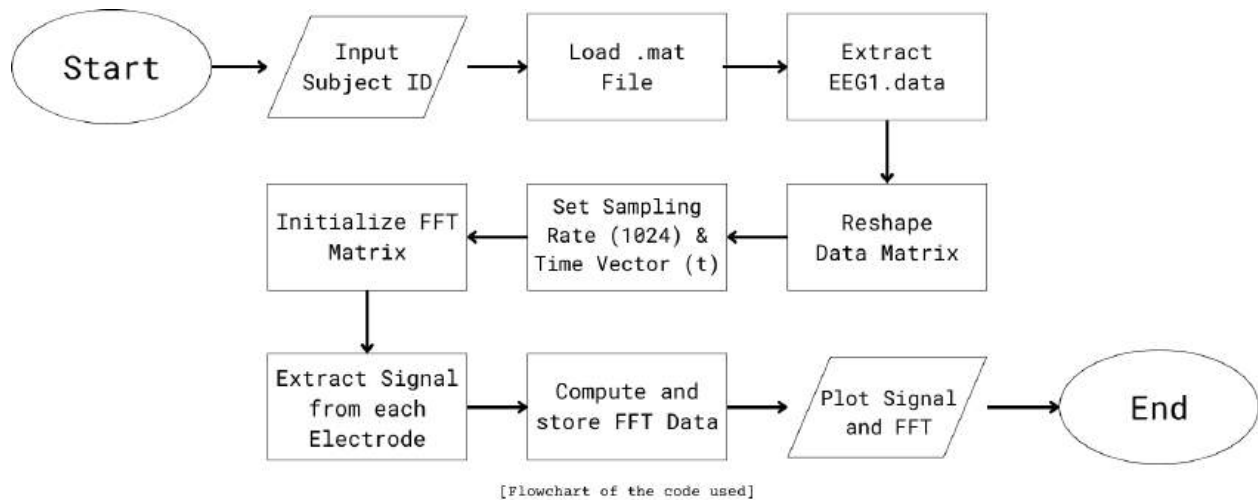
For each of the 64 electrodes, the signal was first extracted and passed to the FFT function to compute the frequency spectrum. The output was then normalized by dividing it by the length of the signal, after which the first half of the data (the single-sided spectrum) was selected, computed, and stored in a matrix.

3.3 | Visualization of EEG Signals

To visually inspect the neural activity for evaluation and to check for data clarity, the MATLAB script visualized both the time-domain and frequency-domain data for each electrode channel.

Time-Domain Plot: The original EEG signal from one electrode was plotted against time, providing an overview of the signal's amplitude variations and general characteristics over the duration of the recording. This plot helped identify any anomalies or trends in the raw signal.

Frequency-Domain Plot: The magnitude of the FFT was plotted against the frequency (in Hz) for the same electrode. This visualization allowed for an assessment of the spectral content of the EEG data, with special attention given to the alpha frequency range. A smoothing function of 80 pts was applied to the frequency data to improve the clarity of the spectrum, highlighting key oscillatory components and eliminating noise.



3.4 | Observational Criteria

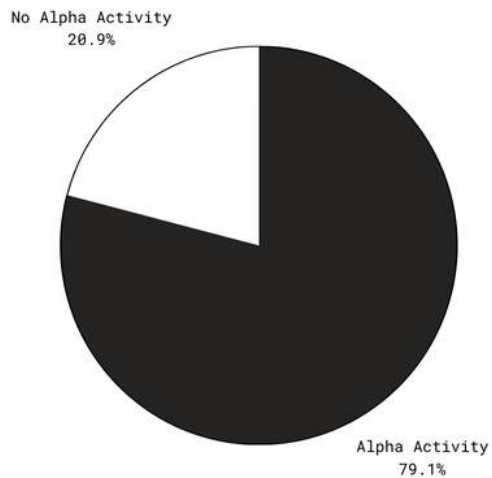
The following criteria were evaluated when looking at each subject's EEG frequency spectra:

- a) Alpha Band Power: The presence of significant power within the approximate range of 8-12 Hz (± 0.5 Hz) was assessed
- b) Double Peak in Alpha Band: In some cases, the alpha band exhibited a double (or triple) peak within the expected frequency range

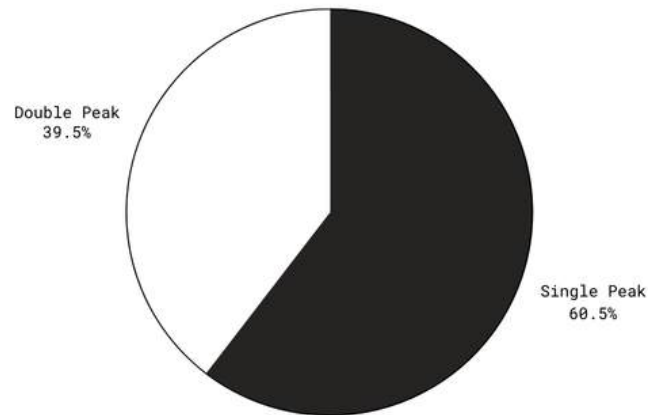
4 | Results

Out of the total participants analyzed, 20.9% ($n = 23$) did not exhibit any resting-state alpha activity, while 79.1% ($n = 87$) displayed detectable alpha activity. Among those with alpha activity, 39.5% ($n = 34$) exhibited two or more peaks in the alpha frequency range. The corresponding pie charts illustrate these distributions.

Distribution of Participants displaying Alpha Activity



Distribution of # of Peaks



4.1 | Insights on Alpha Activity Distribution

The analysis of alpha activity distribution reveals that 20.9% of participants exhibited no detectable resting-state alpha activity, while 79.1% displayed measurable alpha oscillations. This finding emphasizes the need to consider the substantial proportion of individuals lacking detectable alpha activity in neurophysiological assessments. Given that alpha activity serves as a critical marker for various cognitive processes, including attention and memory, the absence of such activity may have significant implications for cognitive performance and mental health.

Furthermore, the presence of alpha activity is often utilized as a biomarker in clinical settings, making it essential to account for individuals who do not conform to typical alpha patterns. This underscores the importance of adopting a more inclusive approach in diagnostic practices and research methodologies. Understanding the dynamics of alpha activity can provide valuable insights into individual differences in cognitive functioning and inform targeted interventions aimed at optimizing mental performance and enhancing overall brain health.

4.2 | Insights on Double Peak Distribution

The observation that 39.5% of participants with detectable alpha activity exhibited double peaks introduces a significant aspect to our understanding of cognitive processing. This phenomenon may reflect complex neural oscillation patterns, potentially associated with enhanced cognitive flexibility or dual-task processing capabilities (Bazanov & Aftanas, 2010).

Previous studies have suggested that variations in alpha peak frequency and amplitude are linked to different attentional and memory-related tasks (Bazanov & Aftanas, 2010). The presence of double peaks could indicate distinct cognitive states, allowing for a more nuanced interpretation of resting-state EEG data. Understanding these patterns may enhance our

comprehension of how individuals process information during resting states and their implications for cognitive efficiency.

Moreover, the mechanisms underlying the emergence of double peaks in alpha activity remain an area requiring further exploration. Investigating these mechanisms could provide deeper insights into the functional significance of double peaks in alpha activity and their relationship to cognitive performance and mental health outcomes. Thus, future research should focus on systematically examining the prevalence and characteristics of double peaks across diverse populations, contributing to a more comprehensive understanding of their role in cognitive processing.

5 | Conclusion

In summary, this research underscores the critical importance of investigating both the absence of alpha activity and the presence of double peaks in alpha oscillation. The findings indicate that a significant portion of individuals—approximately 20.9%—exhibit no detectable resting-state alpha activity. This absence may have profound implications for cognitive performance and mental health, as alpha activity is closely linked to various cognitive processes, including attention, memory, and self-regulation. Recognizing this substantial group within neurophysiological assessments is essential for developing inclusive diagnostic practices and targeted interventions.

Additionally, the discovery that 39.5% of participants with detectable alpha activity displayed double peaks introduces a new dimension to our understanding of cognitive processing. These double peaks may reflect distinct cognitive states or strategies, potentially indicating enhanced cognitive flexibility or efficiency. Understanding the mechanisms underlying these patterns is crucial for elucidating their functional significance in relation to cognitive performance and mental health outcomes.

The gaps in existing literature regarding both absent alpha activity and double peaks highlight the need for further research. Systematic investigations into the prevalence, characteristics, and implications of these phenomena across diverse populations are warranted. Such studies could provide valuable insights into the neurophysiological mechanisms that govern cognitive processes and inform clinical practices aimed at optimizing mental health and cognitive performance.

In conclusion, advancing our understanding of resting-state alpha activity dynamics is vital for enhancing clinical assessments and interventions. By exploring these critical aspects of EEG research, we can better address individual differences in cognitive functioning and develop more effective strategies to support mental health and cognitive enhancement.

Works Cited

- Bazanova, O. M., and D. Vernon. "Interpreting EEG Alpha Activity." *Neuroscience & Biobehavioral Reviews*, vol. 44, May 2013, pp. 94–110.
<https://doi.org/10.1016/j.neubiorev.2013.05.007>.
- Bazanova, O. M., and L. I. Aftanas. "Individual Measures of Electroencephalogram Alpha Activity and Non-verbal Creativity." *Neuroscience and Behavioral Physiology*, vol. 38, no. 3, Feb. 2008, pp. 227–35. <https://doi.org/10.1007/s11055-008-0034-y>.
- Angelakis, Efthymios, et al. "EEG Neurofeedback: A Brief Overview and an Example of Peak Alpha Frequency Training for Cognitive Enhancement in the Elderly." *The Clinical Neuropsychologist*, vol. 21, no. 1, Oct. 2006, pp. 110–29.
<https://doi.org/10.1080/13854040600744839>.
- Loo, Sandra K., et al. "Familial Clustering and DRD4 Effects on EEG Measures in Multiplex Families With Attention-Deficit Hyperactivity Disorder." *PubMed Central (PMC)*, 1 Apr. 2010, www.ncbi.nlm.nih.gov/pmc/articles/PMC2919766.
- Bazanova, O. M., et al. "Biofeedback in Psychomotor Training. Electrophysiological Basis." *Neuroscience and Behavioral Physiology*, vol. 39, no. 5, May 2009, pp. 437–47.
<https://doi.org/10.1007/s11055-009-9157-z>.
- Bazanova, O. M., and L. I. Aftanas. "Individual EEG Alpha Activity Analysis for Enhancement Neurofeedback Efficiency: Two Case Studies." *Journal of Neurotherapy*, vol. 14, no. 3, Aug. 2010, pp. 244–53. <https://doi.org/10.1080/10874208.2010.501517>.
- Klimesch, Wolfgang, et al. "EEG Alpha Oscillations: The Inhibition–timing Hypothesis." *Brain Research Reviews*, vol. 53, no. 1, Aug. 2006, pp. 63–88.
<https://doi.org/10.1016/j.brainresrev.2006.06.003>.
- Clark, C. Richard, et al. "Spontaneous Alpha Peak Frequency Predicts Working Memory Performance Across the Age Span." *International Journal of Psychophysiology*, vol. 53, no. 1, Feb. 2004, pp. 1–9. <https://doi.org/10.1016/j.ijpsycho.2003.12.011>.
- Bazanova, Olga. "Current Interpretation of Electroencephalogram Alpha Activity." *Neuroscience Letters*, vol. 500, July 2011, p. e3. <https://doi.org/10.1016/j.neulet.2011.05.068>.
- Hatlestad-Hall, Christoffer, et al. "BIDS-structured Resting-state Electroencephalography (EEG) Data Extracted From an Experimental Paradigm." *Data in Brief*, vol. 45, Oct. 2022, p. 108647. <https://doi.org/10.1016/j.dib.2022.108647>.
- Chen, He, et al. "EEG Characteristics of Children With Attention-deficit/Hyperactivity Disorder." *Neuroscience*, vol. 406, Mar. 2019, pp. 444–56.
<https://doi.org/10.1016/j.neuroscience.2019.03.048>.
- Foxe, John J., and Adam C. Snyder. "The Role of Alpha-Band Brain Oscillations as a Sensory Suppression Mechanism During Selective Attention." *Frontiers in Psychology*, vol. 2, Jan. 2011, <https://doi.org/10.3389/fpsyg.2011.00154>.
- Hanslmayr, Simon, et al. "Oscillatory Brain Mechanisms for Memory Formation." *Oxford University Press eBooks*, 2024, pp. 711–39.
<https://doi.org/10.1093/oxfordhb/9780190917982.013.25>.

- Hirshkowitz, Max, et al. "National Sleep Foundation's Sleep Time Duration Recommendations: Methodology and Results Summary." *Sleep Health*, vol. 1, no. 1, Mar. 2015, pp. 40–43. <https://doi.org/10.1016/j.sleh.2014.12.010>.
- Klimesch, Wolfgang. "Alpha-band Oscillations, Attention, and Controlled Access to Stored Information." *Trends in Cognitive Sciences*, vol. 16, no. 12, Nov. 2012, pp. 606–17. <https://doi.org/10.1016/j.tics.2012.10.007>.
- . "EEG Alpha and Theta Oscillations Reflect Cognitive and Memory Performance: A Review and Analysis." *Brain Research Reviews*, vol. 29, no. 2–3, Apr. 1999, pp. 169–95. [https://doi.org/10.1016/s0165-0173\(98\)00056-3](https://doi.org/10.1016/s0165-0173(98)00056-3).
- . "EEG Alpha and Theta Oscillations Reflect Cognitive and Memory Performance: A Review and Analysis." *Brain Research Reviews*, vol. 29, no. 2–3, Apr. 1999, pp. 169–95. [https://doi.org/10.1016/s0165-0173\(98\)00056-3](https://doi.org/10.1016/s0165-0173(98)00056-3).
- Pacheco, Luiza Bonfim, et al. "Disentangling Periodic and Aperiodic Resting EEG Correlates of Personality." *NeuroImage*, Apr. 2024, p. 120628. <https://doi.org/10.1016/j.neuroimage.2024.120628>.
- Zioga, Ioanna, et al. "The Role of Alpha Oscillations in Free- and Goal-directed Semantic Associations." *Human Brain Mapping*, vol. 45, no. 10, July 2024, <https://doi.org/10.1002/hbm.26770>.

Trastuzumab Deruxtecan vs. Immunotherapy: A Comparative Analysis of Survival Outcomes in HER2-Positive Breast Cancer By Dylan Gu

ABSTRACT

Breast cancer is the most common cancer worldwide, accounting for approximately 12.5% of all annual cases. In 2020, nearly 2.3 million new cases were reported, with 685,000 deaths. Over the past decade, breast cancer prevalence has risen by approximately 20%. Standard treatments include radiation, hormone therapy, chemotherapy, and targeted therapy, though these often come with significant side effects. Emerging as a promising alternative, immunotherapy leverages the immune system to fight cancer cells. This study compares trastuzumab deruxtecan (T-DXd), a potent treatment for HER2-positive breast cancer, with various immunotherapies. Our meta-analysis of clinical trial data focuses on survival outcomes in trials evaluating different immunotherapies compared to T-DXd. Through our independent analysis, we determined that T-DXd is significantly more effective than immunotherapy with an average progression-free survival (PFS) of 9.5 months versus 6.2 months for immunotherapy, and overall survival (OS) of 24.3 months compared to 19.6 months. This study concludes that while immunotherapy holds significant promise, T-DXd currently provides superior outcomes as a second-line treatment for HER2-positive breast cancer, suggesting potential future combinations of these therapies for better patient care.

Introduction

Breast cancer is a global problem: it is the most commonly diagnosed cancer in women, with an estimated 2.3 million new cases and over 685,000 deaths reported in 2020 (Sung et al., 2021). Although survival rates have steadily increased over the past two decades, breast cancer diagnoses continue to rise, resulting in the need for newer, more effective treatments (Nolan et al., 2023). Improved outcomes have been largely attributable to mammographic screenings (x-ray imaging) and adjuvant therapies - additional cancer treatments given after the primary treatment to minimize the risk of recurrence (Hashim et al., 2016). Breast cancer is predominantly a disease of aging women, with only 5 to 7% of patients diagnosed below the age of 40 years in the world (Brinton et al., 2008) and over 99% of diagnosed cases occur in women (WHO., 2024). Breast cancer arises when breast cells mutate and become cancerous, continuing to multiply and form tumors. A combination of genetic and non-genetic factors such as age, hormonal factors, radiation exposure, and lifestyle contribute to breast cancer incidence (Nolan et al., 2023). Breast cancer cells develop inside the milk ducts or the milk-producing lobules of the breast (Smolarz et al., 2022). Cells not caught quickly could spread rapidly to surrounding tissue, creating tumors or large lumps that further the severity of this disease (Waks, Winer., 2019). Breast cancer is characterized by its heterogeneity, encompassing various subtypes such as hormone receptor-positive, HER2-positive, and triple-negative breast cancers. People often discover they have breast cancer through symptoms like a lump in the breast, changes in breast shape or size, skin dimpling, nipple discharge, or pain. Regular screenings, such as

mammograms, are crucial for early detection, especially since some individuals may not experience noticeable symptoms in the early stages.

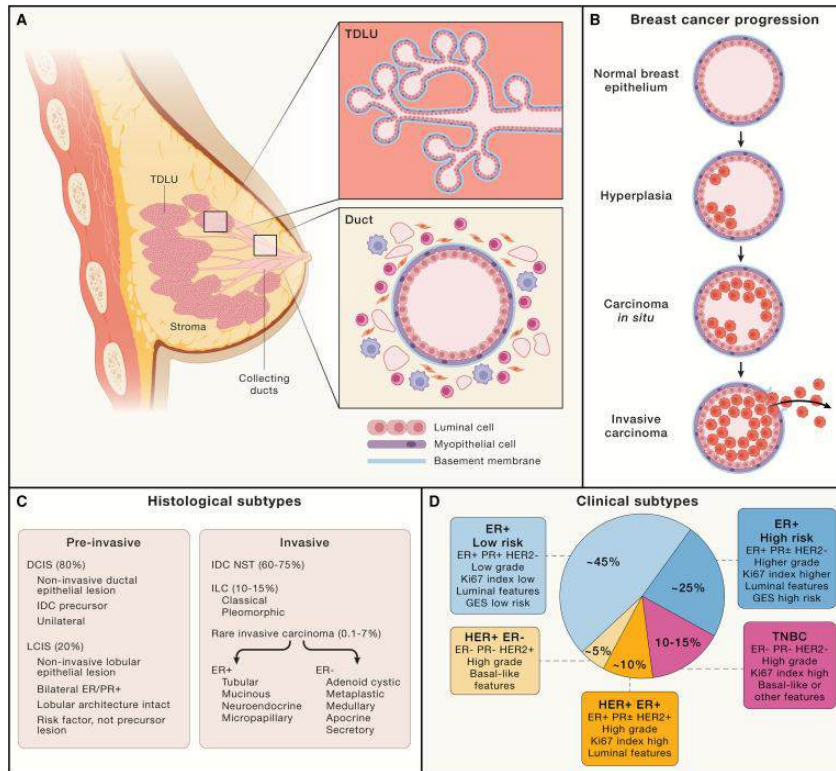


Figure 1: (Nolan et al., 2023) Diagram showing breast structure and various classifications of breast cancer. (A) Schematic of the human breast, depicting the terminal ductal lobular unit (TDLU), the area of the breast where the majority of tumors arise, and view of the epithelial duct tree. (B) Model of breast cancer pathogenesis. (C) Overview of major subtypes for pre-invasive and invasive breast cancer. (D) Overview of the different subtypes of breast cancer, with corresponding frequency percentages in humans.

HER2 Biology

Human epidermal growth factor receptor 2 (HER2) is a protein in the membrane of breast cancer cells that helps breast cancer cells grow quickly (Patel et al., 2020). It is a member of the epidermal growth factor receptor (EGFR) family, which includes four closely related receptors (HER1, HER2, HER3, and HER4) involved in cell signaling pathways that regulate cell proliferation, survival, and differentiation (Rubin & Yarden., 2001). In non-cancerous cells, HER2 helps regulate normal cellular processes by activating signaling pathways that promote healthy cell growth and repair (Yarden & Sliwkowski, 2001). Unlike other members of the EGFR family, HER2 does not have a known ligand that binds to it directly. Instead, it is activated by forming heterodimers with other HER family members, particularly HER3, which can amplify cell signaling and enhance cell growth and survival (Hynes & Lane, 2005). In normal settings, this controlled activation of HER2 is essential for maintaining proper cellular function

and tissue integrity. However, the development or overexpression of HER2 in cancerous settings can result from gene amplification, where multiple copies of the HER2 gene are produced, leading to excessive amounts of the HER2 protein on the cell surface (Slamon et al., 1987). This overexpression can cause continuous activation of growth and survival pathways, even in the absence of external growth signals, driving uncontrolled cell proliferation and contributing to tumor development.

In breast cancer, HER2 becomes problematic when the HER2 gene is amplified, leading to an overproduction of the HER2 protein on the surface of breast cancer cells. This overexpression occurs in approximately 15-20% of breast cancers and is associated with aggressive tumor behavior, increased rates of recurrence, and higher mortality (Hurvitz et al., 2022). The exact cause of HER2 gene amplification is not fully understood, but it is believed to involve complex genetic and epigenetic changes within the cancer cells. These changes may include mutations that affect gene regulation, chromosomal rearrangements, and interactions with other oncogenes or tumor suppressor genes (Moasser, 2007). HER2 overexpression drives cancer progression by continuously activating signaling pathways, such as the PI3K/AKT and MAPK pathways, which promote unchecked cell proliferation and survival, leading to tumor growth and metastasis (Patel et al., 2020).

Trastuzumab Deruxtecan

Trastuzumab deruxtecan (T-DXd) was first approved by the FDA in 2019 for the treatment of HER2-positive breast cancer. Since its approval, numerous studies have demonstrated the efficacy of T-DXd in extending patient survival and reducing the recurrence and progression of tumors. T-DXd is an antibody-drug conjugate (ADC) composed of a humanized anti-HER2 monoclonal antibody linked to a potent cytotoxic payload (Hurvitz et al., 2023). Upon binding to HER2 receptors on malignant cells, T-DXd is internalized, where lysosomal enzymes cleave the linker, releasing the DXd payload. This payload then induces targeted DNA damage, leading to apoptosis in cancer cells (Indini et al., 2021). The DESTINY-Breast01 phase 2 study revealed that T-DXd significantly improved overall survival compared to trastuzumab emtansine (T-DM1) in patients with HER2-positive metastatic breast cancer, achieving the longest reported median progression-free survival. One of the key advantages of T-DXd over T-DM1 is its higher drug-to-antibody ratio (DAR), which delivers more cytotoxic drug molecules per antibody, resulting in a more potent anti-tumor effect. Additionally, T-DXd's payload, DXd, functions as a topoisomerase I inhibitor, causing DNA breaks, whereas T-DM1's payload, emtansine (DM1), is a microtubule inhibitor. This difference in mechanism of action enables T-DXd to more effectively induce cell death, even in cancer cells that may have developed resistance to microtubule inhibitors like DM1 (Ogitani et al., 2016).

While targeted therapies like Trastuzumab Deruxtecan (T-DXd) and chemotherapy are key treatments for breast cancer, they come with significant limitations. One of the primary challenges with these therapies is the potential for drug resistance. Over time, cancer cells may adapt to the selective pressure these treatments apply, resulting in mutations or alterations in

signaling pathways that diminish the therapies' effectiveness. (Modi et al., 2022; O'Shaughnessy et al., 2021). This adaptive resistance is a major hurdle, particularly in patients with advanced or metastatic disease, where treatment options become increasingly limited as resistance develops (Pernas & Tolaney, 2021). In addition, chemotherapy, despite its widespread use and efficacy in certain contexts, also presents significant drawbacks. Chemotherapy agents are cytotoxic, indiscriminately killing rapidly dividing cells (Patel et al., 2022). This broad mechanism of action not only targets cancer cells but also harms healthy cells that naturally divide quickly, such as those in the hair follicles, gastrointestinal tract, and bone marrow, leading to side effects like hair loss, nausea, and increased infection risk (Kumar et al., 2018; Abu-Sbeih et al., 2019). The non-specific nature of chemotherapy also limits its long-term efficacy, causing collateral damage to normal tissues (Nabholtz et al., 2020).

Immunotherapy in Breast Cancer

While targeted therapies like Trastuzumab Deruxtecan (T-DXd) and chemotherapy are key treatments for breast cancer, they come with significant limitations. One of the primary challenges with these therapies is the potential for drug resistance. Over time, cancer cells may adapt to the selective pressure these treatments apply, resulting in mutations or alterations in signaling pathways that diminish the therapies' effectiveness. (Modi et al., 2022; O'Shaughnessy et al., 2021). This adaptive resistance is a major hurdle, particularly in patients with advanced or metastatic disease, where treatment options become increasingly limited as resistance develops (Pernas & Tolaney, 2021). In addition, chemotherapy, despite its widespread use and efficacy in certain contexts, also presents significant drawbacks. Chemotherapy agents are cytotoxic, indiscriminately killing rapidly dividing cells (Patel et al., 2022). This broad mechanism of action not only targets cancer cells but also harms healthy cells that naturally divide quickly, such as those in the hair follicles, gastrointestinal tract, and bone marrow, leading to side effects like hair loss, nausea, and increased infection risk (Kumar et al., 2018; Abu-Sbeih et al., 2019). The non-specific nature of chemotherapy also limits its long-term efficacy, causing collateral damage to normal tissues (Nabholtz et al., 2020).

Immunotherapy offers a novel approach to treating breast cancer by leveraging the body's immune system to fight the disease. Unlike targeted therapies and chemotherapy, which directly attack cancer cells or inhibit specific proteins, immunotherapy works by enhancing the immune system's ability to recognize and destroy cancer cells (Yu et al., 2017). This is primarily achieved by targeting immune checkpoints—proteins like PD-1, PD-L1, and CTLA-4—that normally function as regulators of immune responses, preventing autoimmunity but also enabling tumors to evade immune surveillance (Pardoll, 2012; Sharma & Allison, 2015). By targeting these proteins, signals that tumors rely on to evade the immune system are blocked, allowing T cells to recognize and destroy cancer cells more effectively (Brahmer et al., 2018). Currently, there are various modes of how immunotherapies are carried out.

Vaccines

Cancer vaccines stimulate the immune system to recognize and attack cancer cells by introducing specific tumor-associated antigens (Liu et al., 2019). These vaccines, such as the GLSI-100 vaccine designed to target HER2+ cancer cells, are therapeutics made to elicit an immune response against existing tumors, or preventive, aimed at preventing cancer recurrence. They typically involve the administration of peptide antigens derived from tumor proteins, which activate dendritic cells to present these antigens to T cells thus enhancing the body's immune response against the cancer cells (Van den Eynde et al., 2020).

CAR T-Cell Therapy

CAR T-cell therapy involves genetically engineering a patient's T cells to express chimeric antigen receptors (CARs) that specifically target tumor antigens. The process begins by collecting T cells from the patient's blood, which are then modified in the laboratory to express CARs that recognize specific proteins on the surface of breast cancer cells, such as HER2 (Yang et al., 2022). Once reintroduced into the patient's body, these engineered T cells can proliferate and attack the cancer cells, leading to tumor destruction (Bali et al., 2020). This approach effectively harnesses the specificity of the adaptive immune system to target tumors directly.

Immune Checkpoint Inhibitors

Immune checkpoint inhibitors (ICIs) enhance T cell function by blocking proteins that suppress immune responses. Tumors, including breast cancer, exploit immune checkpoints such as PD-1, PD-L1, and CTLA-4 to evade immune detection (Jalalvand et al., 2021). By inhibiting these checkpoints, checkpoint inhibitors restore T cell activity and boost the immune system's ability to attack cancer cells (Schmid et al., 2018). This process activates effector T cells, increases cytokine production, and ultimately, leads to the destruction of a tumor (Robert et al., 2015). Consequently, these therapies are often linked to improved survival rates and prolonged disease control.

Immunotherapies are quickly emerging as a promising approach for the treatment of HER2-positive breast cancer. These therapies harness the body's own immune system to fight cancer, but have not been tested with and against current approved treatments in both preclinical models and clinical trials. Therefore, the aim of this study is to assess how immunotherapies compare to trastuzumab deruxtecan, a prevalent and proven targeted therapy drug.

METHODS

Trial Selection and Meta-Analysis Criteria

To evaluate the efficacy of trastuzumab deruxtecan (T-DXd) versus immunotherapy in the second-line treatment of HER2-positive breast cancer, a meta-analysis of clinical trial data is conducted. The data is aggregated from a variety of sources, including phase 2 and phase 3 trials

comparing T-DXd and immunotherapy modalities such as ICIs and CAR T-cell therapies. We focused on two main clinical outcomes: progression-free survival (PFS), the time from the start of treatment until the cancer progresses, and overall survival (OS), the time from the start of treatment until death from any cause. We identified and selected relevant clinical trials by conducting a comprehensive literature review through databases such as PubMed, ClinicalTrials.gov, and Web of Science (These studies are shown below). Studies that did not directly compare these therapies or did not report PFS or OS outcomes were excluded. A total of eight trials met our inclusion criteria, providing sufficient data for meta-analysis.

Trial # / paper	Drug tested	Phase	Result	# of Patients
NCT03523585	Trastuzumab Deruxtecan	III	Shown favorable benefit-risk profile of trastuzumab deruxtecan in patients with HER2 positive metastatic breast cancer	406
NCT03529110	Trastuzumab Deruxtecan	III	Among patients with HER2-positive metastatic breast cancer previously treated with trastuzumab and a taxane, the risk of disease progression or death was lower among those who received trastuzumab deruxtecan	261
DEST INY-B reast03 study	Trastuzumab Deruxtecan	III	Trastuzumab deruxtecan showed a significant improvement in overall survival versus trastuzumab emtansine in patients with HER2-positive metastatic breast cancer, as well as the longest reported median progression-free survival	263
DEST INY-B reast 07 study	Trastuzumab Deruxtecan	III	T-DXd combinations show promise, particularly in expanding treatment options for patients with HER2-positive and HER2-low breast cancer	270
NCT03197389	HER2 neu peptide vaccine	I	Produced a median 9.2 month PFS and 26.6 month OS	5
NCT00194714	HER2 peptide vaccine	II	Combination therapy with trastuzumab and a HER2/neu vaccine is associated with minimal toxicity and results in prolonged, robust, antigen-specific immune responses in treated patients.	22
Phase 1B trial of Durva	Durvalumab	I	The RP2D of durvalumab and trastuzumab is standard full doses of both agents. No significant clinical activity was observed in patients with heavily pretreated	15

lumab			HER2-positive PD-L1-negative MBC.	
NCT02492711	Margetuximab plus Chemotherapy	III	Final overall OS analysis did not demonstrate margetuximab advantage over trastuzumab.	266

Statistical Analysis

Two-sample independent t-tests were conducted to compare the PFS and OS between T-DXd and immunotherapy. The p-value threshold for significance was set at 0.05. T-tests were chosen due to its ability to compare means from two independent groups (T-DXd and immunotherapy) to determine whether there is a statistically significant difference. For both endpoints, the mean PFS and OS were compared between T-DXd and immunotherapy across the selected trials. We calculated p-values for each comparison.

RESULTS

Progression-Free Survival (PFS)

The PFS for patients receiving T-DXd was significantly longer compared to those receiving immunotherapy. The t-test revealed a p-value of 0.0177, indicating that T-DXd was significantly more effective at preserving progression-free survival than immunotherapy. On average, patients on T-DXd had a PFS of 21.525 months compared to 9.72 months in patients treated with immunotherapy with a standard deviation of the difference of ±3.14.

Overall Survival (OS)

The OS results similarly showed that T-DXd outperformed immunotherapy. The t-test for OS yielded a p-value of 0.0420, indicating a statistically significant advantage for T-DXd in preserving overall survival. Patients treated with T-DXd had a median OS of 37.32 months, while those treated with immunotherapy had a median OS of 19.16 months, with a standard deviation of the difference of ±4.17.

Median PFS and OS of Patients

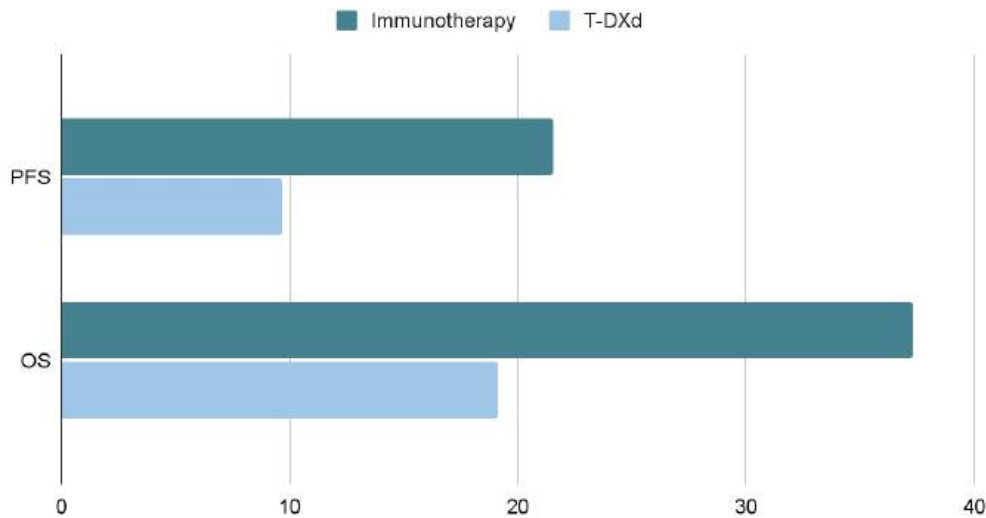


Figure 2: Bar chart of median progression free survival (PFS) and overall survival (OS) for immunotherapy and trastuzumab deruxtecan (T-DXd) patients. PFS means for immunotherapies and T-DXd were 21.525 and 9.72 months respectively. OS means for immunotherapies and T-DXd were 37.32 and 19.16 months respectively.

These results demonstrate that T-DXd offers superior outcomes in both PFS and OS compared to immunotherapy in the second-line treatment of HER2-positive breast cancer. This is possible because immunotherapies are still new, with many still not being approved for public use. As a result, available data for immunotherapies may be skewed by patient demographics for the purpose of the trial (early phase).

DISCUSSION

The results of our meta-analysis indicate that trastuzumab deruxtecan (T-DXd) is significantly more effective than immunotherapy in extending both PFS and OS in patients with HER2-positive breast cancer. We hypothesize this not because immunotherapies are inherently worse than T-DXd, but because immunotherapies are still undergoing testing and are being adjusted daily. These findings suggest that while immunotherapy holds promise, particularly for cancers that are resistant to traditional therapies, it may not yet be ready to supplant T-DXd as a second-line treatment option for HER2-positive breast cancer. However, immunotherapy still offers potential advantages, particularly in cases where drug resistance to T-DXd or other targeted therapies develops. Immunotherapies, such as ICIs, enhance the immune system's ability to recognize and attack cancer cells. This distinct mechanism of action could complement T-DXd or be used sequentially in patients who become resistant to HER2-targeted therapies.

Limitations

While our analysis suggests that T-DXd is more effective today, the long-term effects of combining immunotherapy with targeted therapies like T-DXd have not been fully explored. Many immunotherapies are still currently ongoing clinical trials and have not been released for public use. In addition, sample size, demographics, trial phase, and structure all may have impacted the accuracy of this study. Due to the shortage of publicly available resources for immunotherapy data, many studies used immunotherapies in combination with other treatments, which may skew the full potential of immunotherapies as a treatment. Future research may reveal synergistic benefits between these modalities. Additionally, the heterogeneity of immunotherapy responses across patients with HER2-positive breast cancer underscores the need for personalized treatment approaches.

Immunotherapies in The Future

The discovery and implementation of new HER2-directed therapies in clinical practice has significantly changed how patients with HER2+ diseases are being treated. Although major strides have been made in treating patients with HER2+ diseases, ongoing studies aim to continue improving outcomes for this subset of patients. Whether it is a combination of multiple HER2-directed therapies in various settings or the invention of immunotherapy, the treatment of HER2+ disease has resulted in better outcomes. Progression-free survival and overall survival, compared to the outcomes of previous decades, have been rising and will continue to evolve. For a prevalent disease like breast cancer, finding new and effective treatment options is crucial and necessary for the survival and general health of all cancer patients.

Works Cited

- Adrienne G. Waks, MD. "Breast Cancer Treatment in 2019." *JAMA*, JAMA Network, 22 Jan. 2019, jamanetwork.com/journals/jama/article-abstract/2721183.
- Agostinetto, Elisa, et al. "Immunotherapy for HER2-Positive Breast Cancer: Clinical Evidence and Future Perspectives." *MDPI*, Multidisciplinary Digital Publishing Institute, 25 Apr. 2022, www.mdpi.com/2072-6694/14/9/2136.
- Bartsch R Berghoff et al. "Trastuzumab Deruxtecan in HER2-Positive Breast Cancer with Brain Metastases: A Single-Arm, Phase 2 Trial." *Nature Medicine*, U.S. National Library of Medicine, pubmed.ncbi.nlm.nih.gov/35941372/. Accessed 28 Sept. 2024.
- Gaynor, Nicola, et al. "Immune Checkpoint Inhibitors: Key Trials and an Emerging Role in Breast Cancer." *Seminars in Cancer Biology*, Academic Press, 2 July 2020, www.sciencedirect.com/science/article/pii/S1044579X20301528.
- Hurvitz, Sara A, et al. "Trastuzumab Deruxtecan versus Trastuzumab Emtansine in Patients with HER2-Positive Metastatic Breast Cancer: Updated Results from Destiny-Breast03, a Randomised, Open-Label, Phase 3 Trial - the Lancet." *TheLancet.Net*, 14 Jan. 2023, [www.thelancet.com/journals/lancet/article/PIIS0140-6736\(22\)02420-5/fulltext](http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(22)02420-5/fulltext).
- Keytruda® (PEMBROLIZUMAB) plus Trastuzumab and Chemotherapy Significantly Improved Progression-Free Survival (PFS) versus Trastuzumab and Chemotherapy in First-Line HER2-Positive Advanced Gastric or Gastroesophageal Junction (GEJ) Adenocarcinoma." *Merck.Com*, 28 Sept. 2024, www.merck.com/news/keytruda-pembrolizumab-plus-trastuzumab-and-chemotherapy-significantly-improved-progression-free-survival-pfs-versus-trastuzumab-and-chemotherapy-in-first-line-her2-positive-advanced-gastric/.
- Masoud, Viviana, and Gilles Pagès. "Targeted Therapies in Breast Cancer: New Challenges to Fight against Resistance." *World Journal of Clinical Oncology*, U.S. National Library of Medicine, 10 Apr. 2017, www.ncbi.nlm.nih.gov/pmc/articles/PMC5385433/.
- Modi, Shanu, and Cristina Saura. "ENHERTU® (FAM-Trastuzumab Deruxtecan-Nxki): Official Patient Website." *ENHERTU® (Fam-Trastuzumab Deruxtecan-Nxki) Patient Website*, 11 Dec. 2019, www.enhertu.com/.
- Nolan, Emma, and Geoffrey J Lindeman. "Deciphering Breast Cancer: From Biology to the Clinic." *Science Direct*, 16 Mar. 2023, www.sciencedirect.com/science/article/pii/S0092867423000995.
- Patel, Aena, et al. "The Changing Paradigm for the Treatment of HER2-Positive Breast Cancer." *Cancers*, U.S. National Library of Medicine, 28 July 2020, www.ncbi.nlm.nih.gov/pmc/articles/PMC7464074/.
- Perez, Jose, et al. "Trastuzumab Deruxtecan in Her2-Positive Metastatic Breast Cancer Patients with Brain Metastases: A Destiny-Breast01 Subgroup Analysis." *Cancer Discovery*, U.S. National Library of Medicine, 24 Nov. 2020, pubmed.ncbi.nlm.nih.gov/36255231/.

- Polyak, Kornelia. "Breast Cancer: Origins and Evolution." *The Journal of Clinical Investigation*, American Society for Clinical Investigation, 1 Nov. 2007, www.jci.org/articles/view/33295.
- Radosa, Julia Caroline, et al. "Clinical Data on Immunotherapy in Breast Cancer." *Karger Publishers*, S. Karger AG, 13 Oct. 2020, karger.com/brc/article/15/5/450/53694/Clinical-Data-on-Immunotherapy-in-Breast-Cancer.
- Rubin, I., and Y. Yarden. "The Basic Biology of HER2." *Annals of Oncology*, Elsevier, 6 Jan. 2020, www.sciencedirect.com/science/article/pii/S0923753419544556.
- Slamon DJ et al. "Use of Chemotherapy plus a Monoclonal Antibody against HER2 for Metastatic Breast Cancer That Overexpresses HER2." *The New England Journal of Medicine*, U.S. National Library of Medicine, pubmed.ncbi.nlm.nih.gov/11248153/. Accessed 28 Sept. 2024.
- Smolarz, Beata, et al. "Breast Cancer-Epidemiology, Classification, Pathogenesis and Treatment (Review of Literature)." *MDPI*, Multidisciplinary Digital Publishing Institute, 23 May 2022, www.mdpi.com/2072-6694/14/10/2569.
- Winer Eric P et al. "Nivolumab for Previously Treated Unresectable Metastatic Anal Cancer (NCI9673): A Multicentre, Single-Arm, Phase 2 Study - the Lancet Oncology." *The Lancet.Net*, Apr. 2021, [www.thelancet.com/journals/lanonc/article/PIIS1470-2045\(17\)30104-3/abstract](http://www.thelancet.com/journals/lanonc/article/PIIS1470-2045(17)30104-3/abstract).
- Yarden, Yosef. "Biology of HER2 and Its Importance in Breast Cancer." *Karger Publishers*, S. Karger AG, 26 Oct. 2001, karger.com/ocl/article-abstract/61/Suppl.%202/1/237361/Biology-of-HER2-and-Its-Importance-in-Breast?redirectedFrom=fulltext.

The Effect of Non-governmental Funding on School Dropout Rates in Cambodia

By Jason Tun

Abstract

The educational system in Cambodia faced significant challenges, including high dropout rates, following the oppressive rule of the Khmer Rouge. To support its students and the rebuilding of schools, the Cambodian government has increasingly relied on international funding sources and various non-governmental organizations; this study examines the effect of these funding sources on public school dropout rates. Analyzing data from Cambodia's Ministry of Education, Youth and Sport (MoYES), this article compares the effect of non-governmental funding sources on school dropout rate in primary, lower secondary, and upper secondary grades in all 25 of the country's provinces. Regression analyses showed that only primary schools (grades 1-6) benefited from non-governmental funding, with an overall reduction in dropout rate. These results suggest that allocation of funding should prioritize higher-level grades, in addition to being more transparent and directed towards programs that have demonstrated impact at all grade levels.

Background

Primary and secondary school dropout rates are a significant problem in many developing countries. Many social, environmental, and economic factors may influence dropout rates. Student-related factors such as low attendance and academic achievement, family-related factors such as high poverty rates and parent education, and school-related factors such as teacher absenteeism all have been shown to contribute to school dropout (Hirakawa, 2017; Goksen, 2010; Hanushek, 2006). While many studies examine these factors in developing countries, there are fewer studies specifically addressing student dropout rates in Cambodia's education system.

Cambodia's public schools are generally divided into three distinct educational levels: primary schools (grades 1-6), lower secondary schools (grades 7-9), and upper secondary schools (grades 10-12). Enrollment at all levels has increased drastically over the last decade, especially in primary schools. However, the Cambodian government has largely underfunded the public school system, only spending about 2.6% of its GDP on education between 2006 and 2012, notably lower than the East Asian average of 3.8% and the global average of 4.8% (Brehm, 2016). This has led to a surge in the founding of private schools and a notably high number of non-governmental organizations (NGOs) and international organizations (IOs) operating in the country. For example, in 2008, there were as many as 1,500 registered NGOs and 340 IOs active in Cambodia (Chanboreth & Hach, 2008). Many of these organizations support the education sector, specifically the primary schools, aiming to increase low enrollment rates and decrease the high dropout rates by providing teacher training, school resource donations, and direct funding (Brehm, 2016).

Cambodia's tumultuous past partly accounts for its struggling education system. From 1975 to 1979, the monarchical and repressive Khmer Rouge regime destroyed much of

Cambodia's professional and cultural societies — about a quarter of the population died during the Khmer Rouge reign, including 80% of teachers and 95% of doctors, effectively halting formal education through the oppression of educated and skilled citizens (Suárez & Marshall, 2012). The Khmer Rouge viewed education as a barrier to their primary goal: agricultural expansion (Clayton, 1998). Since 90% of schools were destroyed or re-purposed for Communist activities, many children became illiterate or remained uneducated past basic literacy (Clayton, 1998). Due to the Khmer Rouge's oppression of education, Cambodia education system fell behind, and, even today, the government lacks resources and motivation to develop and catch up, leading to NGOs and international support stepping in.

Recently, the Cambodian government has implemented initiatives to nationally reform education in the years since the Khmer Rouge. The First Education Reform Cycle was introduced in 2000 to invest in educational infrastructure and teacher training in secondary schools; this process was initially successful, with a 35% increase in enrollment, but it eventually failed due to a lack of sustained attention on the project. The Ministry of Education, Youth and Sport (MoEYS) has also started to create charter schools, or New Generation Schools, that receive high investment and increased quality standards (Bredenberg & Kampuchean Action to Promote Education, 2019).

Although the Cambodian government has increased efforts to reform the national education system, NGOs still carry much of the weight in supporting enrollment and curriculum (Chanboreth & Hach, 2008), and dropout rates in Cambodian public schools remain high: 7.2% in grades 1-6, 16.6% in grades 7-9, and 18.5% in grades 10-12 as reported by the Cambodian MoEYS in 2022 (Ministry of Education, Youth, and Sport, 2022). This means that most children are not receiving an education adequate enough to advance the country's development and create meaningful occupations. Many challenges to maintaining school enrollment persist as families continue to rely on their children to work in the year-round agricultural economy. Because of limited government resources, NGOs and IOs maintain a role in providing support through teacher and staff development, investments into school resources, and class curriculums (Escamilla, 2010). This research aims to quantify the effect of NGOs on education outcomes in Cambodia by determining whether NGO/IO and funding from abroad affects school dropout rates in Cambodia. Specifically, this study asks: Does non-governmental funding of schools in Cambodia affect student dropout rates?

Methodology

This project utilizes data from the 2021-2022 governmental report on Education Statistics and Indicators, published by the Cambodian Ministry of Education, Youth, and Sport (MoEYS) (Ministry of Education, Youth, and Sport, 2022). The report provides information on public preschool, primary, and secondary schools in Cambodia, including the number of students and staff, source of funding, and educational outcome measures. The report was publicly available as an open-source dataset published on the webpage "Open Development Cambodia (ODC)" (Ministry of Education, Youth, and Sport, 2022).

To examine the role of funding on educational outcomes in Cambodia, three independent variables were chosen to represent school funding: (1) the amount of funding per student for each province (in Riels, Cambodia's national currency), (2) the number of schools per province obtaining financial support from IOs/NGOs, and (3) the number of schools per province obtaining financial support from other sources abroad, including other nations and foreign donations. The dependent variable was each province's dropout rate, as calculated and reported by MoEYS using the formula in the methods section of the report. Specifically, the dropout rate was defined as the proportion of students remaining in each level of school after subtracting the promotion rate (proportion of students who advanced to the next grade in the following year) and the repetition rate (proportion of students who stayed in the same grade the following year). The dropout rate was calculated after combining the promotion and repetition rates and subtracting them from 100%.

Each variable was analyzed using data from each province ($N = 25$), and analyses were conducted separately for each level of schooling in the Cambodian educational system: primary (grades 1-6), lower secondary (grades 7-9), and upper secondary (grades 10-12) levels. The average dropout rates by level are 8.3 in primary schools, 16.6 in lower secondary schools, and 17.6 in upper secondary schools. Each educational level was examined separately since the dropout rates increased substantially with higher level grades, and each level may be affected differently by the influences of non-governmental support.

An outlier in the province of Koh Kong—whose reported figure for per-student funding was over five times higher than the next highest per-student funding amount in another province—was removed from the analysis to maintain consistency in the amounts of funding per province; it is unclear why this province reported such a high per-student funding rate. Simple linear regression was used to analyze the dropout rate of each education level for each independent variable. The results of multiple regression analyses showed no significance between the variables.

Results

Descriptive statistics for the dependent and independent variables are reported in Tables 1-3. Dropout rates for the three education levels showed a higher dropout rate in the Lower and Upper Secondary grade levels (16.6% and 17.6%, respectively), as compared to the Primary grades (8.3%) (Table 1). These statistics may reflect the difficulties Cambodian schools and families have in prioritizing upper-level education.

Descriptive statistics for the independent variables showed considerable variation among provinces. For example, funding per student ranged from 35,642 riels in Phnom Penh province to 102,781 riels in Kampong Cham province. This variation exists despite excluding the province of Koh Kong, and can be expected, as some provinces are more rural or have fewer students or resources compared to more populated areas, which may receive a greater proportion of funding per student. Likewise, there was considerable variation among the provinces in the number of schools receiving non-governmental funding, ranging from 10 schools in Pailin province to 534

schools in Prey Veng province. The number of total schools in each province could affect the data, as rural provinces have fewer schools than urban provinces, and thus, fewer total schools in that province may receive external funding.

	% Dropout Rate		
	Primary (Grades 1-6)	Lower Secondary (Grades 7-9)	Upper Secondary (Grades 10-12)
AVG	8.3	16.6	17.6
Max	12.2	27.7	30.8
Min	3.6	10.3	9.5
Median	8.3	16.7	16.9
SD	2.5	4.1	5.0
Count	25.0	25.0	25.0

Table 1. Dropout Rate Statistics for Primary, Lower Secondary, and Upper Secondary Schools for 25 provinces in Cambodia 2021-2022. Cumulative dropout rate descriptive statistics for all 25 provinces demonstrate a higher dropout rate in secondary grade levels compared to primary grades, with grades 10-12 having the highest proportion of students dropping out.

Funding per Student (Riel)	
AVG	56,826
Max	102,781
Min	35642
Median	56805.0
SD	13877.3
Count	24

Table 2. Funding per Student (Riel) for 24 provinces in Cambodia in 2021-2022 (excluding Koh Kong). Descriptive statistics for per-student funding shows that the range (max - min) for funding per student is extensive, and the standard deviation is moderately high, indicating considerable variability in per-student funding among provinces.

Number of schools receiving
Funding from Sources Abroad+
IO/NGOs

AVG	235
Max	534
Min	10
Median	243
SD	151
Count	25

Table 3. Number of schools receiving Funding from sources Abroad and International organizations (IOs)/non-governmental organizations (NGOs) for all 25 provinces in Cambodia in 2021-2022. Descriptive statistics on the number of schools receiving non-governmental funding from IOs and NGOs show a very high range (max-min) and standard deviation, meaning there is considerable variability in non-governmental funding among the 25 provinces.

Generally, the analyses show no consistent impact of funding on student dropout rates in Cambodian schools. While there were some fairly weak overall relationships between funding measures and dropout rates, the effect differed depending on the grade levels examined.

Variables	Estimate	Standard Error	t-statistic	p-value	Adjusted R ²	Multiple R ²	F-statistic
Dropout 1 to 6, Funding per Student	7.01E-05	3.52E-05	1.993	0.0588	0.1145	0.153	3.974
Dropout 7 to 9, Funding per Student	3.45E-05	6.34E-05	0.545	0.59128	-0.03153	0.01332	0.297
Dropout 10 to 12, Funding per Student	-7.61E-05	7.69E-05	-0.99	0.333	-0.0008434	0.04267	0.9806

Table 4. Estimates of the regression model between funding per student (in Riels) and dropout rate for each grade level (primary grades 1-6, lower secondary grades 7-9, and upper secondary grades 10-12). There were no statistically significant associations between per-student funding and dropout rates at any grade level.

Variables	Estimate	Standard Error	t-statistic	p-value	Adjusted R ²	Multiple R ²	F-statistic
Dropout 1 to 6, Total							
NGO/IO + Abroad	-0.005274	0.003247	-1.624	0.118	0.06388	0.1029	2.638
Dropout 7 to 9, Total							
NGO/IO + Abroad	0.009634	0.005258	1.832	0.0799	0.08943	0.1274	3.357
Dropout 10 to 12, Total							
NGO/IO + Abroad	0.018479	0.005739	3.22	0.00379*	0.2807	0.3107	10.37

Table 5. Estimates of the regression model between Total NGO/IO + Funding Sources Abroad and dropout rate for each grade level (primary grades 1-6, lower secondary grades 7-9, and upper secondary grades 10-12). The only significant association was in grades 10-12, where an increase in international funding related to increased dropout rates. *statistically significant

Variables	Estimate	Standard Error	t-statistic	p-value	Adjusted R ²	Multiple R ²	F-statistic
Dropout 1 to 6, NGO/IO	-0.05552	0.02479	-2.239	0.0351*	0.1433	0.179	5.015
Dropout 7 to 9, NGO/IO	0.06784	0.04264	1.591	0.125	0.05999	0.09916	2.532
Dropout 10 to 12, NGO/IO	0.13318	0.04767	2.794	0.0103*	0.2209	0.2533	7.804

Table 6. Estimates of the regression model between the number of schools receiving NGO + IO funding and dropout rate for each grade level (primary grades 1-6, lower secondary grades 7-9, and upper secondary grades 10-12). In grades 1-6 an increase in NGO+IO funding was related to decreased dropout rates. However in grades 10-12, the association reversed, with an increase in NGO+IO funding related to higher dropout rates *statistically significant

Variables	Estimate	Standard Error	t-statistic	p-value	Adjusted R ²	Multiple R ²	F-statistic
Dropout 1 to 6, Abroad	-0.005168	0.003555	-1.454	0.16	0.04432	0.08414	2.113
Dropout 7 to 9, Abroad	0.010059	0.005727	1.757	0.0923	0.07995	0.1183	3.085
Dropout 10 to 12, Abroad	0.019237	0.006325	3.042	0.0058*	0.2558	0.2868	9.251

Table 7. Estimates of the regression model between the number of schools receiving support from sources Abroad and the dropout rate for each grade level (primary grades 1-6, lower secondary grades 7-9, and upper secondary grades 10-12). In grades 10-12, an increase in funding sources Abroad related to higher dropout rates *statistically significant

When examining the relationship between per-student funding and dropout rates, there is no significant relationship in any of the grade levels examined (Table 4). Table 5 shows the relationship between the dropout rate and the total number of schools receiving NGO/IO funding plus all funding sources abroad (including other nations and foreign donations) combined. Looking at the combination of these funding sources, there is an unexpected association in grades 10-12 of increased funding with a *higher* dropout rate. Because of this finding, the effect of NGO/IO funding and abroad funding are examined separately.

When only considering the number of schools receiving NGO/IO funding, there is a weak but statistically significant relationship between the number of schools receiving IO/NGO funding and lower dropout rates, for primary grade levels 1-6 (Table 6). For these grade levels, for an additional IO/NGO working in a province, the dropout rate was reduced by 0.05 ($p < .05$). Funding from IO/NGOs explains 17.9% of the explained variation of dropout rates. This regression was statistically significant compared to a null model ($F = 2.638$, $df1 = 1$, $df2 = 23$, $p < .05$).

For the other grade levels, there was no significant relationship (grades 7-9), and in the case of upper secondary grades 10-12, there was a negative relationship, where *increased* IO/NGO funding related to *higher* dropout rates (Table 6). For grades 7-9, we found that for each additional unit increase in the predictor variable, the dropout rate increased by 0.068 ($p = 0.125$). However, the number of NGOs only explains only 5.999% of the variation in dropout rates, and the regression was not statistically significant compared to a null model ($F = 2.532$, $df1 = 1$, $df2 = 23$, $p = 0.125$), suggesting that the number of NGOs does not have a meaningful impact on dropout rates and only explains a very small portion of the variance in dropout rates for grades 7-9. For grades 10-12, we found that when the number of NGOs in a province increased by one, the dropout rate also increased by 0.13 ($p < .05$). The number of NGOs explains 22.09% of the

dropout rate variance. This regression model was statistically significant compared to a null model ($F = 7.804$, $df1 = 1$, $df2 = 23$, $p < .05$), indicating that the number of schools receiving IO/NGO support is significantly related to with an *increase* of dropout rates in grades 10-12.

Finally, when considering the number of schools receiving funding from abroad sources, there was a negative relationship between increased abroad funding and *higher* dropout rates in grades 10-12 (Table 7). For grades 1-6, we observed that for each increase in the number of organizations abroad, the dropout rate decreased by 0.0052 ($p = 0.16$). However, this model was not statistically significant compared to a null model ($F = 2.113$, $df1 = 1$, $df2 = 23$, $p = 0.16$), and it explains only a small portion of the variation in dropout rates. For grades 7-9, we found that for each additional abroad organization increase, the dropout rate increased by 0.0101 ($p = 0.0923$). The predictor variable only explains 11.83% of the variation in dropout rates. Although the model approaches statistical significance, it is not significant at conventional levels ($F = 3.085$, $df1 = 1$, $df2 = 23$, $p = 0.0923$). Finally, for upper secondary education grades 10-12, we observed that for each additional increase in the number of abroad support groups, the dropout rate *increased* by 0.0192 ($p = 0.0058$). The independent variable explains 28.68% of the variation in dropout rates, and this regression model was statistically significant compared to a null model ($F = 9.251$, $df1 = 1$, $df2 = 23$, $p < 0.01$), indicating a meaningful impact of the predictor (funding from abroad) on the dropout rate.

Discussion

Although the public school system in Cambodia is heavily supported by non-governmental funding sources, it remains unclear whether this financial support has benefited the students in Cambodian schools. This study found that the effect of these funding sources, as well as per-student funding, is fairly limited when it comes to student dropout rates. Unexpectedly, the only significant positive relationship found in this study was in Primary levels (grades 1-6), where lower dropout rates were associated with a higher number of schools receiving IO/NGO funding. Not only did the other grade levels not benefit from any funding source or increased per-student funding, the Upper Secondary level (grades 10-12) seemed to be negatively affected by increased non-governmental support.

The factors explaining the lack of benefit for Secondary level grades are likely multifactorial and may vary depending on urban versus rural environments, where students may have differing commute times and family responsibilities outside of school. Pov et al. (2020) examined the causes of lower secondary school dropout rates in Cambodia, and found that student-related factors of grade retention, absenteeism, academic expectations, and private tutoring significantly influenced dropout, whereas family- and school-related factors had no effect on dropout (Pov, 2020). This study, in combination with these results, suggests that funding should be directed at improving attendance and academic achievement goals in Secondary grade levels to reduce student dropout rates. As per-student funding did not have an effect at any grade level, there can be less emphasis on the quantity of funded initiatives, as opposed to the quality of programs.

The government of Cambodia has made recent efforts to improve its educational system, with the introduction of three priority education policies for 2006–2010: ensure equitable access, improve quality and efficiency of educational services, and increase capacity (Tan, 2007). Tan also suggests that social and cultural factors such as teachers charging informal fees to students and poor transparency and accountability may be barriers to improving the educational outcomes (Tan, 2007). Certainly, Cambodia can benefit from any increase in funding to support its school system, but reforming the funding structure to become more transparent should also accompany financial support. NGO/IOs and other funding sources from abroad should also aim to be more transparent in how their funds are spent and the effect of their programs on student outcomes. Finally, the Cambodian government should prioritize Secondary education and analyze and learn how the Primary levels were able to benefit from NGO/IO funding. Although shifting some of the primary school focus onto higher education levels may be difficult as students grow older and have more responsibilities outside of school, it is necessary to further develop Cambodia's education system.

This research did have significant limitations. Finding data about educational outcomes directly from NGOs was difficult, as they often were reluctant to share collected data on specific projects, possibly to protect themselves from the government or obscure results. The Cambodian government should partner with these non-governmental funding sources to better capture and record the effects of NGO/IOs to achieve their main educational goals, instead of working separately or in parallel with these organizations, to try and sustain educational growth. Another limitation of the study is that the data may not be an accurate reflection of NGO/IO's funding because the database reported an aggregated number of NGOs per province and the total funding amount in that district. Provincial factors such as population may change the context of these aggregate numbers. Finally, the quality of the data was also not validated, as data collection methods were not described or verified. Therefore, the results of this paper should be interpreted with caution. More research is needed to determine the effects of non-governmental funding on student outcomes in Cambodia, so that NGO/IOs can collaborate more effectively with the Cambodian government to support all levels of education.

Works Cited

- Bredenberg, K. & Kampuchean Action to Promote Education. (2019). Secondary education in Cambodia: the progress of reform. *Education in Cambodia: From Year Zero to International Standards*. Retrieved July 9, 2024, from https://www.kapekh.org/files/report_file/106-en.pdf
- Brehm, W. (2016). Hybrid Spaces of the “Public” and “Private” [PDF]. In *The Contemporary Landscape of Education in Cambodia* (pp. 271–282). Routledge. <https://doi.org/10.4324/9781315736709>
- Chanboreth, E., & Hach, S. (2008). Aid effectiveness in Cambodia. In *Brookings*. Wolfensohn Center for Development, Brookings Institution. <https://www.brookings.edu/articles/aid-effectiveness-in-cambodia/>
- Clayton, T. (1998). Building the New Cambodia: Educational Destruction and Construction under the Khmer Rouge, 1975-1979. *History of Education Quarterly*, 38(1), 1–16. <https://doi.org/10.2307/369662>
- Escamilla, M. (2010). *NGO-sponsored education and the development of the local civil society sector in Cambodia* [MA Thesis, University of Ottawa]. <https://doi.org/10.20381/ruor-12581>
- Goksen, F. & Cemalcilar, Z. (2010). Social capital and cultural distance as predictors of early school dropout: Implications for community action for Turkish internal migrants. *International Journal of Intercultural Relations*, 34(2), 163–175.
- Hanuskek, E.A. (2006). Do students care about school quality? Determinants of dropout behavior in developing countries. *Journal of Human Capital*, 1(2), 143–151.
- Hirakawa, Y., & Taniguchi, K. (2020). School dropout in primary schools in rural Cambodia: School-level and student-level factors. *Asia Pacific Journal of Education*, 41(3), 527–542. <https://doi.org/10.1080/02188791.2020.1832042>
- Hirakawa, Y. (2017). The causes of school dropout in rural primary schools in Malawi: *Journal of International Cooperation in Education*, 20(1), 1-15.
- Ministry of Education, Youth, and Sport. (2022). *Education Statistics and Indicators, 2021-2022*. Ministry of Education, Youth and Sport. Retrieved August 20, 2024, from https://data.opendevdevelopmentcambodia.net/en/dataset/dd651620-b208-4117-b9c9-d54df047e985/resource/062610f5-45c1-4158-a8cf-503ce874224b/download/public-education-statistics-and-indicators-2021-2022_en.pdf
- Pov, S., Kawai, N., & Murakami, R. (2020). Identifying causes of lower secondary school dropout in Cambodia: a two-level hierarchical linear model. *International Journal of Inclusive Education*, 26(8), 834–847. <https://doi.org/10.1080/13603116.2020.1735542>
- Tan, C. (2007). Education reforms in Cambodia: issues and concerns. *Educational Research for Policy and Practice*, 6(1), 15–24. <https://doi.org/10.1007/s10671-007-9020-3>

Demographic Influences on U.S. Election Outcomes Across Levels: A Machine Learning Approach By Benjamin Qiao

Abstract

Demographic factors such as race and education are strongly correlated with party affiliation in the United States. Using machine learning techniques, this paper uncovers how the relationship between those demographic factors and voting behavior differs *across different election jurisdictions* by analyzing those relationships across senatorial, gubernatorial, and presidential elections. Previous studies surveyed do not incorporate modern machine-learning techniques; incorporating this into the model allows for a numerical scoring of each demographic in its importance of determining election results. The result was a separate importance score per demographic factor for each election jurisdiction analyzed. Among other results, the analysis in this paper showed that racial demographics tend to have the strongest predictive power among all demographic features studied, but have greater importance in presidential elections than in gubernatorial or senatorial elections. This study found that despite large shifts in party platforms and candidate profiles, demographic patterns still hold strong predictive power in how voters will choose candidates in elections across different levels of government.

Introduction

Voters behave differently across various types of elections for different reasons. A governor cannot dictate U.S. foreign policy, and the president cannot manage state infrastructure projects. What people look for on the local level is different from what people look for on the national level; this paper seeks to quantify what that difference is.

Two things can be easily observed with data: the demographics of a county and a county's voting outcomes. In this research, a model was developed using socioeconomic demographics to predict election results from 2010 to 2020, and through analysis, the weight of each demographic parameter in its influence on election results across election levels was identified.

This study's findings confirm that different demographic factors do indeed significantly affect voting patterns depending on the level of office. This study also provides insights into the nuanced ways in which voter behavior and preferences vary by the type of election, contributing to a deeper understanding of electoral dynamics and informing more targeted political strategies.

Literature Review

The influence of demographics on political party affiliation and subsequently voting behavior is well established. In their study identifying how socio-demographic factors determine political party affiliation, Tanuzi and Franklin noted the interaction of specific demographic patterns that align with predicting party affiliation. For example, this study found that for Black Americans there is a 15.5% higher likelihood that the respondent is affiliated with the Democratic Party on a 0.01 significance level, even after controlling for gender, age, education

level, political views, children, income, family income, father's education, and marital status. This study found similar strong relationships between other demographic factors (beyond race) and party affiliation as well.

The trends of partisan affiliation among demographic groups have not changed substantially since 1994 according to a 2018 Pew Research Center study. Among the demographic groups studied,⁷⁸ there were only slight changes in partisan identification for the group, with 37% of women identifying with the Democratic Party in 1994 compared with 39% in 2017, for instance (Pew Research Center). These findings suggest that the *relationship between specific socio-demographic groups and party affiliation* has endured in the period studied.

There have been past studies confirming the existence of a difference in voting behavior depending on election level. Kurtbaş found that voters place different priorities on local elections compared to general ones; the study found that nearly half of the surveyed voters (48.4%) conducted little to no research before voting in the local elections. This research also found that older voters and those with lower education and income levels are more focused on local elections and that younger, more educated, and wealthier voters tend to prioritize general elections. These differences in voting behavior based on demographic differences highlight key areas of analysis later in this paper.

Researchers have previously identified some specific differences in how voters hold incumbent parties and politicians accountable for conditions while they are in office, although the differences identified differ. Atkeson and Partin concluded that governors and senators undergo different types of retrospective evaluations by the constituents, finding that governors are held more accountable for state economic conditions than senators, who are linked to the actions of the presidency. In contrast, an earlier 1990 study determined practically the opposite, concluding that voters hold their governor “neither responsible nor accountable” for their state's economic conditions, placing more responsibility on congressional and presidential incumbents (Stein).

Past research surveyed does not fully incorporate modern machine learning techniques in analyzing the differences in voting outcomes *across* election jurisdictions—this research aims to fill that gap. Machine learning allows the quantitative assessment of factors in contrast with previous methods, which have identified relationships but did not provide a numerical ranking of their relative importance. It also allows for large-scale analysis across multiple elections and jurisdictions, taking advantage of existing data sets instead of having to conduct labor-intensive data collection. Finally, machine learning models can handle complex, nonlinear relationships between variables more effectively than traditional statistical methods like linear regression.

Data/Methods

This study used a predictive model leveraging county demographics to predict county election results.

⁷⁸ Gender, race, education, generation, religion, urban/rural county voter

The county demographic data used in this study is readily available online from 2010 onwards in the US Census Bureau data set from its American Community Survey program. The demographics considered in the predictive model of this study are median household income, unemployment rate, education (proportion of the county population with a bachelor's degree or higher), age (proportion of the county population aged 60 years and older), place of birth (proportion of the county population that was born inside the United States), minority population (proportion of the non-white county population), and the population density. These specific demographics were selected in line with previous studies.

This study used the 5-year studies and its estimates from 2010 to 2020 (the years chosen correspond with each general election: 2010, 2012, 2014, 2016, 2018, and 2020) for every demographic except population density. The 1-year estimates, while more accurate, are only published for geographies with a population of 65,000 or more. Because of data availability reasons, this study used data for population density *only* from a 2020 Census Bureau dataset, which is used as a parameter in models for every year analyzed.⁷⁹ This still captures the essence of whether or not a county is rural or urban, as most county populations have not changed drastically since 2010.

The county election results included in this study come from a dataset created by Algara and Amlani. They include Democratic raw votes, Republican raw votes, and the total two-party votes for each county. A new parameter *result* was created which is the proportion of Democratic votes to the total two-party votes. Using a proportion of votes instead of a binary result (win or loss) allows the model to consider the scale of the victory or loss to adjust the parameter weights accordingly.

Variable	Observations	Mean	Std Dev	Min	Max
Result (Presidential)	5915	0.35	0.15	0.03	0.93
Result (Gubernatorial)	5259	0.38	0.14	0.01	0.93
Result (Senatorial)	7844	0.37	0.15	0.01	0.84

Independent Variables

Variable	Observations	Mean	Std Dev	Min	Max
Median Household Income	12134	44881.45	11679.56	10550.00	125672.00
Age 60 yrs and Older	12134	14.98	435	0.00	51.40

⁷⁹ Used for data availability reasons: www2.census.gov/geo/docs/reference/ua/

Population Density	12134	105.22	166.25	0.04	999.95
Bachelors or higher	12134	18.56	8.40	0.00	68.40
Native Pop. Ratio	12134	0.96	0.05	0.28	1.0
Unemployment Rate	12134	8.11	4.00	0.00	36.00
Percent Minority	12134	15.54	16.20	0.00	96.10

Figure 1. Data Summary Statistics

Both datasets include common metadata including *year* and the county *FIPS* code, which uniquely identifies every county. This allows for the linking of each election result in a county from a specific year to the corresponding demographics of the county in that same year. The Pandas library within Python was used to clean and combine the data. Any observations with missing or highly abnormal data were excluded from this analysis.⁸⁰

Year	FIPS	Median household income	Age 60 years and older	...	Result
2010	01003	\$50,147	15.5	...	0.191781
2010	01005	\$33,129	11.0	...	0.487186
2010	01007	\$41,770	9.7	...	0.244469
...

Figure 2. Sample Data Structure

A random forest algorithm was implemented with the Python library *scikit-learn*, a standard, and interpretable machine learning model, using the variable *result* as the outcome variable of the model and the demographics as predictors.

Analysis and Discussion

⁸⁰ Missing or unusual data includes uncontested elections and demographic data that does not exist for the indicated year and county.

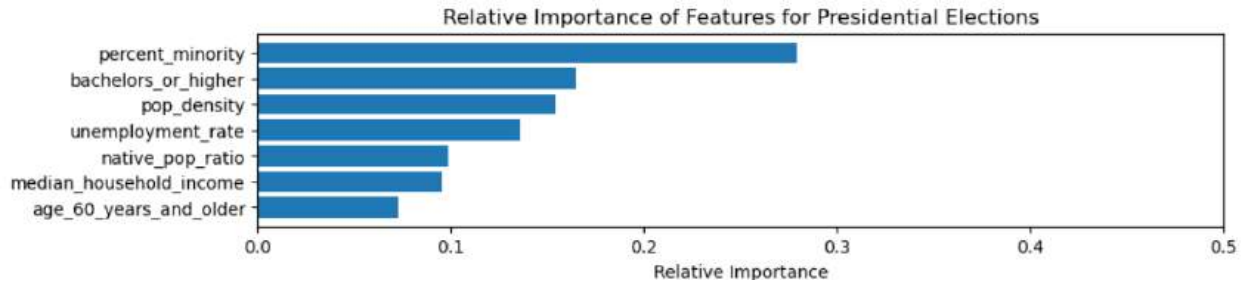


Figure 3. Presidential Election Results ($r^2 = 0.57$)

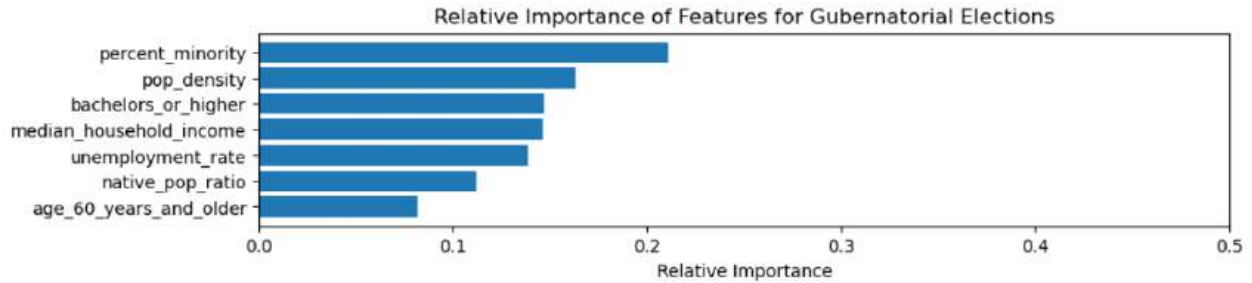


Figure 4. Gubernatorial Election Results ($r^2 = 0.38$)

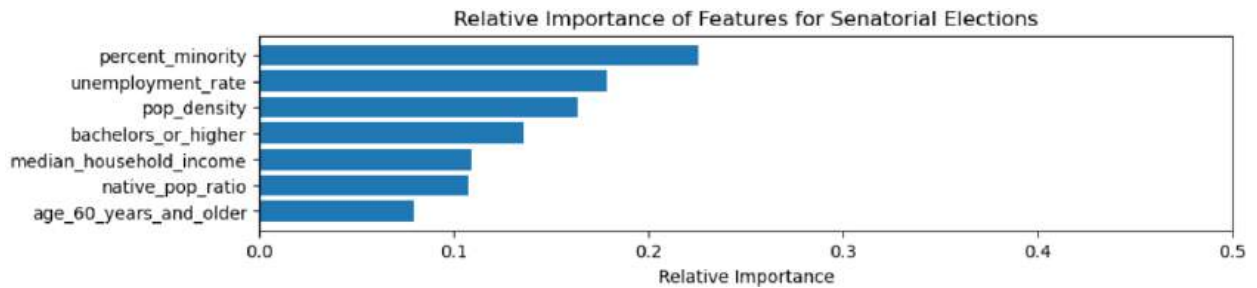


Figure 5. Senatorial Election Results ($r^2 = 0.46$)

Variables	Presidential	Gubernatorial	Senatorial
Median Income	0.10	0.15	0.11
Age	0.07	0.08	0.08
Pop. Density	0.15	0.16	0.16
Education	0.16	0.15	0.14
Native Pop.	0.10	0.11	0.11
Unemployment	0.14	0.14	0.18
Percent Minority	0.28	0.21	0.23

Figure 6. Relative Feature Importance

The data collected in this study reveals interesting trends that could be expanded upon in future research. The differences in the feature importance of most variables are marginal, but three stand out: income, unemployment, and minority population. There could be a variety of causes for these differences: governors may be held more accountable for regional economic conditions, like state income taxes, while the voter turnout among the minority population may be greater for presidential elections.

Other variables have extremely similar feature importance. Age, population density, native population, and education all predict the outcome of each election level by a similar amount. These variables likely are not affected as much as the aforementioned variables by the different characteristics of each office.

The exact *causes* of these differences or similarities are not clear, additional analysis would be helpful to understand the underlying patterns before arriving at any firm conclusion.

The differences in the relationship between social/demographic characteristics and party voting behavior could help interest groups and parties tailor campaign messaging to better align with select voters. While the actual causal effect of demographic differences on election outcomes across office levels is not vast, they may be significant enough to warrant further investigation and potential campaign strategy changes.

Limitations and Future Work

Several obstacles arose during this study. Data availability was a challenge. For example, religion is a notable demographic variable influencing voter party affiliation, but the US Census Bureau does not collect information about religion due to federal laws. Future studies including religion would need to be smaller, limited by samples collected from private researchers. High-quality election data spanning Presidential, Gubernatorial, and Senatorial elections were also sparse. With higher quality and more prevalent election data, this study could analyze additional offices and jurisdictions beyond Presidential, Gubernatorial, and Senatorial elections. Ideally, additional studies could be conducted on the municipal level, where the differences in voter behavior are likely much more pronounced than in the larger-scale elections used in this study.

In the future, this study could be expanded to include data from the 20th century. There is much less demographic data available from the US Census Bureau before 2005 given that is when the American Community Survey (ACS) began (this study used data from 2010 onward because that is the data easily accessible online). A similar analysis without data from the ACS would have to rely on decennial census records, which are only available for every ten years.

A more retrospective study would also have to take shifting political party platforms and the resulting demographic affiliation shifts into account. For example, the Republican Party under Ronald Reagan and Donald Trump are distinctly different; their supporters are likely also distinctly different. Because of party differences through time, it would likely be necessary to develop separate models dividing the “generations” of political parties, as the importance of each demographic may have shifted significantly over the past few decades. Future studies could

compare the importance of each demographic in determining voting outcomes across *party generations* to uncover specific changes in party platforms.

Of note is also how the coefficient of determination (r^2) for each model varies significantly. In the models with a lower r^2 , other demographic variables significant to voter behavior are likely missing (omitted variables). Future analysis could incorporate more demographic variables overall into the models.

Conclusion

The base findings of this study do indeed confirm the results of previous studies: demographic factors predict election results distinctly across office levels. This study identified income, unemployment, and minority population as the demographics with the biggest differences in influencing election results, with age, population density, native population, and education having little to no difference.

This study mainly discusses the *association* of certain demographic variables with election results. This research is not intended to discover *causality*, but rather to identify interesting relationships for further analysis in additional studies. Future research could expand upon how the demographic variables discussed in this study directly relate to voter priorities.

Works Cited

- Algara, Carlos, and Sharif Amlani. *Replication Data for: Partisanship & Nationalization in American Elections: Evidence from Presidential, Senatorial, & Gubernatorial Elections in the U.S. Counties, 1872-2020*. 1.1, Harvard Dataverse, 3 Sept. 2022. *Harvard Dataverse*, <https://doi.org/10.7910/DVN/DGUMFI>.
- Atkeson, Lonna Rae, and Randall W. Partin. "Economic and Referendum Voting: A Comparison of Gubernatorial and Senatorial Elections." *The American Political Science Review*, vol. 89, no. 1, 1995, pp. 99–107. *JSTOR*, <https://doi.org/10.2307/2083077>.
- Census Bureau Data*. <https://data.census.gov/>. Accessed 1 Oct. 2024.
- Kurtbaşı, İhsan. "The Factors Influencing Voting Preferences in Local Elections " An Empirical Study " ." *International Journal of Humanities and Social Science*, vol. 5, no. 9, Sept. 2015, pp. 197–210.
- Pew Research Center. "Trends in Party Affiliation among Demographic Groups." *Pew Research Center*, 20 Mar. 2018, <https://www.pewresearch.org/politics/2018/03/20/1-trends-in-party-affiliation-among-demographic-groups/>.
- Stein, Robert M. "Economic Voting for Governor and U. S. Senator: The Electoral Consequences of Federalism." *The Journal of Politics*, vol. 52, no. 1, 1990, pp. 29–53. *JSTOR*, <https://doi.org/10.2307/2131418>.
- Tanuzi, Besjon, and Justin Franklin. *How Do Socio-Demographic Factors Determine the Political Party Affiliation of Americans? Do Race and Region Affect Differently for Similar Political Views Levels?* 4693183, 12 Dec. 2022. *Social Science Research Network*, <https://doi.org/10.2139/ssrn.4693183>.

The Legalization of Gambling: Does it Really Matter? By Erich Paredes

Abstract

The legalization of gambling is a highly contested and controversial topic, for, while it generates tax revenue, it can increase harmful gambling behavior, which lower-income individuals are more vulnerable to. However, does legalizing gambling really make a difference in the quantity of individuals that gamble and the amount being gambled? How does it then relate to income inequality? Using individual-level data that tracked how many people from various US states have gambled throughout the years and state-level data on commercial gambling revenues, multiple DiD analyses were conducted. Additionally, with the respective states' Gini coefficient and commercial casino gaming revenue, a regression was conducted utilizing the instrumental variable of seat belt usage to control for endogeneity between inequality and gambling. From the DiD analyses, it was found that the legalization of gambling does not have significant effects on the number of individuals that gamble, but significantly impacts gambling revenues. This implies that, while legalizing gambling does not seem to convert people into gamblers, those who were previously already gamblers are likely increasing the amount they're gambling or betting. Through the regression, a positive relationship between gambling revenue and income inequality was found, portraying how legalizing gambling may perpetuate income inequality. However, this should be taken with caution due to a lack of statistical confidence

Introduction

Addictive behaviors have always been a prominent issue within society. It's becoming more prominent in this globalized, digital world, where opportunities to engage in addictive behaviors are increasing. One such example is the rise of online gambling, which has likely drawn more people to gamble given how accessible it currently is. It poses a potential threat to societal wellbeing, as it could lead to more diagnoses of gambling addictions and disorders. However, with the consistent rise in the accessibility of something that's as addicting as online gambling, it's to be expected that this will increase pathological gambling¹.

As such, much thought goes into the processes of legalizing something as addictive as gambling, given the need to consider its implications on public health. However, does it really matter whether gambling is legal? After all, it's no secret that even in regions where the act of gambling is illegal, there are still bound to be black markets operating for gambling. Thus, it puts into question if legalization really makes a significant difference on how gambling can impact societal wellbeing. For if it does not have a significant effect, then the disadvantages that prevent governing bodies from taking the step to legalize gambling may in fact be negligible, instead highlighting any possible merits that may exist with the legalization. Such a finding could potentially alter public policies revolving around the legalization of addictive behavior.

Yet, despite the significance of such findings, it's a topic that has rarely been covered within past research, for many works seemingly presume that legalizing gambling goes hand-in-hand with increased gambling behaviors.

Past research has long since considered the ethicality of legalizing gambling. Many have justified the positive impacts of legalizing gambling through its value in generating tax revenue, particularly in areas where collecting taxation from wealthier individuals and larger corporations can be a challenging process². Additionally, some argue that it's a way to increase funding towards social welfare and spur increased economic growth with Italy being one such example given how gambling tax revenue amounts to over 2% of total tax revenue²⁻⁴. However, this has been open to dissent, as other works have raised concerns over gambling's limited-positive contributions towards the economy when faced against the social costs of problematic gambling, as well as the highly regressive impacts of these gambling taxes⁴⁻⁶. Now, given the broadness and range within the types of gambling, Walker and Jackson (2011) took their analysis a step further, as they investigated the impacts of different forms of gambling on state revenues within the US. Their results concluded that gambling in the form of lotteries and horse racing tended to increase state revenue, whereas casinos and greyhound racing were observed to decrease state revenue⁷.

Beyond tax revenue, another common point of investigation has been the relationship between gambling and income inequality. Many papers have attempted to understand which members of society are most vulnerable to problematic gambling behaviors, for which the common finding seems to revolve around socioeconomic status. Areas with higher income inequality and unemployed individuals with low educational attainment who were receiving social welfare were pointed to as the most at-risk of developing problematic gambling habits⁸⁻⁹. As such, it portrays the dangers of legalizing gambling as the burden of its social costs are likely being carried by the poor, for betting and new generation games (i.e. gambling forms found to cater to the poor) have significantly contributed to growing gambling turnovers and rising gambling disorders¹⁰. The portrayal of a clearer linkage from gambling to income inequality is achieved by Freund and Morris (2006) in their investigations into lotteries and other gambling forms. Through their investigations they concluded that US state lotteries fostered raised income inequality; a finding which was not mirrored by any of the other gambling types analyzed¹¹. However, while gambling raises inequality, numerous studies have pointed towards how an increase in Gini coefficient results in a raise in gambling turnover, particularly lottery expenditures and online gambling^{6,12-13}.

When it comes to studying the effects of increasing gambling opportunities, Govoni et al's (1998) investigation into first-year impacts of Casino Windsor's opening found that there were no significant increases in problematic gambling behaviors¹⁴. While post-casino surveys did yield evidence of higher-spending gamblers, per capita gambling expenditures did not significantly change¹⁴. Now, while this portrays that increasing opportunities for gambling does not necessitate an increase in gambling behaviors and expenditure, Pearce et al's (2008) analysis on New Zealand's neighborhoods presents an opposing view. It found that those living nearest gambling venues were more likely to be gamblers or even problematic gamblers compared to

those living the furthest from these venues¹⁵. Now, while these two papers address the effects of increased opportunities for gambling, neither address how legalizing gambling particularly impacts gambling behaviors. It's only Chóliz's (2015) paper analyzing the effects of legalizing online gambling on problematic gambling that has investigated the legalization component of gambling. Chóliz found that legalizing online gambling significantly increased the number of young pathological gamblers¹. While this does cover the impacts of legalization on problematic gambling behavior, it does not quite illustrate a clear picture of what this implies for the proportion of individuals that gamble or overall quantities of gambling expenditures: two factors that may impact the framework of implementing these legalization policies.

As such, that's what this paper intends to understand. It does so by conducting an analysis on gambling within various U.S. states, where attitudes towards gambling have seen recent changes on both an individual and political scale. The Wire Act of 1961 has long since prevented states from legalizing online gambling, as it states that bookmakers are not allowed to take in sports bets over the phone¹⁶. However, in 2011, the Department of Justice repealed the initial interpretation of this act, noting that it particularly only covers sports betting¹⁷. Hence, under that technicality, other forms of gambling were permissible if legalized by states¹⁷. Thus, it resulted in consequent legalizations by various states with six states offering online casinos¹⁷. However, this narrative has been complicated by the Department of Justice's return to its prior interpretation of the Wire Act of 1961 in 2019, but this change in understanding has not prevented most states from their ongoing efforts to legalize online gambling¹⁷.

More recently, the Supreme Court's overruling of *Murphy v. National Collegiate Athletic Association* case in 2018 marked a new era for gambling in the U.S., as it gave way for individual states to slowly legalize online and retail sports betting¹⁸⁻¹⁹. Prior to this change in legislation, Nevada had been the only state where sports betting was legal due to a prior law in 1992 that allowed for its exemption¹⁸. However, as a result of this ruling, 38 states have now legalized sports betting, depicting the changing landscape in the U.S.' attitudes towards gambling²⁰.

On a more individual scale, the American Gaming Association's 2024 survey shows that 88% of Americans now perceive casino gambling to be an acceptable behavior, be it for themselves or others²¹. With 55% of Americans having partaken in some type of gambling in the past year and casino visitations being at an all-time high of 49%, the legal acceptance of gambling seems to reflect the improved perceptions and openness that Americans have towards this recreational activity²¹. It seems that this positive change in perception traces back to Americans' growing recognition of the benefits that gambling has had on the U.S. economy²¹. Thus, it portrays why U.S. states were viable choices for this investigation, as the improvements in how gambling is seen opens up opportunities to analyze legalizations of new forms of gambling or liberalizations in current policies surrounding gambling.

Within this paper, I will be answering the following research question. Does legalizing gambling really make a difference in both the quantity of individuals that gamble and the amount being gambled? How does it then relate to income inequality? It was hypothesized that legalizing

gambling will not significantly increase the proportion of individuals that gamble, but it will raise overall measures of gambling revenue, as pre-existing gamblers increase the quantities they're gambling. Furthermore, it was also hypothesized that the legalization of gambling would not significantly increase income inequality.

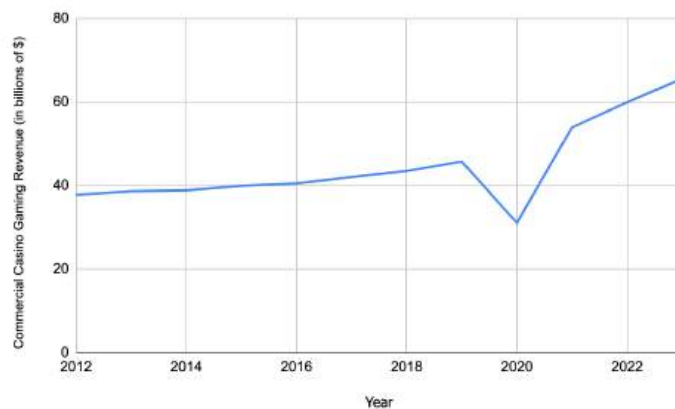
This paper will test these hypotheses through the use of more recent datasets and methodologies like DiD analyses and regressions with instrumental variables, which will further set it apart from past literature on these topics and allow it to uniquely contribute towards increasing knowledge on the legalization of gambling and income inequality.

Within this paper, it was found that legalizing gambling did not increase the proportion of individuals that gambled, albeit we can't be confident with the significance of this conclusion due to limitations in data. However, it did significantly increase commercial casino gaming revenues. Although, there may have been biases impacting these results. Moreover, the regression analysis concluded that legalizing gambling does increase income inequality, but the strength of this conclusion is questionable due to low statistical confidence.

Summary Statistics

Looking at gambling trends within the US, Figure 1 shows how casino gambling has been increasing within the past decade as reflected by growing annual commercial casino gaming revenues. There's a notable spike downwards that takes place in 2020, which is likely attributable to the outbreak of the COVID-19 pandemic. The consequent lockdowns prevented casino operations for a few months; thus, explaining the anomaly in the trend. However, post-pandemic, casinos have returned to pre-pandemic levels of activity, continuing the trend of increased casino gaming activity.

Figure 1: USA Commercial Casino Gaming Revenue from 2012 to 2023²²



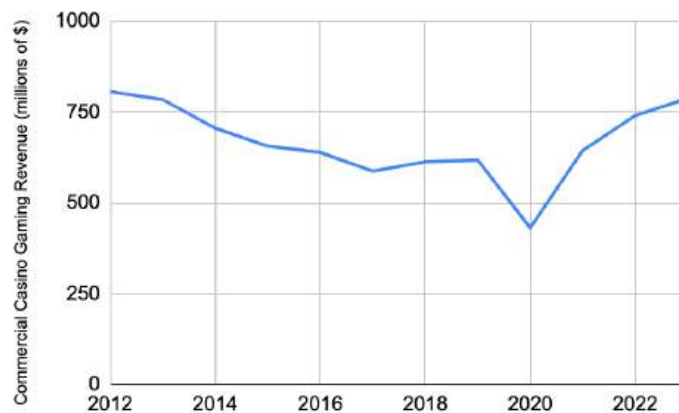
Looking at the individual states, however, there is observable heterogeneity within trends. The only notable similarity shared within all individual trends is the evident decrease in commercial casino gaming revenue during 2020, caused by temporary halts to casino operations. However, otherwise, the tendencies of graphing the states' commercial casino gaming revenues pre-pandemic ranged from consistently increasing, decreasing, or being relatively stagnant.

These distinctions remained even during the post-pandemic period. Albeit, none of the observed states displayed any decreases in commercial casino gaming revenue after the pandemic.

To better show these distinctions, Figures 2, 3, 4, and 5 display trends that were observed across the chosen states.

Figure 2 shows an evident decrease in the trends of commercial casino gaming revenue from 2012 to 2017. This particular trend is only observable within West Virginia, where an average of \$36,398,359 was lost annually with the largest year-on-year drop in revenue being approximately \$79,149,000²². It, however, does see an increasing trend from 2018 onwards — with the exception of 2020 — with West Virginia’s commercial casino gaming revenue in 2023 being nearly comparable to its revenues in 2012.

Figure 2: West Virginia’s Commercial Casino Gaming Revenue from 2012 to 2023²²



In Figure 3, a different trend is shown, depicting commercial casino gaming revenues that remain stagnant throughout 2012 to 2023 with 2020 being the only exception. This trend can only be seen in the state of South Dakota, where, after the pandemic, revenues returned to pre-pandemic levels. Similar to Figure 3, Figure 4 depicts a relatively stagnant trend in pre-pandemic years. However, post-pandemic, the commercial casino gaming revenue within South Dakota spikes upwards to a new equilibrium. Comparatively, South Dakota’s average revenues pre- and post-pandemic are \$101,073,000 and \$138,133,000, respectively²². It’s a trend that can also be seen in other states such as Colorado, Pennsylvania, Iowa, and Michigan.

Figure 3: Missouri's Commercial Casino Gaming Revenue from 2012 to 2023²²

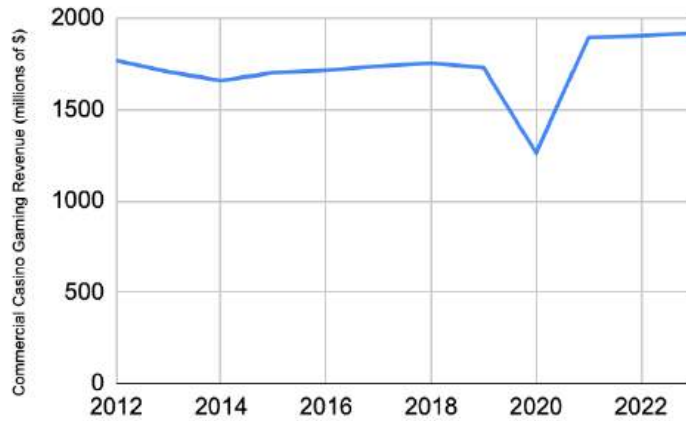
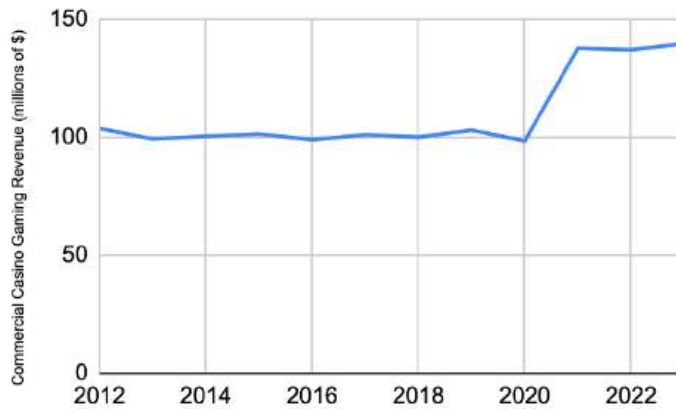
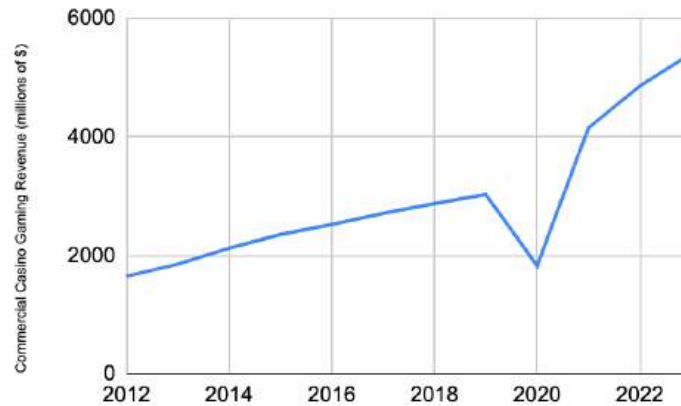


Figure 4: South Dakota's Commercial Casino Gaming Revenue from 2012 to 2023²²



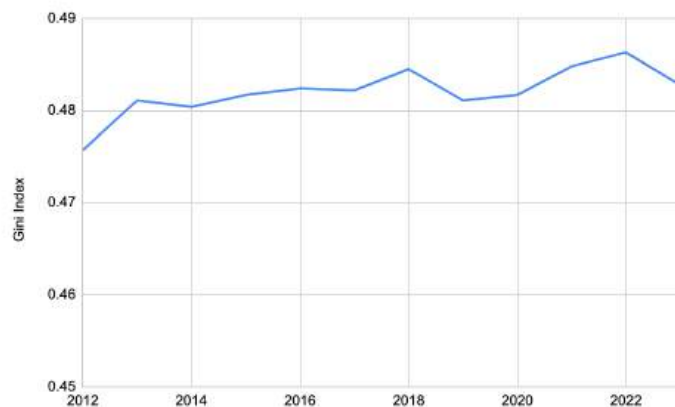
Finally, in Figure 5, the graph portrays a consistently increasing trend in commercial casino gaming revenues from 2012 to 2023, which was only disrupted by 2020's downward spike. The rate of increase does increase post-pandemic, going from an average annual increase in revenue of \$172,669,870 to \$418,192,026²². Beyond Illinois, this trend of increasing revenues is also observable in other states, including Florida, Ohio, Maryland, and Maine.

Figure 5: Illinois' Commercial Casino Gaming Revenue from 2012 to 2023²²



Now, looking at trends in income inequality, Figure 6 shows that income inequality has not seen any significant changes from 2012 to 2023 with the Gini coefficient having a range of 0.0106²³. However, the Gini coefficient is growing, albeit slightly, even achieving its maximum value of 0.4863 in 2022²³. This means that income inequality within the US has slowly been rising. This may partially be attributed to the continuous increase in gaps between the wealth of upper-, middle-, and lower-income families²⁴. Middle-income families are significantly losing their share of the cumulative wealth, shifting from 32% in 1983 to 17% in 2016²⁴. Whereas upper-income families have seen significant increases in their share of the wealth, seeing a 19% growth within the past three decades²⁴.

Figure 6: US Income Inequality from 2012 to 2023²³



Methodology: Data

For the purposes of this paper, data was collected on 3 sets of variables: gambling metrics, income inequality metrics, and seat belt usage.

When it came to gambling metrics, data on aggregated individual-level gambling behavior was collected for the states of Pennsylvania, Massachusetts, and North Carolina. This

Finally, data was collected on seat belt usage through the National Highway Traffic Safety Administration for all the chosen states from 2013 to 2022³³. Albeit, it is worth noting that for a majority of the states, data on seat belt usage for 2020 was missing.

Methodology: Analysis

To test the first part of the hypothesis, multiple DiD analyses were conducted on both the individual-level and state-level gambling data, as a means of observing the effect of policy changes on trends within treatment and control variables pre- and post-treatment.

When conducting state-by-state comparisons for the DiD analysis of individual-level data, Pennsylvania was taken as a treatment variable given that the state legalized sports betting in 2017, before then officially launching retail sports betting in 2018 and online sports betting in 2019³⁴. Meanwhile, North Carolina and Massachusetts acted as control states, for, during the years of comparison (i.e. 2012 and 2020 for North Carolina, 2013/2014 and 2019 for Massachusetts), neither of these states legalized any new forms of gambling³⁵⁻³⁶.

Beyond state-by-state comparisons, a DiD analysis was also conducted on region-by-region comparison of individual-level data in North Carolina for the years of 2014, 2016, 2018, and 2020. For which, aggregated data for regions 1 and 2 acted as treatment variables, given that a liberalization in gambling policies took place within Cherokee County in region 1, where the state government permitted the opening of a 2nd tribal casino in 2015^{35,37}. This is particularly significant given that North Carolina does not permit the operation of commercial casinos, and a 3rd tribal casino was only permitted to operate starting from 2021³⁵. As such, comparisons in trends of the proportion of individuals that gambled on both a monthly and annual basis were made for regions 1 and 2 against the control variables of regions 3 through 8, where a change in gambling policies did not take place within the chosen timeframe. Multiple DiD analyses were run for this, as individual comparisons were made between regions 1 and 2 and each control region, alongside a collective analysis that compared regions 1 and 2 to the collective trends in all other regions. Further distinctions in DiD analyses were made, as it was run for both the annual and monthly data, and for each set of comparisons, the DiD was run twice to include and exclude the year 2020, where anomalies may have been present due to the impact of the COVID 19 pandemic.

A DiD analysis was also conducted for the comparisons of macro-level data on commercial casino gaming revenue. This was conducted individually for the years of 2018 to 2023, during which states legalized sports betting. It took into account two methodologies. There was the not all controls methodology, which only considered states that hadn't legalized sports betting by 2023 as control variables. The second methodology was the all controls methodology, where control variables were not just limited to states that had yet to legalize sports betting. Instead, it also included states that legalized sports betting after the year that was being analyzed as the treatment year for the DiD analysis; however, it only used pre-treatment data for these states. For instance, if the DiD analysis was studying legalizations that took place in 2019, states that legalized sports betting in 2020 to 2023 were still included within the dataset, but only for

the years prior to their legalization of sports betting. The use of multiple methodologies was a good measure to ensure the validity of findings derived from the DiD analysis. For additional clarity on the policy changes that dictated which states were considered as treatment and control variables, refer to Table 1.

Table 1: Changes in the legality of sports betting across the observed states.^{34, 38-59}

State	Policy Change
CO	Legalized sports betting in May 2020
DE	Legalized retail sports betting in June 2018
FL	Legalized retail sports betting in December 2023
IL	Legalized sports betting in June 2019, but retail sports betting was not launched until March 2020
IN	Legalized sports betting in May 2019, but retail sports betting was not launched until September 2019
IA	Legalized sports betting in August 2019
KS	Legalized sports betting in July 2022, but it was not launched until September 2022
LA	Legalized sports betting in June 2021, but retail sports betting was not launched until October 2021
ME	Legalized sports betting in July 2022, but it was not launched until November 2023
MD	Legalized sports betting in May 2021, but retail sports betting was not launched until December 2021
MI	Legalized sports betting in December 2019, but retail sports betting was not launched until March 2020
MS	Legalized retail sports betting in August 2018
MO	Has not legalized sports betting
NJ	Legalized sports betting in June 2018
NM	Legalized retail sports betting in October 2018
NY	Legalized sports betting in 2013, but retail sports betting was not launched until July 2019
OH	Legalized sports betting in 2021, but it was not launched until January 2023
OK	Has not legalized sports betting

PA	Launched retail sports betting in November 2018
RI	Legalized sports betting in June 2018, but didn't launch retail sports betting until November 2018
SD	Legalized sports betting in September 2021
WV	Legalized sports betting in March 2018, but retail sports betting was not launched until August 2018

Through the above DiD analyses, the following equation can be constructed, demonstrating the effect of legalizing gambling on gambling behaviors and expenditure. It's worth noting that a simplified DiD analysis was conducted without accounting for control variables, which is reflected in the equation below⁶⁰.

$$y_{i,t} = \beta_0 treat_{i,t} + \beta_1 post_{i,t} + \beta_2 treat * post_{i,t} + \lambda_{i,t}$$

Within the equation, the variable i represents the chosen state while t represents the chosen. Meanwhile, $treat_{i,t}$, $post_{i,t}$, and $treat * post_{i,t}$ are binary variables, which are represented as either 0s or 1s depending on the values for i and t . Now, $treat_{i,t}$ is a 1 if i is a treated state and a 0 if it's a control state. Meanwhile, $post_{i,t}$ is a 1 if the chosen t takes place on or after sports betting was legalized, but if it took place pre-legalization, then it would be a 0. Finally, $treat * post_{i,t}$ is the interaction variable, which only takes on the value of 1 if i is a treated state and t is a year on or after sports betting was legalized; otherwise, a 0 would be substituted. The variables β_0 , β_1 , and β_2 are the coefficients of the $treat_{i,t}$, $post_{i,t}$, and $treat * post_{i,t}$ variables, respectively, while the variable $\lambda_{i,t}$ is the error term: all 4 of which are derived from the DiD analysis. The variable of $y_{i,t}$ is then the metric for gambling behaviors and attitudes in state i and year t . Within the micro-level analysis, this represents the proportion of individuals that gamble in a year, whereas, in the macro-level analysis, it represents commercial casino gaming revenue.

The interpretation of a few variables would slightly vary for the North Carolina micro-level analysis. Firstly, i would represent one of North Carolina's regions, as opposed to a state. As such, the binary variable $treat_{i,t}$ would only be a 1 if the chosen i is region 1/2, where a new tribal casino was opened. Meanwhile, $post_{i,t}$ would only be a 1 if the chosen t took place after 2015, which is when the policy change that allowed for the opening of a 2nd tribal casino took place. Finally, interaction variable $treat * post_{i,t}$ would only be a 1 if i is region 1/2 and t is greater than 2015.

Now, to test the second half of the hypothesis, a regression analysis was conducted between the commercial casino gaming revenue and Gini coefficient. However, in order to achieve causality between these two variables, an instrumental variable is needed, as it could control for the endogeneity between gaming revenue and income inequality that arises due to reverse causality.

The instrumental variable, seat belt usage, was chosen, as it's a measure of risky behavior with individuals who always wear seat belts being more risk-averse. Thus, it presented a possibility for causality with gambling revenues, as gambling is also another measure of risky behavior. There's also room to argue that seat belt usage and Gini coefficient exhibit exogeneity, as the inclusion of seat belts across all vehicular models that cater to passengers implies that incomes should not have an impact on seat belt usage. However, there arises a potential for endogeneity since lower-income individuals may not be as inclined to use seat belts, given their lack of awareness or education on the potential dangers that come with not wearing a seat belt.

Nonetheless, seat belt usage was not a weak instrument according to the weak instrument test, which was conducted through the first-stage regression between seat belt usage and gambling revenues. As such, the regression between gambling revenue and income inequality was run with seat belt usage being integrated as an instrumental variable for gambling revenue. The regression also included control variables of unemployment rates, educational attainment beyond high school, educational attainment beyond a bachelor's degree, and GDP per capita. However, given potential endogenous relationships between seat belt usage and income inequality, additional statistical tests like the Manual Residual Inclusion Test and Durbin-Wu-Hausman Test for Endogeneity were conducted, as a means of observing the possibility of endogeneity.

To better show the parts of the regression analysis, below are the equations that can be constructed from the above investigation⁶¹.

$$gambyr_{i,t} = \gamma_0 seatbelt_{i,t} + \gamma_1 unemploy_{i,t} + \gamma_2 gdp_{i,t} + \gamma_3 hs_{i,t} + \gamma_4 bachelor_{i,t} + \delta_{i,t}$$

$$gini_{i,t} = \alpha_0 \widehat{gambyr}_{i,t} + \alpha_1 unemploy_{i,t} + \alpha_2 gdp_{i,t} + \alpha_3 hs_{i,t} + \alpha_4 bachelor_{i,t} + \epsilon_{i,t}$$

Similar to the DiD analysis, variables i and t represent the chosen state and year, respectively. Now, the instrumental variable $seatbelt_{i,t}$ represents the percentage of individuals that always wore seat belts in state i and year t . Meanwhile, $unemploy_{i,t}$, $gdp_{i,t}$, $hs_{i,t}$, and $bachelor_{i,t}$ are the control variables that correspondingly represent unemployment rates, GDP per capita, the percentage of people that were high school graduates or higher, and the percentage of people that received their bachelor's degree or higher in state i and year t . Variables γ_0 , γ_1 , γ_2 , γ_3 , γ_4 , α_0 , α_1 , α_2 , α_3 , and α_4 are the coefficients of the instrumental and control variables, which can be found by running the regression. Similarly, variables $\delta_{i,t}$ and $\epsilon_{i,t}$ are the error terms that

are derived from the regression analysis. Now, $gambyr_{i,t}$ represents the actual commercial casino gaming revenues of state i during year t . However, $\widehat{gambyr}_{i,t}$ represents the fitted commercial casino gaming revenues, measuring predicted gambling revenues by not accounting for the error term $\delta_{i,t}$ in the first equation. Finally, $gini_{i,t}$ represents the Gini coefficient, as a value from 0 to 1, for state i during year t .

Results and Discussion: Micro-Level DiD Analysis

The DiD analysis on state-by-state comparisons found that the change in the percentage of individuals that gamble in Pennsylvania after the legalization of sports betting decreased by 9.6% and 14.8% more compared to the change in North Carolina and Massachusetts, respectively (see in Table 2). However, there are limitations within the results, as it only considers 2 years of data for 2 states within each DiD analysis, presenting potential issues in data interpretation. Nonetheless, while we’re not confident of the significance of these findings due to limitations in the availability of data on individual-level gambling behaviors, this does suggest that legalizing gambling will not increase the proportion of individuals that gamble.

Table 2: Results of the DiD analysis on state-by-state comparisons for individual-level gambling behaviors.

Treatment	Control	
PA	NC	-9.6
PA	MA	-14.8

On the other hand, the DiD analyses on regional comparisons in North Carolina that observed the effects of liberalizations in gambling policies yielded insignificant p-values, as seen in Table 3. This indicates that all variations found within the trends of the control and treatment regions after the policy change are highly probable to merely be 0. Thus, implying that the opening of another tribal casino in North Carolina does not significantly impact the percentage of people that gamble in a month or year.

Table 3: Results of the DiD analysis on region-by-region comparisons for individual-level gambling behaviors in North Carolina.

All Years						All Years Excluding 2020					
Individuals that Gamble in a Year			Individuals that Gamble in a Month			Individuals that Gamble in a Year			Individuals that Gamble in a Month		
Co	Di	P-	Co	Di	P-	Co	Di	P-	Co	Di	P-

Control Region (s)	DiD Coefficient	Value	Control Region (s)	DiD Coefficient	Value	Control Region (s)	DiD Coefficient	Value	Control Region (s)	DiD Coefficient	Value
Al1	1.394	0.576	Al1	-0.400	0.825	Al1	1.150	0.594	Al1	-0.150	0.942
3	-0.633	0.643	3	-0.247	0.148	3	-0.350	0.752	3	-1.700	0.265
4	1.200	0.597	4	1.800	0.432	4	0.850	0.506	4	2.800	0.340
5	0.133	0.969	5	0.200	0.915	5	0.350	0.945	5	0.300	0.871
6	5.500	0.197	6	-0.263	0.311	6	4.400	0.132	6	-2.700	0.379
7	1.600	0.634	7	-1.767	0.466	7	0.850	0.750	7	-2.100	0.229
8	0.567	0.768	8	2.467	0.320	8	0.800	0.739	8	2.500	0.409

Results and Discussion: Macro-Level DiD Analysis

As for the DiD analyses conducted on the macro-level data tracking commercial casino gaming revenue, by referencing Table 4, it can be seen that the DiD coefficients are predominantly positive and the majority of the p-values are indeed significant. As such, the general trend shows that there is a positive effect in commercial casino gaming revenue after legalization, as evident in Figures 8 and 9. Within these figures we can see how the trends in control and treated states diverge after the legalization takes place. In Figure 8, from 2018 to 2019, there's an evident increase in commercial casino gaming revenues within the treated states, whereas there's a decrease in control states. And while both control and treated states faced decreases in commercial casino gaming revenue from 2019 to 2020, this is largely due to the external impacts of the COVID 19 pandemic. This is exemplified by the fact that after 2020, we can see the long-term effects of legalizing gambling on commercial casino gaming revenue for the treated states, as their revenues significantly increase, which isn't paralleled within the

control states. Similarly, in Figure 9, it's evidently noticeable that after the legalization of sports betting in 2020, treated states experienced a large spike in their commercial casino gaming revenue that control states didn't experience.

Going back to Table 4, the only anomaly within this observation is the year 2021, where the p-values are insignificant and the DiD coefficients are either comparatively small or negative. This can largely be explained due to the exceptional circumstances faced by the treatment states in 2021: Louisiana and South Dakota. Louisiana was still recovering from the impacts of Hurricane Ida, and as such, this may have prevented individuals from gambling as much even despite the legalization of sports betting. As for South Dakota, it was likely facing extreme competition from its neighboring state of Nebraska. As Nebraska legalized casino gaming horse tracks, its neighboring state of Iowa was noted to have seen decreases in casino revenue⁶². As such, it's highly probable that South Dakota also lost customers due to Nebraska's policy change.

Table 4: Results of the DiD analysis on state-by-state comparisons for commercial casino gaming revenue.

Not All Controls Methodology			All Controls Methodology		
Legalization Year	DiD Coefficient (millions in \$)	P-Value	Legalization Year	DiD Coefficient (millions in \$)	P-Value
2018	316.2002	0.03	2018	200.3824	0.057
2019	549.3663	0.001	2019	387.6265	0.012
2020	793.1368	0	2020	752.8179	0
2021	70.13092	0.378	2021	-63.9092	0.64
2022	385.0499	0.01	2022	306.3784	0.101
2023	1573.448	0	2023	1573.448	0

Figure 8: The effects of the legalization of sports betting on commercial casino gaming revenues in 2018.

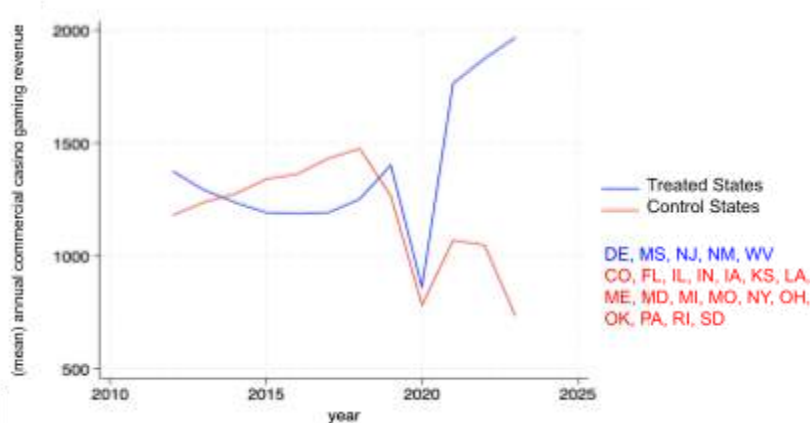
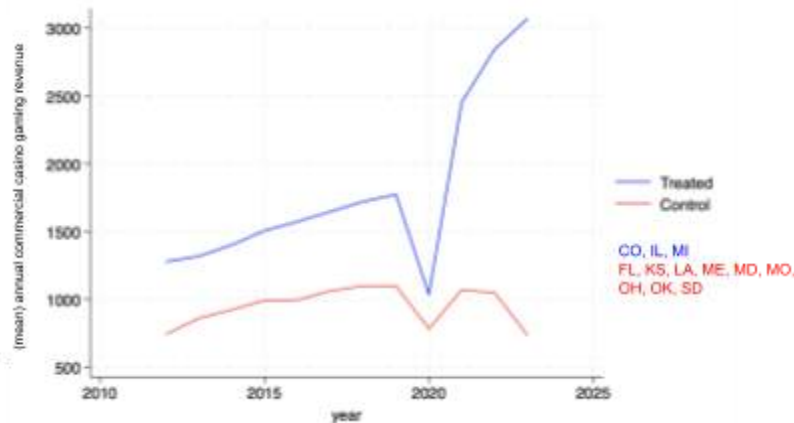


Figure 9: The effects of the legalization of sports betting on commercial casino gaming revenues in 2020.



South Dakota's special case of cross-border movement portrays the potential for bias that may have impacted the findings of the DiD analyses. A past study has found that the legalization of sports betting within a neighboring state resulted in decreased casino gambling revenues for the state that had yet to legalize sports betting⁶³. This is because the legalization of a new form gambling encourages cross-border movements towards the state where it is now legal. Thus, it increases admissions to the treated state, while reducing admissions to the control state⁶³. As such, considering the DiD analysis for 2018 where the neighboring states of New Mexico and Oklahoma were considered as treatment and control states, the legalization of sports betting in New Mexico increased state commercial casino gaming revenues while reducing revenues in Oklahoma.^{28, 53, 56} This amplifies the effect of legalizing gambling on commercial casino gaming revenues, for, in reality, the differences in the changes to gambling revenues are accounted for by cross-border movements, as opposed to the legalization. Thus, this bias is inflating the results of the DiD, for control states are affecting the results of treated states.

Beyond the bias associated with gamblers crossing state borders, there's room for further bias within how individuals who partook in illegal sports betting have just shifted towards gambling legally as a result of the legalization. Similarly, this amplifies the effect of legalizing gambling, as commercial casino gaming revenues pre-legalization would have been higher if illegal sports betting revenues were included. However, efforts to account for this bias are severely limited by the unavailability of data on revenues from illegal gambling. Nonetheless, even with this particular bias, the merits of legalizing gambling still hold, as shifting illegal sports betting to legal sports betting does raise tax revenues for the state.

Results and Discussion: Regression Analysis

The instrumental variable of seat belt usage, when correlated with gambling revenue, yielded a positive, significant weak correlation of 0.2564. However, upon running the

weak-instrument test through the first-stage regression between seat belt usage and gambling revenue, the F-statistic of 14.98, exceeding the baseline threshold of an F-statistic of 10, showed that seat belt usage was not a weak instrumental variable.

There are, however, existing issues within the potential endogeneity between seat belt usage and gambling revenue, as studies have observed that lower-income individuals were less likely to wear seat belts, potentially due to variations in educational levels that lead those in higher incomes to be more aware of the risks that come with not wearing a seat belt⁶⁴. As such, there was a need to run the following statistical tests.

As seen in Table 5, the p-values achieved ranged between 0.1211 to 0.255, implying that there's between a 12.11% to 25.5% chance of getting a value that's even more extreme than the results achieved and it still being exogenous. While ideally these p-values would've been higher to increase reassurance within the needed exogenous relationship between seat belt usage and income inequality, this was the best that we could work with. However, it would be ideal for future research to look into other options for stronger instrumental variables that could increase confidence within the regression findings.

Table 5: Statistical tests looking into the endogeneity of seat belt usage and income inequality.

Statistical Test	P-Values
Manual Residual Inclusion Test	0.255
Durbin-Wu-Hausman Test for Endogeneity (Robust Score)	0.1435
Durbin-Wu-Hausman Test for Endogeneity (Robust Regression)	0.1211

Thus, conducting the regression for gambling revenues and gini coefficient yields a regression coefficient of $1.86e^{-6}$ (see in Table 6), which implies that increasing gambling revenue by \$1,000,000 would lead to an increase in the Gini coefficient by $1.86e^{-6}$.

Contextually, Croatia's and Kyrgyzstan's Gini coefficient of 0.289 and 0.288 in 2021, respectively, has a difference of 0.001⁶⁵. However, in order to increase Kyrgyzstan's Gini coefficient by that incremental 0.001 difference, it would take a \$1 billion increase in gambling revenue.

However, it is worth noting that the p-value is 0.202, indicating that there's a 20.2% chance that the coefficient is actually 0. Therefore, we should be cautious upon implementing policies given these results.

Table 6: Regression analysis between gambling revenue and Gini coefficient.

Gini Coefficient	Regression	P-Value
------------------	------------	---------

	Coefficient	
Gambling Revenue	1.86e ⁻⁶	0.202

Conclusion

As such, based on the data analysis, we can conclude that legalizing gambling will not increase the proportion of individuals that gamble within a population. It will, however, significantly increase commercial casino gaming revenue, implying that individuals that were previously gamblers are now gambling in higher quantities. Thus, this suggests that while legalizing gambling will not draw more people to gamble, it does have positive benefits of raising tax revenues that may go towards investments in social welfare. However, it's worth noting to take these conclusions with some caution due to the presence of biases or lack of confidence within the significance of findings.

Moreover, given the positive relationship between gambling revenue and income inequality, this implies that the legalization of gambling, which significantly increases gambling revenues, may also have a positive relationship with income inequality. It's worth taking this conclusion with some caution as well due to high p-values and, consequently, low statistical confidence. Nevertheless, regardless of that, this finding shouldn't necessarily force policy makers to detract from legalizing gambling. For if tax revenues from the legalization are managed well in its distributions to the improvement of social welfare, this has the potential to offset the increase in income inequality that arises from higher gambling revenues caused by the legalization of gambling. It's not to say that legalizing gambling is entirely harmless, as previous literature has shown evidence of its contribution towards the rise of gambling disorders¹.

Now, there's still much room for future research on this topic. Most prominently, a stronger instrumental variable should be used to test the validity of the concluded positive relationship between gambling revenue and income inequality found within this paper. However, other areas of research that can be further pursued would be analyzing non-aggregated individual-level data on gambling behaviors that would allow researchers to track how individuals themselves change their gambling attitudes and behaviors as a result of legalization. Furthermore, understanding the applicability of these findings within other countries outside the scope of the United States, as well as other commonly covered nations like Italy and Spain, could be another potential avenue for future researchers to investigate. Cultural variations could shift how legalization impacts gambling behaviors and expenditures; as such, diversifying the countries analyzed would be viable. A potential candidate for research could be Japan, where gambling legislations are seeing changes as plans to open its very first gambling resort were approved last year⁶⁶.

Acknowledgments

Special thanks to Jose Maximo for his mentorship and guidance in writing this paper.

Works Cited

- Chóliz, Mariano. "The challenge of online gambling: the effect of legalization on the increase in online gambling addiction." *J Gambl Stud* 32 (2016): 749–756.
<https://doi.org/10.1007/s10899-015-9558-6>
- Nikkinen, Janne. "The legalization of dangerous consumption: a comparison of cannabis and gambling policies in three US states." *Addiction Research & Theory* 25, no. 6 (2017): 476–484. <https://doi.org/10.1080/16066359.2017.1366455>
- Walker, Douglas M. Jackson, John D. "New goods and economic growth: evidence from legalized gambling." *Review of Regional Studies* 28, no. 2 (1998): 47-69.
<https://doi.org/10.52324/001c.8887>
- Gandullia, Luca. Leporatti, Lucia. "Distributional effects of gambling taxes: empirical evidence from Italy." *J Econ Inequal* 17 (2019): 565–590.
<https://doi.org/10.1007/s10888-019-09423-9>
- Günay, Hamdi Furkan. "Taxation as a policy instrument for social and economic effects of gambling." *Journal of Gambling Issues* (2023). <https://doi.org/10.4309/MKME8615>
- Ferri, Giovanni. Lagravinese, Raffaele. Resce, Giuliano. "Inequality and gambling." *Economics & Statistics Discussion Papers* (2024).
<https://ideas.repec.org/p/mol/ecsdp/esdp24095.html>
- Walker, Douglas M. Jackson, John D. "The effect of legalized gambling on state government revenue." *Contemporary Economic Policy* 29, no. 1 (2011): 101 - 114.
<https://doi.org/10.1111/j.1465-7287.2010.00198.x>
- Canale, Natale. Vieno, Alessio. Lenzi, Michela. Griffiths, Mark D. Borraccino, Alberto. Lazzeri, Giacomo. Lemma, Patrizia. Scacchi, Luca. Santinello, Massimo. "Income inequality and adolescent gambling severity: findings from a large-scale Italian representative survey." *Frontiers in Psychology* 8 (2017).
<https://doi.org/10.3389/fpsyg.2017.01318>
- Henkel, Dieter. Zemlin, Uwe. "Social inequality and substance use and problematic gambling among adolescents and young adults: a review of epidemiological surveys in Germany." *Current Drug Abuse Reviews* 9, no. 1 (2016): 26-48.
<https://doi.org/10.2174/1874473709666151209114023>
- Resce, Giuliano. Lagravinese, Raffaele. Benedetti, Elisa. Molinaro, Sabrina. "Income-related inequality in gambling: evidence from Italy." *Rev Econ Household* 17 (2019): 1107 - 1131. <https://doi.org/10.1007/s11150-019-09468-9>
- Freund, Elizabeth A. Morris, Irwin L. "Gambling and income inequality in the States." *Policy Studies Journal* 34, no. 2 (2016): 265 - 276.
<https://doi.org/10.1111/j.1541-0072.2006.00169.x>
- Bol, Thijs. Lancee, Bram. Steign, Sander. "Income inequality and gambling: a panel study in the United States (1980–1997)." *Sociological Spectrum* 34 (2014); 61 - 75.
<https://doi.org/10.1080/02732173.2014.857196>

- Pabayo, Roman. Patel, Priya. Patte, Karen A. Leatherdale, Scott T. “Income inequality and the odds of online gambling among a large sample of adolescents in Canada.” *J Gambl Stud* 40 (2024): 289–305. <https://doi.org/10.1007/s10899-023-10194-2>
- Govoni, Richard. Frisch, G. Ron. Rupcich, Nicholas. Getty, Heather. “First year impacts of casino gambling in a community.” *J Gambl Stud* 14 (1998): 347–358. <https://doi.org/10.1023/A:1023021009398>
- Pearce, J. Mason, K. Hiscock, R. Day, P. “A national study of neighbourhood access to gambling opportunities and individual gambling behaviour.” *Journal of Epidemiology & Community Health* 62 (2008): 849. <https://doi.org/10.1136/jech.2007.068114>
- “American Online Gaming.” Casino City. Accessed Oct 4, 2024. <https://us.casinocity.com/online-gaming/>.
- “State by State Betting in the USA.” BettingUSA. Last modified Aug 18, 2024. Accessed Oct 4, 2024. <https://www.bettingusa.com/states/>.
- Liptak, Adam. Draper, Kevin. “Supreme Court Ruling Favors Sports Betting.” New York Times. May 14, 2018. <https://www.nytimes.com/2018/05/14/us/politics/supreme-court-sports-betting-new-jersey.html>.
- “Murphy v. National Collegiate Athletic Association.” Wikipedia. Accessed Aug 7, 2024. https://en.wikipedia.org/wiki/Murphy_v._National_Collegiate_Athletic_Association.
- “Interactive U.S. Map: Sports Betting.” American Gaming Association. Accessed Oct 4, 2024. <https://www.americangaming.org/research/state-gaming-map/>.
- “American Attitudes Towards Gaming 2024.” American Gaming Association. Aug 28, 2024. <https://www.americangaming.org/resources/american-attitudes-towards-gaming-2024>.
- “United States Commercial Casino Gaming: Monthly Revenues.” University of Nevada, Las Vegas Center for Gaming Research. pp 1-43. Sept 2024. https://gaming.library.unlv.edu/reports/national_monthly.pdf. Accessed Aug 7, 2024.
- “Gini Index of Income Inequality.” United States Census Bureau. Accessed Aug 5, 2024. <https://data.census.gov/table/ACSDT1Y2023.B19083?q=Gini%20index>.
- Horowitz, Juliana Menasce. Igielnik, Ruth. Kochhar, Rakesh. “Trends in income and wealth inequality.” Pew Research Center. Jan 9, 2020. <https://www.pewresearch.org/social-trends/2020/01/09/trends-in-income-and-wealth-inequality/>.
- “Behavioral Risk Factor Surveillance System (BRFSS).” North Carolina State Center for Health Statistics. Accessed Aug 8, 2024. <https://schs.dph.ncdhhs.gov/data/brfss/2020/nc/risk/topics.htm#ga>.
- “Enterprise Data Dissemination Informatics Exchange.” Pennsylvania Department of Health. Accessed Aug 9, 2024. <https://www.phaim1.health.pa.gov/EDD/WebForms/BRFSSstate.aspx>.
- Williams, Robert J. Volberg, Rachel A.. Zorn, Martha. Stanek, Edward J. Evans, Valerie. “Six Year Longitudinal Study of Gambling and Problem Gambling in Massachusetts.” pp

42. April 16, 2021.
https://massgaming.com/wp-content/uploads/MAGIC-Six-Year-Longitudinal-Study-of-Gambling-and-Problem-Gambling-in-Massachusetts_Report-4.16.21.pdf.
- “United States.” MapChart. Accessed Sept 8, 2024. <https://www.mapchart.net/usa.html>.
- “Educational Attainment.” United States Census Bureau. Accessed Sept 12, 2024.
<https://data.census.gov/table/ACSST5Y2020.S1501?q=educational%20attainment&g=040XX00US08,10,12,17,18,19,20,22,23,24,26,28,29,32,34,35,36,39,40,42,44,46,54&moe=false>.
- “Labor Force Statistics from the Current Population Survey: Unemployment Rate.” Bureau of Labor Statistics, U.S. Department of Labor. Accessed Sept 5, 2024.
<https://data.bls.gov/toppicks?survey=la>.
- “Federal Reserve Economic Data.” Federal Reserve Bank of St. Louis. Accessed Sept 8, 2024. <https://fred.stlouisfed.org/release/tables?rid=118&eid=259194#snid=259200>.
- “GDP by State.” Bureau of Economic Analysis, U.S. Department of Commerce. Accessed Sept 13, 2024. <https://www.bea.gov/data/gdp/gdp-state>.
- “National Highway Traffic Safety Administration.” United States Department of Transportation. Accessed Sept 15, 2024. <https://cdan.dot.gov/STSI/stsi.htm#>.
- T. Maher, H. Palattella. Pennsylvania Sports Betting – Is Sports Betting Legal In Pennsylvania? *Forbes*, updated Jul 25, 2024,
<https://www.forbes.com/betting/legal/is-sports-betting-legal-in-pennsylvania/>.
Accessed Aug 20, 2024
- “Gambling in North Carolina.” Wikipedia. Accessed Aug 8, 2024.
https://en.wikipedia.org/wiki/Gambling_in_North_Carolina.
- “Gambling in Massachusetts.” Wikipedia. Accessed Aug 8, 2024.
https://en.wikipedia.org/wiki/Gambling_in_Massachusetts.
- “Directors.” NCALHD. Accessed Sept 9, 2024. <https://ncalhd.org/directors/>.
- Maher, Tyler. Pempus, Brian. “Colorado Sports Betting – Is Sports Betting Legal In Colorado?” *Forbes*. Aug 28, 2024.
<https://www.forbes.com/betting/legal/is-sports-betting-legal-in-colorado>.
- Maher, Tyler. Palattella, Henry. “Delaware Sports Betting – Is Sports Betting Legal In Delaware?” *Forbes*. Last modified Jul 25, 2024. Accessed Aug 20, 2024.
<https://www.forbes.com/betting/legal/is-sports-betting-legal-in-delaware>.
- Maher, Tyler. Pempus, Brian. “Florida Sports Betting – Is Sports Betting Legal In Florida?” *Forbes*. Jul 1, 2024.
<https://www.forbes.com/betting/legal/is-sports-betting-legal-in-florida>.
- Maher, Tyler. Palattella, Henry. “Illinois Sports Betting – Is Sports Betting Legal In Illinois?” *Forbes*. Sep 3, 2024.
<https://www.forbes.com/betting/legal/is-sports-betting-legal-in-illinois>.

Maher, Tyler. Palattella, Henry. "Indiana Sports Betting – Is Sports Betting Legal In Indiana?" Forbes. Last modified Jul 25, 2024. Accessed Aug 20, 2024. <https://www.forbes.com/betting/legal/is-sports-betting-legal-in-indiana>.

Palattella, Henry. "Iowa Sports Betting – Is Sports Betting Legal In Iowa?" Forbes. Aug 16, 2024. <https://www.forbes.com/betting/legal/is-sports-betting-legal-in-iowa>.

Maher, Tyler. Palattella, Henry. "Kansas Sports Betting – Is Sports Betting Legal In Kansas?" Forbes. Last modified Jul 25, 2024. Accessed Aug 20, 2024. <https://www.forbes.com/betting/legal/is-sports-betting-legal-in-kansas>.

Maher, Tyler. Palattella, Henry. "Louisiana (LA) Sportsbooks: Best Sports Betting Sites & Apps – September 2024." Forbes. Oct 2, 2024. <https://www.forbes.com/betting/legal/louisiana-sports-betting-sites>.

Maher, Tyler. Palattella, Henry. "Maine Sports Betting – Is Sports Betting Legal In Maine?" Forbes. Last modified Jul 25, 2024. Accessed Aug 20, 2024. <https://www.forbes.com/betting/legal/is-sports-betting-legal-in-maine>.

Maher, Tyler. "Maryland Sports Betting – Is Sports Betting Legal In Maryland?" Forbes. May 16, 2024. <https://www.forbes.com/betting/legal/is-sports-betting-legal-in-maryland>.

Maher, Tyler. Palattella, Henry. "Michigan Sports Betting – Is Sports Betting Legal In Michigan?" Forbes. Last modified Jul 25, 2024. Accessed Aug 20, 2024. <https://www.forbes.com/betting/legal/is-sports-betting-legal-in-michigan>.

Maher, Tyler. Palattella, Henry. "Is Sports Betting Legal In Mississippi (MS)?" Forbes. Aug 14, 2024. <https://www.forbes.com/betting/legal/is-sports-betting-legal-in-mississippi>.

Maher, Tyler. Palattella, Henry. "Missouri Sports Betting – Is Sports Betting Legal In Missouri?" Forbes. Last modified Jul 25, 2024. Accessed Aug 20, 2024. <https://www.forbes.com/betting/legal/is-sports-betting-legal-in-missouri>.

Maher, Tyler. Palattella, Henry. "Nevada Sports Betting – Is Sports Betting Legal In Nevada?" Forbes. Last modified Jul 25, 2024. Accessed Aug 20, 2024. <https://www.forbes.com/betting/legal/is-sports-betting-legal-in-nevada>.

Palattella, Henry. "New Jersey Sports Betting – Is Sports Betting Legal In New Jersey?" Forbes. Last modified Jul 25, 2024. Accessed Aug 20, 2024. <https://www.forbes.com/betting/legal/is-sports-betting-legal-in-new-jersey>.

"New Mexico Sports Betting Updates." Online Gambling. Accessed Aug 20, 2024. <https://www.onlinegambling.com/states/nm/sports-betting/>.

Beare, Nick. "Top 9 NY Sportsbooks – New York Sports Betting 2024." LSR. Last modified Oct 8, 2024. Accessed Oct 6, 2024. <https://www.legalsportsreport.com/ny>.

Henry, Megan. "Ohio placed more than \$7.65 billion in sports bets in 2023." Ohio Capital Journal. Feb 6, 2024. <https://ohiocapitaljournal.com/2024/02/06/ohio-placed-more-than-7-65-billion-in-sports-bet-in-2023>.

- Maher, Tyler. Palattella, Henry. "Oklahoma Sports Betting – Is Sports Betting Legal In Oklahoma?" *Forbes*. May 10, 2024.
<https://www.forbes.com/betting/legal/is-sports-betting-legal-in-oklahoma/>.
- Maher, Tyler. Palattella, Henry. "Rhode Island (RI) Sportsbooks: Best Sports Betting Sites & Apps – September 2024." *Forbes*. Last modified Oct 7, 2024. Accessed Aug 20, 2024. <https://www.forbes.com/betting/legal/rhode-island-sports-betting-sites>.
- Maher, Tyler. Palattella, Henry. "South Dakota (SD) Sportsbooks: Best Sports Betting Sites & Apps – September 2024." *Forbes*. Last modified Oct 7, 2024. Accessed Aug 20, 2024. <https://www.forbes.com/betting/legal/south-dakota-sports-betting-sites>.
- Maher, Tyler. Palattella, Henry. "West Virginia (WV) Sportsbooks: Best Sports Betting Sites & Apps – September 2024." *Forbes*. Oct 1, 2024.
<https://www.forbes.com/betting/legal/west-virginia-sports-betting-sites>.
- Torres-Reyna, Oscar. "Differences-in-Differences." Princeton University. pp 4. August 2015.
<https://www.princeton.edu/~otorres/DID101.pdf>
- "Instrumental Variable: Definition & Overview." *Statistics How To*. Accessed Oct 5, 2024.
<https://www.statisticshowto.com/instrumental-variable>.
- Cobb, Sam. "How Nebraska's legalization of casinos has affected the state." *Nebraska News Service*. May 15, 2024.
https://www.nebraskanewsservice.net/news/how-nebraska-s-legalization-of-casinos-has-affected-the-state/article_cb848496-127d-11ef-bac2-d38579699bc7.html.
- Can, Ege. Nichols, Mark W. "The effects of sports betting on cross-border substitution in casino gambling." *Public Finance Review* 52, no. 5 (2024): 563-609.
<https://doi.org/10.1177/10911421241248506>.
- Chu, May. "STATISTICAL BRIEF #62: Characteristics of Persons Who Seldom or Never Wear Seat Belts, 2002." Agency for Healthcare Research and Quality. Dec 2004.
https://meps.ahrq.gov/data_files/publications/st62/stat62.shtml.
- "Gini Index coefficient - distribution of family income." *The World Factbook*. Accessed Oct 5, 2024.
<https://www.cia.gov/the-world-factbook/field/gini-index-coefficient-distribution-of-family-income/country-comparison/>.
- Miller, Monica. "Japan approves plan to open its first casino." *BBC*. Apr 14, 2023.
<https://www.bbc.com/news/business-65271857>.

Livestock Ponds and Their Significance to Avian Communities By Rayyan Ibrahim

Abstract

Livestock ponds are a significant resource for many species across California's rangeland. The management of the ponds can determine the degree to which they benefit native wildlife and which species derive the most benefit from them. Birds can use these ponds for many crucial activities, including feeding, drinking, and bathing. The presence of man-made ponds may alter their behavior and foraging habits while also possibly allowing for the existence of habitat beneficial to certain imperiled species. This introductory study aims to assess how manipulatable environmental factors affect the species diversity of avian communities using the ponds through data collected through an observational study of two livestock ponds within the Sunol Regional Wilderness. I also aimed to see if there was any link between livestock ponds and a potential disproportionate advantage they may offer to non-native bird species, most notably the European Starling (*Sturnus vulgaris*). The results show a lower species richness of the bird community at the pond closer to a hiking trail. The proportion of insectivorous birds is notably higher at the more secluded pond site as compared to the more exposed site, likely due to a combination of several factors, including human activity, vegetation, and the presence of Western Pond Turtles (*Actinemys marmorata*) in one of the ponds. Starlings, having only been observed once throughout the study, did not seem to be disproportionately benefiting from the existence of the ponds.

Introduction

The sale of domestic cattle (*Bos taurus*) resulted in a cash income of 3.63 billion dollars in California, placing cattle as the state's third most important agricultural commodity (2022-2023 California Agricultural Statistics Review, n.d.). Of California's 100 million acres of land, 38 million are classified as range and pasture land (*Ranching in California – California Cattlemen's Association*, n.d.). Cattle grazing is a contentious issue, often attributed to treading damage of foliage and the high enteric CH₄ emissions associated with livestock production (Franzluebbers, 2020; Sheath & Boom, 1997). Moderate-to-light grazing can shape grasslands, altering the relative abundances of species, both alien and native, in an ecosystem (Hayes & Holl, 2003; Kimball & Schiffman, 2003; Michaels et al., 2021). Grazed lands showcase increases in abundance and species richness of native forbs, likely due to the decreased light competition caused by the reduced height of perennial grasses, which are fed on by ruminants (Hayes & Holl, 2003). Plant variability and native plant abundance are increased through moderate grazing (Gennet et al., 2017; Michaels et al., 2021) due to the reduction in competition from dominant grass species and the creation of bare patches, which facilitate the growth of native species (Michaels et al., 2021). The reduction of litter as a result of cattle grazing can form habitat that is beneficial for some bird species, such as the Horned Lark (*Eremophila alpestris*) (Gennet et al., 2017), but can damage other wildlife populations (Jones, 2000). Areas subject to moderate grazing in Central California showed a higher abundance of Western Fence Lizards (*Sceloporus*

occidentalis) and Western Skink (*Plestiodon skiltonianus*) (Riensch, 2005). This could have further implications on the populations of the Alameda Whipsnake (*Masticophis lateralis euryxanthus*) (Riensch, 2005), a threatened species that feeds primarily on lizard species, including the Western Fence Lizard (Swaim, 1994). Grazing can also have spatiotemporal effects on the activity of large mammals such as mule deer, causing more nocturnal behaviors (Caldwell & Klip, 2023). Across grazed land, cattle ponds serve as a habitat and resource for wildlife (Swartz & Miller, 2021). Cattle ponds generally lack emergent vegetation and tend to have high nitrogen and phosphorus content (Knutson et al., 2004). Due to the lack of fish in most cattle ponds (Knutson et al., 2004), the ponds can sometimes sustain significant amphibian populations (Knutson et al., 2004; Swartz & Miller, 2019). Despite the adverse effects that cattle can have on amphibian survival (Schmutzer et al., 2008), Central California cattle ponds serve as vital breeding habitat for endangered amphibians such as the California Red-legged Frog (*Rana draytonii*) and California Tiger Salamander (*Ambystoma californiense*) (Ford et al., 2013). Cattle ponds are especially beneficial to the California Tiger Salamander due to the creation of preferential habitat (Ford et al., 2013) and the ability of the California Tiger Salamander to adapt its diet to one more suitable for the available prey in cattle ponds (Manenti et al., 2017). Birds also benefit from the habitat provided by cattle ponds, though ponds at different successional stages harbor drastically different bird diversity (Zamora-Marín et al., 2021). Larger ponds can even harbor water birds, such as waterfowl and shorebirds (Evans & Kerbs, 1977).

Cattle ponds throughout Central California may be an essential asset to preserving biodiversity in anthropogenically managed rangeland, balancing the economic benefits that cattle bring (2022-2023 *California Agricultural Statistics Review*, n.d.) with the protection of species using the ponds as a resource. Central Valley and its surrounding foothills experience a Mediterranean climate with an unpredictable water supply (Byrd et al., 2015), so ponds are a valuable resource to all wildlife. The impacts of cattle ponds on Central California's amphibians have been well studied (Doubledee et al., 2003; Ford et al., 2013; Messerman et al., 2021), but the use of cattle ponds by other wildlife in Central California's rangeland has been seemingly ignored. Globally, cattle ponds serve as areas of high diversity and abundance for aquatic invertebrates, especially dragonflies (Manenti et al., 2017; Paz et al., 2020), possibly due to the lack of fish to serve as predators. Locally, such an abundance of insects would likely serve as a food source for many native insectivorous birds, such as the Black Phoebe (*Sayornis nigricans*), Western Kingbirds (*Tyrannus verticalis*), and swallows (*Hirundinidae*) (Blancher & Robertson, 1987; Bumelis et al., 2022; Freehling & Johnson, 2018). Wild mammals have also been observed to congregate around artificial water sources (Destefano et al., 2000). While some suggest that increased wildlife concentration may produce a higher risk of predation, there is little evidence to support the notion that artificial water sources cause a substantially higher loss of prey species to predation at said water sources (Destefano et al., 2000; Simpson et al., 2011). There is little information regarding the ponds causing the disproportionate propagation of non-native bird species, particularly European Starlings (*Sturnus vulgaris*), as they tend to associate with cattle (Källander, 2004). European Starlings have been observed consuming livestock feed in feedlots

(*Starlings a Costly Problem for Livestock Producers*, n.d.) and developing a commensalistic relationship with cattle in which the starlings feed both on the insects that are attracted to the cow manure and those that are flushed out of the grass by the cattle (Källander, 2004). While the magnitude of damage caused by the starlings on native bird populations is contested (Koch et al., 2012; Koenig, 2003), starlings cause significant agricultural damage and can be a disease vector to livestock and humans (Linz et al., 2007).

Further research needs to be done to assess the conservation value that cattle ponds serve for non-amphibian vertebrates, particularly birds, and whether or not the level of activity that the ponds receive, both by humans and cattle, affects the species distribution of avian life at the ponds. This paper assesses bird and small mammal diversity at cattle ponds with varying levels of both cattle and human traffic. Such research aims to see if cattle ponds may be a valuable resource to preserving rangeland biodiversity and suggest any potential changes that may improve the quality of such ponds as resources for native wildlife.

Methods

Data collection occurred in cattle ponds at the Sunol Regional Wilderness near Sunol, California. The park is managed by the East Bay Regional Park District, which also manages 72 other parks in the region. The habitat of the park can generally be described as grassland with dispersed thickets of oak. Vegetation was predominantly characterized by non-native grasses such as barley (*Hordeum sp.*), Wild Oat (*Avena fatua*), Slender Wild Oat (*Avena barbata*), Common Soft Brome (*Bromus hordeaceus*), Ripgut Brome (*Bromus diandrus*), and Compact Brome (*Bromus madritensis*). Native grasses such as California Melic (*Melica californica*), Big Squirreltail (*Elymus multisetus*), and Pine Bluegrass (*Poa secunda*) were also present on the landscape. The park is home to a vast diversity of birds, including California Quail (*Callipepla californica*), Western Meadowlark (*Sturnella neglecta*), Yellow-billed Magpie (*Pica nuttalli*), and Western Bluebird (*Sialia mexicana*). The park is also home to a large population of European Starlings (*Sturnus vulgaris*). The park is grazed by cattle, which drink from a system of cattle ponds and troughs.

Two ponds were selected for data collection, one which faced virtually no human traffic (over 150 meters from the nearest hiking trail) and one which was around 5 meters from a moderately popular hiking trail. This paper will refer to these ponds as Pond A and Pond B, respectively. Pond A was more sheltered by grassy hills on both sides, was smaller, and contained limited emergent vegetation on one side of the pond. The perimeter of the pond was extremely uneven due to a large amount of hoof imprints in the mud. More cattle traffic was observed at Pond A. Pond B also had hoof prints but to a lesser extent. Pond B had a large amount of floating vegetation but no apparent emergent vegetation. Both ponds supported a large amount of invertebrate life. Dragonflies were most commonly sighted around the ponds, with the predominant species present being Common Whitetail (*Plathemis lydia*), Cardinal Meadowhawk (*Sympetrum illotum*), and Flame Skimmer (*Libellula saturata*). Common Green Darner (*Anax junius*) and Variegated Meadowhawk (*Sympetrum corruptum*) were also sighted, albeit less than

the previously mentioned species. Pond B was also notable for a substantial population of Western Pond Turtle (*Actinemys marmorata*).

Data collection was focused on the use of the ponds by birds and small mammals. Due to the elusive nature of many small rodents, the primary focus of the data collection with respect to small mammals was the California Ground Squirrel (*Otospermophilus beecheyi*). Both ponds were observed in hour-long intervals on randomly selected days throughout June and July. Ponds were surveyed either from 9:30 AM-10:30 AM or 11:00 AM-12:00 PM. The ponds alternated between the two time blocks. The species diversity and relative abundance of birds and squirrels at the ponds were collected through visual survey from a site within 50 meters of the pond. The observer sat at a visually obscured vantage point with a full view of the pond at least 25 meters from the edge of the water. Observations were recorded as use of the pond site if recorded within 5 meters of the edge of the waterline at either pond. Behavioral data was collected on instances of feeding, drinking, and bathing at the ponds. Binoculars and a camera were used to reliably identify the birds being studied.

Results

12 species of bird and 1 rodent species were observed directly as they were using the cattle ponds (within 5 meters of the pond). The following species were observed at the ponds: Western Bluebird (*Sialia mexicana*), Black Phoebe (*Sayornis nigricans*), Western Kingbird (*Tyrannis verticalis*), Mourning Dove (*Zenaida macroura*), California Quail (*Callipepla californica*), Lesser Goldfinch (*Spinus psaltria*), Lark Sparrow (*Chondestes grammacus*), Lazuli Bunting (*Passerina amoena*), Western Meadowlark (*Sturnella neglecta*), Yellow-billed Magpie (*Pica nuttalli*), Violet-Green Swallow (*Tachycineta thalassina*), European Starling (*Sturnus vulgaris*), and California Ground Squirrel (*Otospermophilus beecheyi*).

Species Name	Diet	Behaviors Observed	Ponds Observed
Black Phoebe (<i>Sayornis nigricans</i>)	Insectivorous (Gall & Fernández-Juricic, 2009)	F, D	A, B
Western Bluebird (<i>Sialia mexicana</i>)	Insectivorous (Herlugson, 1982)	F, D	A, B
Western Kingbird (<i>Tyrannis verticalis</i>)	Insectivorous (Blancher & Robertson, 1987)	F	A
Mourning Dove (<i>Zenaida macroura</i>)	Granivorous (Hayslette & Mirarchi, 2001)	F, D	A, B
California Quail (<i>Callipepla californica</i>)	Granivorous (Glading et al., 1940)	D	B
Lesser Goldfinch (<i>Spinus psaltria</i>)	Granivorous (Coutlee, 1968)	D, B	A

Lark Sparrow (<i>Chondestes grammacus</i>)	Insectivorous (Augustine & Baker, 2013)	F	A
Lazuli Bunting (<i>Passerina amoena</i>)	Omnivorous (Wheelock, 1912)	D	A
Western Meadowlark (<i>Sturnella neglecta</i>)	Omnivorous (Bryant, 1913)	N	A
Yellow-billed Magpie (<i>Pica nuttalli</i>)	Omnivorous (Parker et al., 2001)	F	A
Violet-Green Swallow (<i>Tachycineta thalassina</i>)	Insectivorous (Garlick et al., 2014)	F, D	A, B
European Starling (<i>Sturnus vulgaris</i>)	Omnivorous* (Fischl & Caccamise, 1987)	F	B

* indicates non-native

F = Feeding, D = Drinking, B = Bathing, N = No notable behaviors observed

TABLE 1.

Caption: Table 1. The table displays the complete list of birds observed at both pond sites. The dietary habits (insectivorous, granivorous, or omnivorous) are displayed alongside the behaviors observed at the ponds and which sites they were observed to have used. Feeding was defined as engaging in the act of foraging or hunting, whether successful or not.

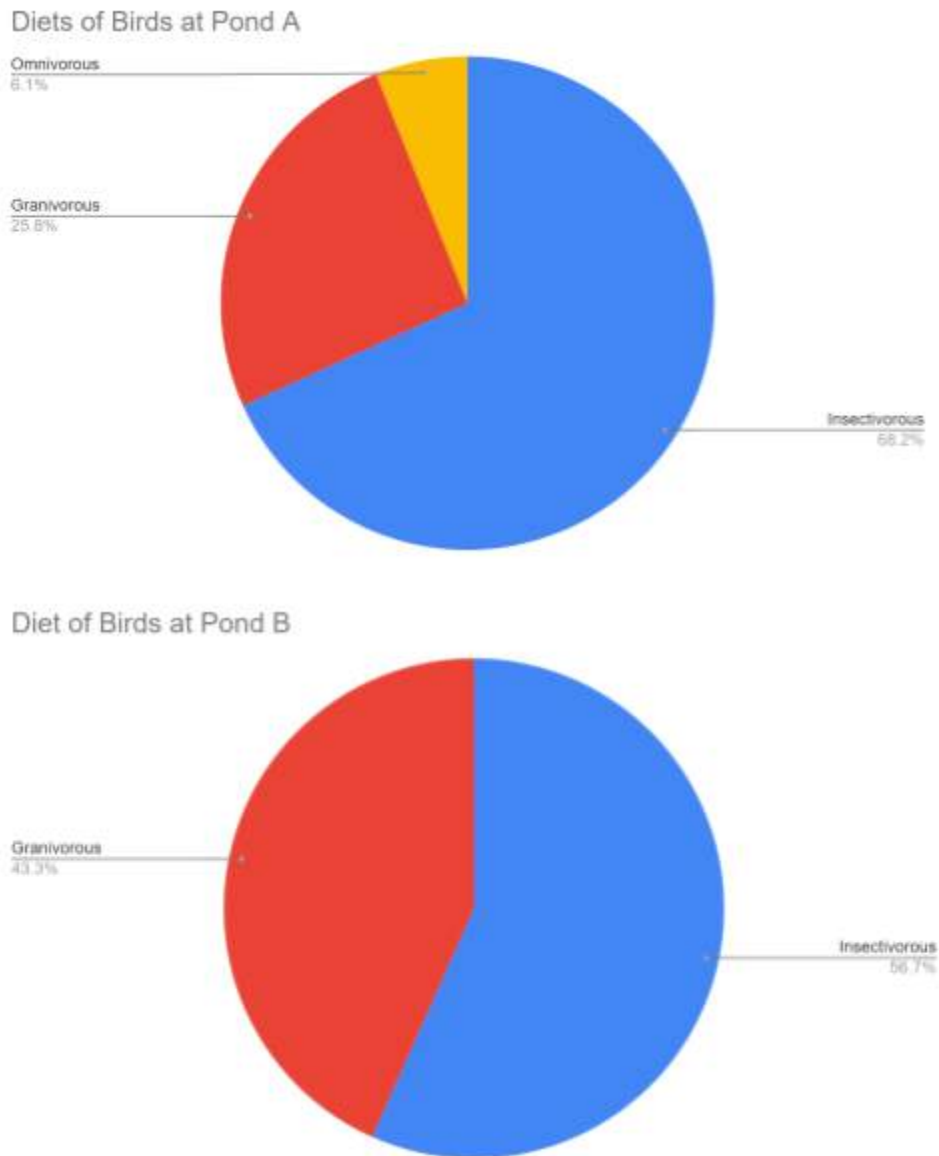


FIGURE 1

Caption: Figure 1. The two pie charts compare the usage of the two pond sites by birds of different feeding classes. European Starlings were omitted from this data as there was only one observation of them, and the count of starlings in the flock was not exact.

Although the two livestock pond sites were within 3 kilometers of each other, the bird activity was substantially different. Pond A had a species richness of 9, species evenness of 0.668, and a calculated Shannon Diversity Index of 1.47 for the observed bird population. 50.75% of all sightings were of Western Bluebirds (*Sialia mexicana*), of which all cases reported signs of feeding/foraging behavior. Of the granivorous birds, 88.24% were Mourning Doves

(*Zenaida macroura*), and all other sightings were Lesser Goldfinch (*Spinus psaltria*). All birds sighted at Pond A were native, although a lone European Starling (*Sturnus vulgaris*) was sighted in a nearby cattle trough.

Pond B had a species richness of 5, species evenness of 0.847, and a calculated Shannon Diversity Index of 1.36. The most populous bird at Pond B was the Mourning Dove (*Zenaida macroura*), accounting for 35.48% of all bird sightings and 84.62% of all granivorous bird sightings. 63.64% of all Mourning Dove (*Zenaida macroura*) sightings at Pond B involved drinking. Pond A had 0 sightings of Mourning Dove (*Zenaida macroura*) drinking. All doves sighted at Pond A were either foraging or resting. Of the insectivorous population at Pond B, 47.06% were Black Phoebe (*Sayornis nigricans*), and 52.94% were Violet-Green Swallows (*Tachycineta thalassina*). Of the Violet-Green Sparrow (*Tachycineta thalassina*) sightings, 77.78% involved drinking, and 55.56% involved feeding. Of the Black Phoebe (*Sayornis nigricans*) sightings, 25.00% involved drinking, and 87.50% involved feeding. European Starlings (*Sturnus vulgaris*) were again left out of this calculation due to inexactness in the count of the flock. The flock was sighted once at Pond B and consisted of approximately 30 members. Out of all bird specimens, 2.04% were unable to be identified down to the species level by the observer due to the bird being obscured or flying away before an identification was made.

California Ground Squirrels (*Otospermophilus beecheyi*) were sighted at both ponds, making up 6.94% of observations at Pond A and 11.43% of observations at Pond B. 88.89% of all ground squirrel observations reported actively drinking from the water line. Ground squirrel activity was reported from the vicinity (100 meters) of both ponds. Both ponds had ground squirrels actively foraging in the vicinity, which reacted negatively to human presence (running away and making alarm calls) in response to both the observer and nearby hikers. Both juvenile and adult ground squirrels were observed actively drinking at both ponds. The ground squirrels showed no apparent preference for drinking in groups, as both solitary and group drinking were observed at both pond sites.

The only living mammals other than ground squirrels observed at the sight were cattle, which were only directly observed at Pond A. Evidence of cattle was observed at Pond B through the presence of hoofprints and feces. Evidence of other mammal activity at the ponds included a Mule Deer carcass found at the edge of Pond B. Other vertebrate species observed at the ponds were California Red-legged Frogs (*Rana draytonii*; A, B), Pacific Chorus Frogs (*Pseudacris regilla*; A only), Western Pond Turtles (*Actinemys marmorata*; B only), Diablo Range Garter Snakes (*Thamnophis atratus*; A only), and Western Fence Lizards (*Sceloporus occidentalis*; A, B).

Discussion

Human-modified environments often allow for the increased spread of invasive species, with cattle ponds explicitly being a part of the spread of the American Bullfrog into less suitable habitat (Nelson & Piovia-Scott, 2022). The American Bullfrog has been a significant contributor to the decline of native anurans such as the Foothill Yellow-legged Frog (*Rana boylei*) (Moyle,

1973). While neither of the ponds showed evidence of American Bullfrogs, Pond B was actively used by European Starlings (*Sturnus vulgaris*), another invasive species to California. Nothing in the data suggests that the cattle ponds were being disproportionately used by starlings. No starlings visited Pond A despite them being seen in the vicinity. While European Starlings (*Sturnus vulgaris*), like many invasive birds, thrive in human-modified environments (Ibañez et al., 2023), more evidence is needed to evaluate the extent to which large, permanent water sources such as cattle ponds facilitate their spread and propagation.

Pond A's higher diversity can likely be attributed partially to its seclusion and the lack of human activity at the pond. Species less tolerant of human activity would be less frequent visitors to a pond next to an active trail. Certain members of the Columbidae family have been known to become habituated to human presence (Gendall et al., 2015), potentially explaining the proportion of Mourning Doves (*Zenaida macroura*) present at Pond B. Pond A also had a higher proportion of insectivorous birds. This may be attributed to the presence of Western Pond Turtles (*Actinemys marmorata*) in Pond B. Western Pond Turtles feed on larval invertebrates (Hays et al., 1999), possibly resulting in a lower concentration of flying insects at Pond B than at Pond A. This could also be attributed to differences in habitat, with Pond A being more shielded and containing emergent vegetation.

The methods of this experiment could be improved with more effective concealment of the observer with the use of a hide. Birds less tolerant of human activity would have been less likely to visit the ponds in the presence of a visible researcher. Another potential flaw in the experimentation was that certain birds could have been obscured from the view of the observer. Birds with cryptic coloration would be less likely to be spotted than those with vibrant colors. Birds closer to the observer were also more likely to be spotted, allowing for potential bias in the observer disproportionately recording species that are more tolerant of humans and, therefore, more willing to be near the observer. While the bias was limited by having the observer sit at a vantage point where the entire pond was visible, this possibility for bias will always occur in the case of visual surveys for bird populations. There must be a balance between being close enough to spot birds reliably and far enough not to disturb their behaviors. Other methods that could be considered for future research would be camera trapping, mist netting, or audio survey.

The sample size of this research was relatively small, only focusing on two cattle ponds in the Sunol Regional Wilderness. A broader study on this topic involving more study sites is necessary for direct conclusions to be drawn between specific environmental factors and their effects on avian diversity at the pond sites. The current data is insufficient to make educated decisions on pond management for the stewardship of healthy, diverse avian populations in the California rangeland.

Conclusion

This research was an introductory study into the relevance of cattle ponds to avian population management. Cattle ponds are large, foreign bodies of water that may come to the benefit of certain species and the detriment of others. Even though this study only focused on

two ponds, there were noticeable differences in the observational data gathered at both sites. The diversity of avian communities at the two pond sites was noticeably different, likely due to the variance in environmental factors. This could have implications for how we manage livestock ponds in the future.

Efforts to manage cattle ponds in a way that balances amphibian populations and rancher needs are already underway. Yet, amphibians are not the only creatures affected by these foreign bodies of water dotting California rangeland. Broader studies on the use of these ponds by birds can result in a better understanding of how native wildlife uses these ponds and how we can selectively manage the ponds to maximize beneficial resources available to threatened species across a rapidly changing rangeland.

Works Cited

- 2022-2023 California Agricultural Statistics Review. (n.d.).
- Augustine, D. J., & Baker, B. W. (2013). Associations of Grassland Bird Communities with Black-Tailed Prairie Dogs in the North American Great Plains. *Conservation Biology*, 27(2), 324–334. <https://doi.org/10.1111/cobi.12013>
- Blancher, P. J., & Robertson, R. J. (1987). Effect of Food Supply on the Breeding Biology of Western Kingbirds. *Ecology*, 68(3), 723–732. <https://doi.org/10.2307/1938478>
- Bryant, H. C. (1913). The Economic Value of the Western Meadowlark in California. Richardson.
- Bumelis, K. H., Cadman, M. D., & Hobson, K. A. (2022). Endogenous biomarkers reveal diet partitioning among three sympatric species of swallows. *Ornithology*, 139(1), ukab078. <https://doi.org/10.1093/ornithology/ukab078>
- Byrd, K. B., Flint, L. E., Alvarez, P., Casey, C. F., Sleeter, B. M., Soulard, C. E., Flint, A. L., & Sohl, T. L. (2015). Integrated climate and land use change scenarios for California rangeland ecosystem services: Wildlife habitat, soil carbon, and water supply. *Landscape Ecology*, 30(4), 729–750. <https://doi.org/10.1007/s10980-015-0159-7>
- Caldwell, M. R., & Klip, J. M. K. (2023). Influence of Cattle Grazing on the Spatiotemporal Activity of Wildlife. *Rangeland Ecology & Management*, 88, 122–128. <https://doi.org/10.1016/j.rama.2023.02.011>
- Coutlee, E. L. (1968). Maintenance Behavior of Lesser and Lawrence's Goldfinches. *The Condor*, 70(4), 378–384. <https://doi.org/10.2307/1365933>
- Destefano, S., Schmidt, S. L., & deVos, J. C. (2000). Observations of Predator Activity at Wildlife Water Developments in Southern Arizona. *Journal of Range Management*, 53(3), 255. <https://doi.org/10.2307/4003428>
- Doubledee, R. A., Muller, E. B., & Nisbet, R. M. (2003). Bullfrogs, Disturbance Regimes, and the Persistence of California Red-Legged Frogs. *The Journal of Wildlife Management*, 67(2), 424–438. <https://doi.org/10.2307/3802783>
- Evans, K. E., & Kerbs, R. R. (1977). Avian use of livestock watering ponds in Western South Dakota. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station.
- Fischl, J., & Caccamise, D. F. (1987). Relationships of Diet and Roosting Behavior in the European Starling. *The American Midland Naturalist*, 117(2), 395–404. <https://doi.org/10.2307/2425982>
- Ford, L. D., Hoorn, P. A. V., Rao, D. R., Scott, N. J., Trenham, P. C., & Bartolome, J. W. (2013). California Tiger Salamanders.
- Franzuebbers, A. J. (2020). Chapter 2 - Cattle grazing effects on the environment: Greenhouse gas emissions and carbon footprint. In M. Rouquette & G. E. Aiken (Eds.), *Management Strategies for Sustainable Cattle Production in Southern Pastures* (pp. 11–34). Academic Press. <https://doi.org/10.1016/B978-0-12-814474-9.00002-5>
- Freehling, M., & Johnson, K. (2018). Black Phoebe (*Sayornis nigricans*) Predation on Variegated

- Meadowhawk (*Sympetrum corruptum*). 30, 17–18.
- Gall, M. D., & Fernández-Juricic, E. (2009). Effects of Physical and Visual Access to Prey on Patch Selection and Food Search Effort in a Sit-and-Wait Predator, the Black Phoebe. *The Condor*, 111(1), 150–158. <https://doi.org/10.1525/cond.2009.080016>
- Garlick, N. W., Newberry, G. N., & Rivers, J. W. (2014). An Assessment of Nestling Diet Composition in the Violet-Green Swallow (*Tachycineta thalassina*). <https://bioone.org/journals/northwest-science/volume-88/issue-1/046.088.0109/An-Assessment-of-Nestling-Diet-Composition-in-the-Violet-Green/10.3955/046.088.0109.full>
- Gendall, J., Lill, A., & Beckman, J. (2015). Tolerance of disturbance by humans in long-time resident and recent colonist urban doves. *Avian Research*, 6(1), 7. <https://doi.org/10.1186/s40657-015-0018-x>
- Gennet, S., Spotswood, E., Hammond, M., & Bartolome, J. W. (2017). Livestock grazing supports native plants and songbirds in a California annual grassland. *PLOS ONE*, 12(6), e0176367. <https://doi.org/10.1371/journal.pone.0176367>
- Glading, B., Biswell, H. H., & Smith, C. F. (1940). Studies on the Food of the California Quail in 1937. *The Journal of Wildlife Management*, 4(2), 128–144. <https://doi.org/10.2307/3795650>
- Hayes, G. F., & Holl, K. D. (2003). Cattle Grazing Impacts on Annual Forbs and Vegetation Composition of Mesic Grasslands in California. *Conservation Biology*, 17(6), 1694–1702. <https://doi.org/10.1111/j.1523-1739.2003.00281.x>
- Hays, D., McAllister, K., Richardson, S., & Stinson, D. (1999). Washington State Recovery Plan for the Western Pond Turtle.
- Hayslette, S. E., & Mirarchi, R. E. (2001). Patterns of Food Preferences in Mourning Doves. *The Journal of Wildlife Management*, 65(4), 816–827. <https://doi.org/10.2307/3803031>
- Herlugson, C. J. (1982). Food of Adult and Nestling Western and Mountain Bluebirds. *The Murrelet*, 63(2), 59–65. <https://doi.org/10.2307/3533829>
- Ibañez, L., Palacio, F. X., Maragliano, R. E., & Montalti, D. (2023). The presence of an invasive bird, the Common Starling, in an urban landscape: Habitat use and relationships with other bird species. *Journal of Ornithology*, 164(3), 537–546. <https://doi.org/10.1007/s10336-023-02047-x>
- Jones, A. (2000). Effects of Cattle Grazing on North American Arid Ecosystems: A Quantitative Review. *Western North American Naturalist*, 60(2), 155–164.
- Källander, H. (2004). Starlings *Sturnus vulgaris* and cattle—A widespread feeding association. *Ornis Svecica*, 14(1–2), Article 1–2. <https://doi.org/10.34080/os.v14.22795>
- Kimball, S., & Schiffman, P. M. (2003). Differing Effects of Cattle Grazing on Native and Alien Plants. *Conservation Biology*, 17(6), 1681–1693.
- Knutson, M. G., Richardson, W. B., Reineke, D. M., Gray, B. R., Parmelee, J. R., & Weick, S. E. (2004). AGRICULTURAL PONDS SUPPORT AMPHIBIAN POPULATIONS. *Ecological Applications*, 14(3), 669–684. <https://doi.org/10.1890/02-5305>
- Koch, A. J., Martin, K., & Aitken, K. E. H. (2012). The relationship between introduced

- European Starlings and the reproductive activities of Mountain Bluebirds and Tree Swallows in British Columbia, Canada. *Ibis*, 154(3), 590–600.
<https://doi.org/10.1111/j.1474-919X.2012.01242.x>
- Koenig, W. D. (2003). European Starlings and Their Effect on Native Cavity-Nesting Birds. *Conservation Biology*, 17(4), 1134–1140.
<https://doi.org/10.1046/j.1523-1739.2003.02262.x>
- Linz, G. M., Homan, H. J., Gaulker, S. M., Penry, L. B., & Bleier, W. J. (2007). EUROPEAN STARLINGS: A REVIEW OF AN INVASIVE SPECIES WITH FAR-REACHING IMPACTS. *Managing Vertebrate Invasive Species*.
<https://digitalcommons.unl.edu/nwrcinvasive/24>
- Manenti, R., Zanetti, N., Pennati, R., & Scari, G. (2017). Factors driving semi-aquatic predator occurrence in traditional cattle drinking pools: Conservation issues.
<https://doi.org/10.4081/jlimnol.2016.1447>
- Messerman, A. F., Clause, A. G., Catania, S. V. L., Shaffer, H. B., & Searcy, C. A. (2021). Coexistence within an endangered predator–prey community in California vernal pools. *Freshwater Biology*, 66(7), 1296–1310. <https://doi.org/10.1111/fwb.13717>
- Michaels, J., Batzer, E., Harrison, S., & Eviner, V. T. (2021). Grazing affects vegetation diversity and heterogeneity in California vernal pools. *Ecology*, 102(4), e03295.
<https://doi.org/10.1002/ecy.3295>
- Moyle, P. B. (1973). Effects of Introduced Bullfrogs, *Rana catesbeiana*, on the Native Frogs of the San Joaquin Valley, California. *Copeia*, 1973(1), 18–22.
<https://doi.org/10.2307/1442351>
- Nelson, N., & Piovia-Scott, J. (2022). Using environmental niche models to elucidate drivers of the American bullfrog invasion in California. *Biological Invasions*, 24(6), 1767–1783.
<https://doi.org/10.1007/s10530-022-02744-3>
- Parker, K. C., Parker, A. J., & Vale, T. R. (2001). Vertebrate Feeding Guilds in California’s Sierra Nevada: Relations to Environmental Condition and Change in Spatial Scale. *Annals of the Association of American Geographers*, 91(2), 245–262.
<https://doi.org/10.1111/0004-5608.00241>
- Paz, V. de, Baños-Picón, L., Rosas-Ramos, N., Tobajas, E., Tormos, J., & Asís, J. D. (2020). The role of artificial ponds in maintaining dragonfly populations in an intensified farmland landscape. A case of study in Zamora, Spain. *Journal of Natural History*.
<https://www.tandfonline.com/doi/abs/10.1080/00222933.2020.1850901>
- Ranching in California – California Cattlemen’s Association. (n.d.). Retrieved June 22, 2024, from <https://calcattlemen.org/ranching-in-california/>
- Riensch, D. L. (2005, December). Lizard abundances in various managed Central California grasslands. California State University, East Bay.
- Schmutzer, A. C., Gray, M. J., Burton, E. C., & Miller, D. L. (2008). Impacts of cattle on amphibian larvae and the aquatic environment. *Freshwater Biology*, 53(12), 2613–2625.
<https://doi.org/10.1111/j.1365-2427.2008.02072.x>

- Sheath, G. W., & Boom, C. J. (1997). Impact of beef cattle grazing systems on treading damage and forage supply. *Proceedings of the New Zealand Grassland Association*, 87–92. <https://doi.org/10.33584/jnzg.1997.59.2271>
- Simpson, N. O., Stewart, K. M., & Bleich, V. C. (2011). What have we learned about water developments for wildlife? Not enough! *CALIFORNIA FISH AND GAME*, 97(4).
- Starlings a Costly Problem for Livestock Producers. (n.d.). Plone Site. Retrieved October 8, 2024, from <https://www.ag.ndsu.edu/news/newsreleases/2015/jan-12-2015/starlings-a-costly-problem-for-livestock-producers>
- Swaim, K. E. (1994). *Aspects of the Ecology of the Alameda Whipsnake Masticophis Lateralis Euryxanthus*. Calif. State University, Hayward.
- Swartz, T. M., & Miller, J. R. (2019). Managing farm ponds as breeding sites for amphibians: Key trade-offs in agricultural function and habitat conservation. *Ecological Applications*, 29(7), e01964. <https://doi.org/10.1002/eap.1964>
- Swartz, T. M., & Miller, J. R. (2021). The American Pond Belt: An untold story of conservation challenges and opportunities. *Frontiers in Ecology and the Environment*, 19(9), 501–509. <https://doi.org/10.1002/fee.2381>
- Wheelock, I. G. (1912). *Birds of California: An Introduction to More Than Three Hundred Common Birds of the State and Adjacent Islands*. A.C. McClurg & Company.
- Zamora-Marín, J. M., Zamora-López, A., Jiménez-Franco, M. V., Calvo, J. F., & Oliva-Paterna, F. J. (2021). Small ponds support high terrestrial bird species richness in a Mediterranean semiarid region. *Hydrobiologia*, 848(7), 1623–1638. <https://doi.org/10.1007/s10750-021-04552-7>

Design and Analysis of a Modified Knee Brace to Provide Stabilization and Reduce Inflammation from Patellar Tendonitis in Volleyball Players

By Boston Richardson

Abstract

On average, 40-50% of volleyball players worldwide today suffer from patellar tendonitis, or “jumper’s knee.” However, hundreds of knee brace or strap designs on the market today fail to provide long-term relief for athletes because those designed for severe tendonitis are unsuitable for high activity levels. The knee brace design analyzed here contains an adjustable lacing system above the knee and added padding over the patella bone to serve as a knee pad for players. Research was conducted to compare different foams that would be used to protect the knee. A light and shock-absorbent foam will allow athletes to perform movements with a normal range of motion while providing additional support. The study involved material comparisons and range-of-motion tests on a prototype brace. A structural analysis machine tested the tensile modulus of the micro-adjustable lacing system along with a velcro knee strap. The range of motion of the knee in a rapid prototype was measured and compared to that of a knee with an overlapped brace and knee pad. The design was proven to have a larger range of motion compared to a traditional brace and pad. The adjustable lacing system was also shown to have a greater tensile modulus than the velcro strap. Overall, the brace design is far more durable and adaptable to activity levels required in volleyball while providing the same support as available designs on the market.

Introduction

Patellar tendonitis or “jumper’s knee” is most common in volleyball players, affecting 40-50% of athletes worldwide. Patellar tendonitis is defined as micro-tears in the patellar tendon, which connects your knee cap to the shin bone. Repeated jumping and landing movements in gameplay cause microtears in the tissue, weakening the knee [4] and causing inflammation of the patellar tendon. For elite players, the injury often results in tendon rupture [7].

The easiest and most common ways to treat this injury are tendon straps that reduce tensile stress (Figure 1), and stabilizing braces, to track the patella and provide proper alignment (Figure 2). While patellar tendon straps are often successful in reducing swelling and pain in the knee, they reduce the range of motion during activity. On the other hand, stabilizing braces allow for a larger range of motion during activity, however, they do not provide long-term effects for mild to severe cases [9]. In volleyball players, any type of knee support combined with required knee pads will also reduce the range of motion of the knee. In short, all viable treatment options for patellar tendonitis, excluding surgery, are suboptimal.

The mission of this project was to design a brace that is effective in the treatment of patellar tendonitis and offers stability and pain relief, even during intense activity. A brace design that incorporates a tendon strap below the patella and knee tracking will provide correction and pain relief for patients. Micro-adjustable lacing system, in replacement of velcro, will add

durability and longevity to the brace. An added insert over the knee for a knee pad will allow for a wider range of motion for volleyball players by eliminating the need for two accessories.

Figure 1: *Jumper's Knee Strap* [5]



Figure 2: *Tru-Pull Knee Brace* [6]



Methods

Design

The initial design of a modified knee brace for comfort and stability was motivated by the widespread issue of patellar tendonitis in volleyball players caused by jumping during gameplay. To approach this problem, research was conducted to understand how knee flexion in jumping causes micro tears in the patellar tendon and what different treatment options-such as patellar tendon straps and knee tracking braces- can offer. Several constraints, such as the overall fit and structure of the brace, were considered along with multiple ways to improve existing knee braces to determine which components would enhance the design. These include an added strap above the knee to reduce inflammation and enhance patellar stability, an advanced adjustment system to increase longevity and durability, and an added knee padding to protect the patella and improve the range of motion.

Testing and Prototype Development

A literature search was conducted to determine which material type is suitable for a knee pad insert. The density, compression set (permanent deformation remaining after compressing it), and tensile modulus (ratio of tensile stress to strain when undergoing deformation) were compared for different brands of silicon gel foam, polyurethane foam, and ethylene vinyl acetate (EVA) foam to determine their ability to absorb impact. Silicone gel foam is most commonly used in electronic component protection, polyurethane foam is used for bedding and furniture, and EVA is used in the soles of athletic shoes.

A rapid prototype of the design (Figures 3 and 4) was also created using a traditional Hely & Weber knee sleeve, an EVA foam knee pad, and a BOA® cord and dial fastening system. The BOA® system is most commonly used in mountain biking and snowboard shoes and consists of a small nylon cord and a dial that can be turned to tighten or loosen the cord. An EVA foam knee pad, chosen for its lightweight and shock-absorbing qualities, was hand-sewn into the

inside center of the knee sleeve. Small slits were cut around the sleeve every 2 cm, one ring of slits just below the knee pad and another 2.5 cm below that. Then, the fastening system was woven between the slits, all the way around the knee sleeve.



Figures 3: (A) *Rapid Prototype Knee Sleeve With Inserted Knee Padding and Weaved BOA® Adjustment System.* (B) *Prototype on A Player's Knee*

The fastening system's tensile strength was tested and compared to a velcro patellar strap (Figure 1) using a Vernier 1,000 N structural analysis machine that pulled the materials up to 40N and logged the elongation of the material every millisecond. The data was graphed on a stress vs. strain graph and a linear fit was set to the line using $Y = mx$. Using a Mindateq Protractor application, the range of motion for the rapid prototype brace, the Tru Pull brace alone, and the Tru Pull brace with an added knee pad were measured for 5 different subjects. The subjects' initial full flexion squat angle and angle with each added variable were recorded and the angles in each apparatus were compared to one another using a bar graph.

Results

To test the validity and actuality of the design, three different foams were researched and compared to find the most effective and reliable, and three experiments were conducted on the design's unique components: the tensile strength of the BOA® cord and dial system (Figure 4) and the effect of the prototype brace on the knee's range of motion (Figure 5).

Table 1 compares the material properties of three different foam materials and brands. The tensile strength, density, and compression set were shown to be the largest for silicone gel due to its overall hardness.

Table 1: Comparing the Three Most Common Shock Absorption Foams for Knee Padding: aGel [1], LAMATEK [3], Shocktec [8]

Foam Type	Brand	Density (lbs/ft ³)	Tensile Modulus (lbs/in ²)	Compression Set (%)
Silicone Gel	aGel	65.55	340.8	15
Ethylene Vinyl Acetate Foam	Lamatek	3.8	143	7
Polyurethane Elastomer Foam	Shocktec	65.44	51.0	2.34

EVA and Polyurethane Elastomer foam tend to have a much lower compression set, however, EVA’s lower density makes it much lighter. These properties will ensure that the chosen material will be lightweight, durable, and absorb impact when volleyball players fall or dive on the ground.

Figure 5 compares the stress vs. strain of a traditional velcro tendon strap (Figure 1) and a BOA® cord and dial system using a Vernier Structures and Materials Tester. Ten total data points for each material were graphed using Microsoft Excel. The slope of the lines on the Stress vs. Strain graph depicts the tensile modulus of each material.

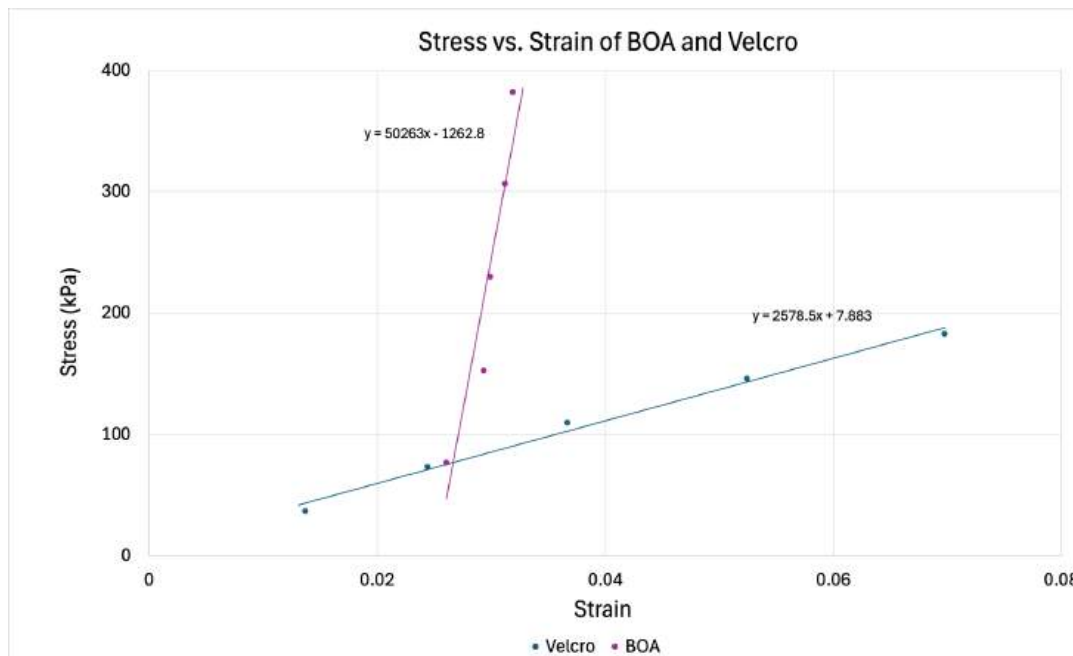


Figure 5: *Stress vs. Strain Graph of BOA® and Velcro Strap Using Vernier Structures and Materials Tester.*

The slope of the BOA®, 50263 kPa, indicates that its tensile modulus is higher than that of the velcro at 2578.5 kPa. With a higher tensile modulus, the BOA® adjustment system will increase the longevity of the brace design and apply constant pressure to the patellar tendon to reduce the effects of tendonitis.

Figure 6 describes and compares the range of motion allowed by the prototyped BOA® knee sleeve and the Tru Pull knee brace with and without additional knee padding. Five people were tested on their normal squat angle and with each added variable for 5 trials. Each data point is an average of the subject's range of motion.

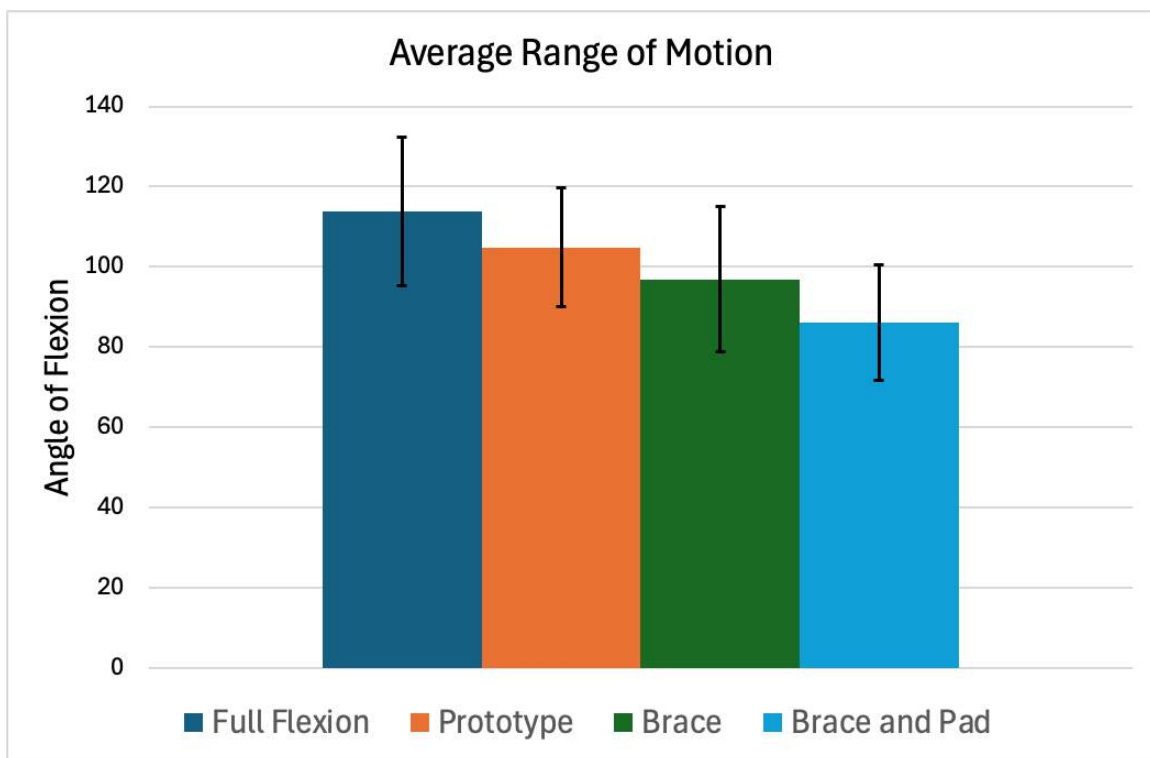


Figure 6: *Average Range of Motion Among 5 Participants (Note: all angles measured in degrees from the positive y-axis)*

The added layers of a brace and a knee pad bunch up at the knee when players bend their knees to jump. This stops the knee at a smaller angle from the vertical and results in a weaker and less explosive jump. The Tru-Pull brace reduced the range of motion from the prototype by as much as 15.5° and 11.4° on average. Tru-Pull with an added knee pad reduced the range of motion from the prototype by as much as 30.8° and by 18.8° on average. With less bulkiness, the prototype allows for a greater range of motion and, therefore, more power in the jump.

Discussion

The purpose of the design process was to create a brace that would limit the effects of patellar tendonitis and offer stability and pain relief, even during high-impact activities. The results of specific research and testing validated the design's durable adjustment system and increased range of motion compared to traditional designs. Unlike conventional, rigid, and bulky knee braces, the design would support players' knees while adding comfort and a greater range of motion. The adjustable BOA® system applies constant pressure to the patellar tendon and increases the longevity of the design. The innovative construction not only addresses the specific needs of athletes with patellar tendonitis but also incorporates feedback from users to ensure optimal functionality.

Table 1 shows a literature search gathered the material properties of the three common foam paddings. The materials were evaluated on their density, compression set, and tensile modulus, which help to determine a material's ability to absorb and withstand repeated impacts. When compared, EVA foam is the best option for protective knee padding. Its lightweight form and low compression set make it suitable for high activity levels and repeated impact. Despite Polyurethane's low compression set and silicone gel's high tensile strength, both are far too heavy to be worn during gameplay. Therefore, EVA is the most recommended for protection and performance. The limited scope of the design and experiments excludes pricey or rare foams such as PORON® urethane foam, which may be more effective in protecting the knee and suitable for gameplay.

As seen in Figure 5, the slope of the line displaying the BOA® system's tensile modulus is much higher than that of the velcro patellar strap. The BOA® system is less likely to break or to come loose during gameplay, ensuring constant pressure on the patellar tendon to reduce inflammation. Unlike velcro, the BOA® system does not wear down after repeated use, has a wider range of sizes, and can be adjusted efficiently. In summary, the BOA® system's superior tensile modulus and durability make it a more reliable choice for maintaining consistent support on the patellar tendon compared to traditional velcro straps, enhancing performance and stability during gameplay.

As seen in Figure 6, the subjects' range of motion was greatly decreased by the Tru-Pull knee brace with and without a knee pad. The structure of the Tru-pull knee brace includes plastic supports along the sides which restricts subjects' knee flexion. Several subjects noted that it was very "unnatural" and "uncomfortable" when squatting and jumping. However, the prototype reduced the participant's range of motion by almost none, due to its simple, sleeve-like structure. Along with uncomfotability, reduced knee flexion will also affect the athlete's performance. During gameplay, a player would see an even greater decrease in vertical jump height and slower movements due to a less powerful muscle twitch. Restricted motion will have long-term effects, potentially directing stress to surrounding muscles and creating additional overuse injuries. While the Tru-Pull knee brace offers structural support, its negative effect on athletic performance highlights the need for more effective designs that balance support with freedom of movement.

An attempt to measure the brace's ability to track and stabilize the patella using the Digital Imaging Correlation Engine was inconclusive due to the program's lack of precision. Despite this setback, the design's structural similarities to traditional knee braces provide a basis for its expected effectiveness in treating patellar tendonitis. The design incorporates features proven to support the knee joint and alleviate strain on the patellar tendon. Future testing with more precise tools could offer a clearer assessment of its performance. Until then, the design's foundational elements suggest that it offers similar effects to traditional knee braces, such as the Tru-Pull brace, in the treatment of patellar tendonitis.

Several details of the design have the potential to limit its overall success. In the prototype, the BOA® adjustable system did not distribute pressure evenly around the knee. For the BOA® to be effective in a manufactured design, it must be weaved through rigid holes and placed on a rigid base that will translate the pressure to every point along the front and back of the knee. The BOA® dial must also be covered by a layer of padding or fabric to prevent it from damage when hitting the ground. Several factors in the experimental design may have influenced the results, including imprecise angle measurements and the lack of resources to test the brace's overall effectiveness in treating patellar tendonitis. Flexion angles were measured using a digital protractor that lacked defined reference lines and only provided a rounded measurement, which could lead to inaccuracies in assessing the prototype's impact on a player's range of motion. Future researchers may use imaging technology to accurately analyze the brace's increased range of motion, its ability to treat patellar tendonitis and its long-term effects on performance. This data can also be used to improve the design and predict its success. Further experimentation may also discover the long-term effects of an increased range of motion on a player's performance. Ultimately, while this study lays the groundwork for a promising treatment approach to patellar tendonitis in volleyball players, further refinement and rigorous testing are essential to ensure its effectiveness and safety on the court.

Conclusion

Patellar tendonitis is the leading injury in volleyball players, affecting almost half of players worldwide. Since patellar tendonitis is an overuse injury, the condition is almost incurable for active athletes. Treatment options off of the court are not suitable for intense movement and activity, leaving athletes with pain and discomfort when playing. This study attempted to find an effective treatment option that is suitable for gameplay while implementing an alternative adjustment system. Experimentation has proven the design to be durable, comfortable, and have the potential to improve performance.

Works Cited

1. Alpha GEL. “αGEL Products (Shock Absorption/Gel Sheets (θ-GEL)).” Accessed August 22, 2024. https://taica.co.jp/gel/en/product/shock_absorption/theta_sheet.html.
2. De Vries, Astrid J. “The Impact of Patellar Tendinopathy on Sports and Work Performance in Active Athletes.” Routledge. Accessed August 20, 2024. <https://www.tandfonline.com/doi/epdf/10.1080/15438627.2017.1314292?needAccess=true>.
3. LAMATEK. “Technical Data Sheet: 4# EVA Ethyl Vinyl Acetate Blend » LAMATEK.” Accessed September 2, 2024. <https://lamatek.com/data-sheets/closed-cell/ev-series-medium-4lb-eva>.
4. Massachusetts General Hospital. “Patellar Tendonitis - Treatments & Symptoms | Massachusetts General Hospital.” Accessed August 17, 2024. <https://www.massgeneral.org/orthopaedics/sports-medicine/conditions-and-treatments/patella-tendinitis>.
5. Mueller Sports Medicine. “JUMPERS KNEE STRAP BLACK OSFM | Golf | By Sport | Open Catalog | Mueller B2C Canada.” Accessed August 15, 2024. <https://www.muellersportsmed.com/b2c-ca/en/Open-Catalog/By-Sport/Golf/Jumper%27s-Knee-Strap/p/992>.
6. OrthoBracing. “DonJoy Tru Pull Lite Knee Brace.” OrthoBracing. Accessed August 26, 2024. <https://orthobracing.com/products/donjoy-tru-pull-lite>.
7. Roizman, Tracy. “Patellar Tendonitis - Rheumatology Advisor.” Rheumatology Advisor, November 30, 2022. <https://www.rheumatologyadvisor.com/ddi/patellar-tendonitis/>.
8. SHOCKtec. “SPECS & MSDS - SHOCKtec.” Accessed August 17, 2024. <https://shocktec.com/specs-and-msds/#physical-and-chemical-testing>.
9. Sprouse, Ryan A. “Braces and Splints for Common Musculoskeletal Conditions | AAFP.” American Family Physician, 15 November 20. <https://www.aafp.org/pubs/afp/issues/2018/1115/p570.html>.

A Thorough Analysis of Machine Learning Models For Sports Results Prediction

By Adit Singh

Abstract

AI machine learning has become widely used in order to predict outcomes through learning and processing particular inputs. This paper is a review of a project that uses AI's fundamental properties in order to predict the chances of American Football teams winning games. This project aims to use the statistics measured during the course of this game to properly predict the probability of a team winning their next game. This paper will revolutionize sports analytics and the world of coaching by predicting a team's success rate in a season based off of their statistics and how they compare to numbers needed to be prosperous in the league. This way coaches can train their players up in specific areas that can measure up to these analytics. Furthermore, analysts will now be able to pinpoint who the successful teams really are and observe how they translate to the bigger stage of championship success. The high accuracy of this project shows that these models can be used not just in the realms of American football, but across all sports on the world stage.

Introduction

Sport result predictions across the board could be a pivotal move for many sports experts and players. These individuals are often asked for their predictions for forthcoming matches, which then appear online or in newspapers. Other stakeholders who are interested in predicting sport results include bookmakers and sports betting platforms, as well as gamblers who bet on match results, or for certain events occurring within a particular match. Sport result prediction is also potentially useful to players, team management and performance analysts in identifying the most important factors that help to achieve winning outcomes, upon which appropriate tactics can be identified ((Bunker, R., & Susnjak, T. The Application of Machine Learning Techniques for Predicting Results in Team Sport: A Review. *arXiv*. **1912**, 11762 (2023))). This concept applies to the players and experts in American Football as well, thus forming the basis of this project. But before understanding how AI can perform machine learning tasks by using statistical data from football teams, it is important to understand how this game started.

American football came into existence on November 12, 1892 ((Football HOF, P. (2019, January 20). Birth of pro football. <https://www.profootballhof.com/football-history/nov-12-birth-of-pro-football/>)). Its roots stem from the sports soccer and rugby, two longtime popular sports. The very first game was played between two universities, Harvard and Yale. The game was initially played as more of a rugby styled game, but eventually transitioned into a "scrummage", trying to kick it forward through the mass of players—Harvard opted for "heeling it out," or kicking the ball backward to a teammate. The further transformation of English rugby into American football came chiefly through the efforts of Walter Camp.

A famous rugby player, Camp is often regarded as “the father of American Football,” would pioneer the new rules that would help transform rugby into modern day American Football. He proposed that the possession of the ball be given to one of the two teams. Camp also proposed a rule that a team must advance the ball 5 yards or lose 10 in three downs (plays), or it would be obliged to surrender the ball to the other side. Camp was also responsible for having 11 players on a side, for devising a new scoring system in 1883 with two points for a touchdown, four points for the goal after a touchdown, and five points for a field goal (a field goal became worth three points in 1909, a touchdown six points in 1912), for creating the quarterback position, for marking the field with stripes, and for proposing several other innovations, but it was those two simple rules adopted in 1880 and 1882 that fundamentally created American football ((*Oriard, M. (2024, July 22). American football.* <https://www.britannica.com/sports/American-football>)).

After the crucial rule changes, the play of the game was relatively open, featuring long runs and numerous lateral passes, as in rugby. In 1888 Camp proposed that tackling below the waist be legalized, in order to offset the advantage of speedy backs streaking around the ends. The new sport emerged as a popular pastime in the college setting, as that is where the sport arose. Professional football would eventually find its footing in 1920, as the American Professional Football Association, which changed to the National Football League (NFL) in 1922, with Jim Thorpe as its nominal president. The NFL took its modern shape in 1933 under new president Joe Carr, when it was reorganized in two five-team divisions of big-city clubs whose leaders would meet at the end of the season for the NFL title. This would soon develop into the modern day sport with two conferences, each with 4 divisions of 32 teams total, each fighting its way to the top of their respective conferences and face off in the season finale at the superbowl.

If you are unfamiliar with the game of American football, here is a quick summary. Teams advance the ball down the field in a series of set-piece plays (called 'downs'). The play ends when the ball-carrier is tackled to the ground, goes out of bounds, or if a forward pass hits the ground. Teams have to gain at least 10 yards in 4 plays to keep possession and earn a new set of 4 downs. Teams score 6 points for a touchdown, 3 points for a field goal and, after a touchdown, can kick for one extra point or pass or run for two extra points. On each play teams can choose to run (by handing the ball off to a running back) or pass (the quarterback throws to a receiver). On fourth down teams will usually try to kick a field goal or punt the ball away if they don't think that they can gain enough yards for a first down. Otherwise, they have to surrender possession of the ball to the opposition at the end of the play. Players on the offensive team (the team with the ball) can protect the player with the ball by blocking opposing tacklers. Defenders can use their hands to push or pull opponents out of the way to get to the ball carrier. Only the ball carrier can actually be tackled and pulled to the ground ((*IFAF. (2018, March 12). A GUIDE TO TACKLE FOOTBALL.*

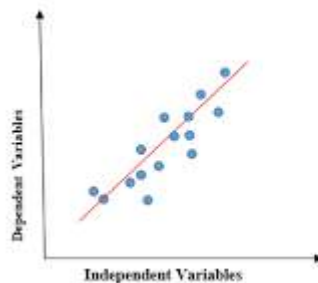
<https://www.americanfootball.sport/a-guide-to-tackle-football/#:~:text=The%20Basics,-Teams%20>

[20score%206&text=Teams%20advance%20the%20ball%20down.new%20set%20of%204%20d
owns. \)\)](#)

Methods

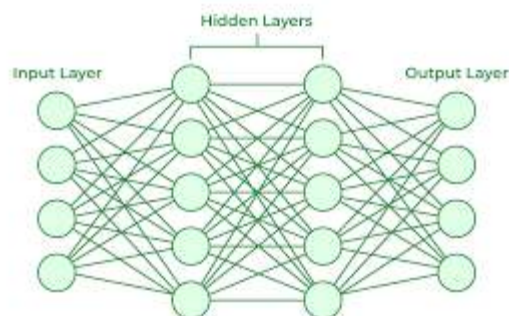
Training and Testing

In most AI models, we create a training dataset and a testing dataset. The training data will involve the model equation to be tested on known data with a known output in order to validate accuracy, while the testing dataset will involve the AI model being used on a new dataset for a new prediction. With an image, the training data will be a known image where the model will hopefully output that exact image. If not perfectly outputted, the weights and pixels numbers are adjusted to get a correct output.



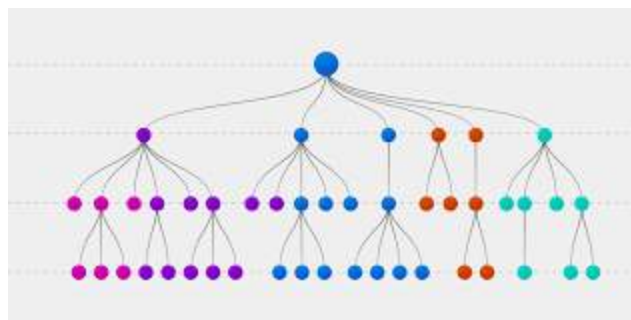
Linear Regression

Different AI machine learning algorithms can be used to analyze given inputs, then those inputs are put through certain conditions in a specific model structure to reach a certain output. One of these models is called a linear regression model. Linear regression is a fundamental machine learning algorithm used for predicting numerical values based on input features. It assumes a linear relationship between the features and the target variable. The model learns the coefficients that best fit the data and can make predictions for new inputs. Let's take an image for instance. Images are made up of a layer of different pixels, each corresponding to a specific color or shade. Each pixel represents a number, and these numbers are used as the input for the AI linear regression model. The model processes each pixel one at a time, then adds them all together to output the full image in pixels. However, with 3 dimensional images, there is an extra layer of pixels that represent the background features of the image. Therefore, the weights will be multiplied to each pixel, then a summation is taken.



Neural Network

A more popular form of AI machine learning is the neural network model. Neural networks are used in a large number of web services including Chat GPT. These models will typically take multiple inputs and narrow them down to one final prediction. Neural networks contain multiple layers of neurons that are connected to each other by intersecting lines which represent the weights. Each neuron represents the specific parameters of the inputs as they flow through each layer in the network. For example, when processing an image, each neuron in the network represents a pixel, and, as aforementioned, each pixel represents a number. Each collection of pixels make up a layer in the neural network. The last layer of the network contains a row of pixels from 0 to 10. When an image or dataset is processed, it lights up certain neurons in one layer, which activates a pattern of neurons in the next layer, then this continues until a specific number neuron is activated in the final layer to predict the final output. With an image, each layer picks up a certain edge of line which corresponds to a specific part of the next layer. Each layer then combines the edges and lines to form the image features. The lines that connect each layer of neurons represent the weights that can capture the pattern of neurons in the input. Each weight is a number that connects the neurons from one layer to a specific neuron in the next layer. The weights are all multiplied to each activated neuron in the previous layer and then a summation is taken. This sum represents the number of the neuron that will be activated in the next layer. Like with linear models, a training set and testing set is used. When training the data, if we discover a difference between our desired output and the output of our model, we measure the amount of error between our output and the actual output, then take an average using the mean squared error formula. Using this error margin, we modify our weights by either increasing or decreasing them and then running the model again until we get the desired output.



Decision Trees

The final model that is commonly used in most machine learning algorithms is decision trees. These trees are made up of an initial root at the top and branches that extend towards the bottom, with the bottom of the tree representing the prediction. The root of the tree represents the initial conditions that the input faces. In a decision tree, the conditions result in either a “yes” or “no” decision. If the input desires an output with the specified condition, the input moves down the “yes” branch, and if not, it proceeds down the “no” branch. On each of these branches, the input will reach a node which represents another condition. This condition will also involve a “yes” or “no” decision, whether the output has this condition or not. Then, the input proceeds to another “yes” or “no” branch. This continues until the output is reached. These decision trees can be further developed by being combined into a random forest. A random forest takes multiple

decision trees, combines them, then takes the average of all the predictions to get the final output.

This project's AI model takes its data from subsets of stats from each of the 32 teams in the National Football League(NFL). The dataset in this project in particular uses a CSV file with the data collected into rows and columns. The rows is a specific team from each year, going from the year 1998 to 2019. The columns are different season stats for each team. The headers are the names of each statistic, such as team wins, scoring percentage, and yards per play ((Talbot. (2020, January 16). American Football Team Stats (1998 - 2019). <https://www.kaggle.com/datasets/ttalbitt/american-football-team-stats-1998-2019>)).

Using this dataset, a linear AI model was constructed, taking into account two subsets of statistics and storing them into x and y variables. The x variable functions as the input data, the subset of data that functions as the specific columns in the dataset that we choose as our "target variables", or in other words the statistics in the dataset we would like to use to "train" the model in order to predict our y variable, the variable that represents our AI model prediction. For example, the columns '#of yards', 'PF', 'number of yards/play', 'score percentage', 'number of 1st downs'" in the aforementioned dataset is accessed, then stored in an input variable called 'x'. The number of yards stat refers to the total amount of yards gained by a team throughout the season through offensive plays. PF stands for "points for" team throughout a season, or the total amount of points scored by a team throughout the season. The number of yards per play refers to the number of yards gained on each offensive play through a team's season. Score percentage is the frequency of which the team scores points throughout the season, and number of first downs refers to how often the offense is able to advance the ball down the field by 10 yards or more. These numbers all contribute to success in American Football, so the higher these statistics are, the more wins the team will get. Among these columns, a single column will be targeted for the first coefficient of the model, for example the column scoring percentage will be stored in a variable called "target column." This variable is then fit into the linear regression model code, training the model to predict the new variable "wins". Additionally the training performance is around the same as the testing performance, which shows that our model is not overfitting. Overfitting occurs when the machine learning model accurately predicts the results for the training data, but not for the testing data.

Results

All this is coded in python, the software pycharm. When coding this, we made sure to import a linear model for the machine learning portion of the project, as well as the dataset in a CSV file. Once these resources were imported, the variables were declared and assigned as mentioned above. The result was successfully outputting the prediction of team wins for the next season.

This result is done by using a linear regression model that is imported into the software into the early line of code. The column that the model targeted from the input variable set was the score percentage variable, which was automatically measured in comparison to the teams

with the most wins in the dataset. In order to do this, the model takes into account the weights of the dataset and multiplies it to the scoring percentage variables. These weights come from the team records in the rest of the dataset and the other columns. Once these weights are multiplied to the original input, the final product is the output, which in this case is the number of wins for the chosen team. This process is performed on the known wins amount in the dataset then is tested by using the model to predict the number of wins for the team going into the new season.

We constructed a table to measure the results of our multiple models and their accuracy in determining the output. Here is the table:

Model	Accuracy	Precision	Recall	F1 Score
Linear Model	0.72	0.68	0.7	0.69
Neural Network	0.84	0.8	0.83	0.82
Decision Tree	0.78	0.75	0.76	0.75
Random Forest	0.82	0.81	0.79	0.8

The accuracy here represents the percentage of correct predictions in the model. The precision refers to the quality of the predictions made by the model. The recall further ensures effectiveness by measuring the frequency of correct predictions in each model. The F1 score takes the mean of the precision and recall of the models to measure the overall reliability of the models. The numbers above thus represent that these models that we coded are effective in determining outputs.

This next table shows the MSE we calculated for each model. MSE stands for mean squared error. It is the average amount of error between the correct output and the actual output in a model. We take this number into account when adjusting our weights so that we can reach a correct prediction during the training.

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Mean
Error
Squared

Model	MSE
Linear Model	12.5
Neural Network	10.3

Model	MSE
Decision Tree	14.2
Random Forest	11.7

To better understand the machine learning process let's look at my hometown American Football team, the Detroit Lions. The Lions went 12-5 during the 2023 season, reaching the conference championship final, one game short of the superbowl. They were the 5th highest scoring team in 2023, scoring 461 total points and 58 total touchdowns ((*ESPN*. (2024, February 2). Detroit Lions Team Stats 2023. https://www.espn.com/nfl/team/stats/_/type/team/name/det)). Using these stats, we can use the same AI model mentioned above to predict the number of wins the Detroit Lions will have in 2024. Using these scoring numbers as our input variable, the model takes the same scoring statistics from all the other teams during the season as weights, and multiplies it to the input variable. Then an output variable is declared to act as the results of the equation. We start by performing this process on the 2023 statistics to see if we can reach the 12 win number that the Lions achieved last season, in order to train the model first. We keep adjusting the weights until we reach this number, then apply this model to some statistics we have at the start of the next season to predict their 2024 record. This prediction would come out to be approximately 13 wins, 4 losses.

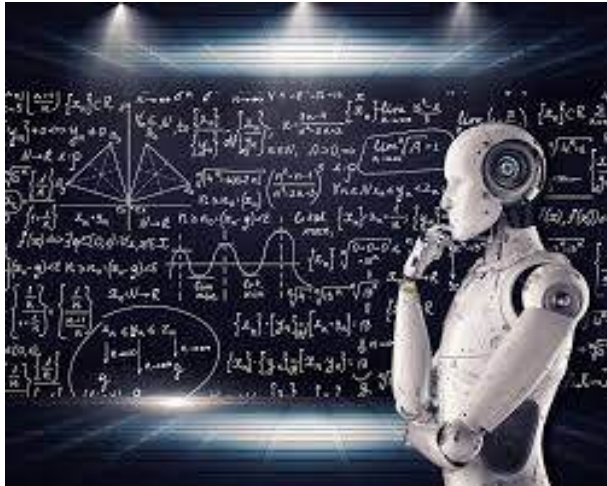
Conclusion

Add open question (future potential directions)

In summary, this paper covers all the bases of specific AI machine learning models, as well as the process and data involved in the machine learning project involving American football. We begin by discussing the potential implications of sporting predictions to players, sports experts, and other analysts in their quest to find certain advantages and disadvantages in reaching winning outcomes. We further discuss how this relates specifically to American Football by transitioning into its history and development. This history helps give a little more context of the sport to those unfamiliar, then the rules of the sport are explained in order to give an explanation of how many of football's statistics are measured. Furthermore, we explored the common AI models and how they are able to reach certain outputs. The models we covered consisted of linear regression, neural networks, decision trees, and random forests. These models were all incorporated into the project, which used a dataset taking 32 NFL teams statistics from 1998 to 2019 in order to offer an input for the model. We selected 5 specific columns for the input and put them through all the possible models to reach the final output, which was the number of wins. Finally, we tested the reliability of each of the models by calculating accuracy and F1 scores, as well as the MSE. All this can contribute to major advances in sporting predictions and strategies that are developed towards future sports games, overall resulting in a whole new world of sports.

Works cited

- University, C. (2022, October 12). New AI Algorithms Predict Sports Teams' Moves With 80% Accuracy.
<https://scitechdaily.com/new-ai-algorithms-predict-sports-teams-moves-with-80-accuracy/>
- Agarwal, R. (2020, October 22). The Use of Machine Learning in Predicting Sports Match Outcomes. <https://mlwhiz.com/blog/2020/10/22/sports-match-outcome-guest/>
- Bunker, R., & Susnjak, T. The Application of Machine Learning Techniques for Predicting Results in Team Sport: A Review. *arXiv*. **1912**, 11762 (2023)
- Li, J & Fang, A. In search of the next generation of training sets for language models. *arXiv* **2406**, 1179(2024)
- Football HOF, P. (2019, January 20). Birth of pro football.
<https://www.profootballhof.com/football-history/nov-12-birth-of-pro-football/>
- Talbott. (2020, January 16). American Football Team Stats (1998 - 2019).
<https://www.kaggle.com/datasets/ttalbitt/american-football-team-stats-1998-2019>
- ESPN. (2024, February 2). Detroit Lions Team Stats 2023.
https://www.espn.com/nfl/team/stats/_/type/team/name/det
- IFAF. (2018, March 12). A GUIDE TO TACKLE FOOTBALL.
<https://www.americanfootball.sport/a-guide-to-tackle-football/#:~:text=The%20Basics,-Teams%20score%206&text=Teams%20advance%20the%20ball%20down,new%20set%20of%204%20downs.>
- Oriard, M. (2024, July 22). American football.
<https://www.britannica.com/sports/American-football>



Proof of Zsigmondy's Theorem

Nayer Sharar

September 6, 2024

Abstract

Zsigmondy's Theorem is a classical result in number theory. In this article, we provide a proof of Zsigmondy's Theorem using cyclotomic polynomials, Mobius Inversion and Lifting the Exponent Lemma. We also provide brief introductions to these topics and also mention the proofs of associated lemmas used.

Contents

1	Introduction	2
2	Roots of Unity	3
3	Arithmetic Functions and Mobius Inversion	4
4	Cyclotomic Polynomials	6
4.1	Definition	7
4.2	Lemma	7
5	LTE and Other Important lemmas	11
6	Combining the Results and the Proof	13

1 Introduction

Statement

Let a, b be two co prime numbers and let $n > 1$ be another integer. There exists a prime divisor of $a^n - b^n$ that does not divide $a^k - b^k$ for any $k \in \{1, 2, \dots, n - 1\}$, except for the following cases:

1. $a = 2$ and $b = 1$, and $n = 6$,
2. $n = 2$ and $a + b$ is a power of 2.

This prime divisor is called a **Primitive Prime Divisor** of $a^n - b^n$ [4].

Applications of Zsigmondy's Theorem

Zsigmondy's theorem is a powerful result and can be applied to a variety of number theory problems. While a complete demonstration of its usage is beyond the scope of this paper, we will mention a few simple applications.

Application 1

Problem: Let p, q be prime numbers with $p > q > 2$. Prove that $2^{pq} - 1$ has at least three distinct prime factors. (*Polish Math Olympiad 2010*)

Solution: Suppose z is a prime dividing $2^p - 1$. Then $2^p \equiv 1 \pmod{z}$, which implies $2^{pq} \equiv 1 \pmod{z}$. Similar results hold for any prime r dividing $2^q - 1$. Applying Zsigmondy's theorem, $2^{pq} - 1$ must have a primitive prime divisor t which does not divide $2^p - 1$ or $2^q - 1$. Therefore, t, r , and s must be distinct primes, which completes the proof.

Application 2

Problem: Prove that a sequence defined by $a_n = 3^n + 2^n$ contains no three terms in geometric progression. (*Romania*)

Solution: Assume, for contradiction, that there are three terms a_i, a_k, a_j in geometric progression with $i < k < j$. Then $a_i \cdot a_j = a_k^2$. By Zsigmondy's theorem, a_j must have a prime factor that a_k does not, leading to a contradiction.

Note that these problems are not trivial to solve, illustrating the power of Zsigmondy's theorem.

Outline of the Proof of Zsigmondy's Theorem

We will prove Zsigmondy's theorem using the following ideas:

- Roots of Unity and Cyclotomic Polynomials,
- Mobius Inversion,

- Lifting the Exponent Lemma.

This paper will provide a brief introduction to these topics and combine them to prove Zsigmondy's theorem.

2 Roots of Unity

To understand Roots of Unity, it is essential to be familiar with complex numbers and their basic properties. Assuming familiarity with complex numbers, our discussion will be brief.

Complex Numbers

A complex number z is of the form $a + bi$, where i is defined as $\sqrt{-1}$. The number a is called the **real part** of z , and b is called the **imaginary part**. The conjugate of $z = a + bi$ is $\bar{z} = a - bi$. In the Argand plane, a complex number $z = a + bi$ represents the point (a, b) in the Cartesian plane. The argument of z , often denoted as $\arg(z)$, is the angle made by the complex number z with the real axis in the positive direction. The modulus of z , denoted by $|z|$, is equal to $\sqrt{a^2 + b^2}$ [2].

Complex numbers follow algebraic rules for addition and subtraction, where the operations on the real parts and imaginary parts are applied separately. The multiplication of two complex numbers z_1 and z_2 results in a complex number z with:

$$\arg(z) = \arg(z_1) + \arg(z_2) \quad \text{and} \quad |z| = |z_1| \cdot |z_2|.$$

Moreover, a complex number z can be represented by:

$$z = r (\cos \theta + i \sin \theta),$$

where θ is its argument and r is its modulus. Furthermore:

$$z \cdot \bar{z} = |z|^2.$$

Roots of Unity

A complex number z is called an **n -th root of unity** if it satisfies the equation:

$$z^n = 1.$$

Note that $|z|^n = 1$ implies $|z| = 1$. If $\arg(z) = \theta$, then $\arg(z^n) = n\theta$. Since $\arg(1) = 2k\pi$ for all integers k , the possible values of θ are $\frac{2\pi k}{n}$.

There are exactly n distinct values of z , and the n -th roots of unity can be expressed as:

$$z_k = \cos\left(\frac{2\pi k}{n}\right) + i \sin\left(\frac{2\pi k}{n}\right) \quad \text{for } k = 0, 1, 2, \dots, n-1.$$

A complex number z is called a **primitive n -th root of unity** if it is an n -th root of unity and there does not exist a k such that $n > k$ and $z^k = 1$. The primitive n -th roots of unity are numbers ζ_n^k where $n > k$ and $n, k \in \mathbb{N}$.

3 Arithmetic Functions and Mobius Inversion

In this document, we briefly discuss some common notations and properties of certain number-theoretic functions, focusing on Mobius functions and the Mobius inversion formula which will be used in our proof later.

Basic Definitions

- An **arithmetic function**, often called a 'number-theoretic function', is a function $f : \mathbb{N} \rightarrow \mathbb{C}$.
- An arithmetic function is called **multiplicative** if $f(mn) = f(m) \cdot f(n)$ for all m, n such that $\gcd(m, n) = 1$. Furthermore, it is called **completely multiplicative** if $f(mn) = f(m) \cdot f(n)$ for all m, n without any condition on $\gcd(m, n)$.

Common Arithmetic Function

For a positive integer n , let $d(n)$ denote the number of positive divisors of n . Mathematically, $d(n)$ can be expressed as:

$$d(n) = \sum_{i=1}^k (r_i + 1)$$

where $n = \prod_{i=1}^k p_i^{r_i}$ is the prime power factorization of n .

Properties

Multiplicative Functions

Let f be a multiplicative function. Then, the function $F(n) = \sum_{d|n} f(d)$ is also multiplicative.

Proof. It suffices to show $F(mn) = F(m) \cdot F(n)$. Note that each divisor of mn can be expressed as a unique divisor of m times a divisor of n since $\gcd(m, n) = 1$. Hence,

$$F(mn) = \sum_{d|mn} f(d) = \left(\sum_{d|m} f(d) \right) \left(\sum_{d|n} f(d) \right) = F(m) \cdot F(n).$$

□

Mobius Function

The **Mobius function** $\mu(n)$ is defined as follows:

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ (-1)^m & \text{if } n \text{ is square-free and has } m \text{ prime divisors,} \\ 0 & \text{if } n \text{ is not square-free.} \end{cases}$$

The Mobius function is multiplicative. The proof follows from induction and the properties of number-theoretic functions.

Another important property is:

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases}$$

Mobius Inversion Formula

The **Mobius Inversion Formula** states that if F and f are functions such that:

$$F(n) = \sum_{d|n} f(d),$$

then:

$$f(n) = \sum_{d|n} \mu(d) \cdot F\left(\frac{n}{d}\right).$$

Proof.

$$\sum_{d|n} \mu(d) \cdot F\left(\frac{n}{d}\right) = \sum_{d|n} \mu(d) \cdot \sum_{t|\frac{n}{d}} f(t) = \sum_{dt|n} \mu(d) \cdot f(t) = \sum_{t|n} f(t) \cdot \sum_{d|\frac{n}{t}} \mu(d).$$

If n and t are distinct, then $\frac{n}{t} \neq 1$ and thus:

$$\sum_{d|\frac{n}{t}} \mu(d) = 0.$$

Thus,

$$\sum_{d|n} \mu(d) \cdot F\left(\frac{n}{d}\right) = f(n).$$

□

[3]

Special Case for Multiplicative Functions

A special case of the Mobius Inversion formula for multiplicative functions is:

$$F(n) = \prod_{d|n} f(d),$$

then:

$$f(n) = \prod_{d|n} F\left(\frac{n}{d}\right)^{\mu(d)}.$$

Proof.

$$\prod_{d|n} F\left(\frac{n}{d}\right)^{\mu(d)} = \prod_{d|n} \left(\prod_{t|\frac{n}{d}} f(t) \right)^{\mu(d)} = \prod_{d|n} f(t)^{\mu(d)} = \prod_{t|n} f(t)^{\sum_{d|\frac{n}{t}} \mu(d)} = f(n),$$

since $\sum_{d|\frac{n}{t}} \mu(d) = 0$ if n and t are distinct and 1 if $n = t$.

□

4 Cyclotomic Polynomials

Consider the congruence

$$x^{\phi(n)} - 1 \equiv 0 \pmod{n}.$$

The solutions to this congruence are precisely the integers that are coprime to n , including the primitive roots modulo n . Now, look at the polynomial

$$P(x) = x^{\phi(n)} - 1.$$

Its roots are the $\phi(n)$ -th roots of unity, which are complex numbers. So how do integer solutions arise for this polynomial when considered modulo n ? Could it be that the primitive roots of $P(x)$ are actually the primitive roots modulo n ?

The answer is affirmative. For example, take $n = 7$. Then

$$P(x) = x^6 - 1 = (x - 1)(x + 1)(x^2 + x + 1)(x^2 - x + 1).$$

The roots of this polynomial are the 6th roots of unity. Among these, the roots of $x^2 - x + 1$ are the primitive 6th roots of unity. When considering modulo 7, the primitive roots modulo 7 are 3 and 5, which surprisingly also satisfy the congruence

$$x^2 - x + 1 \equiv 0 \pmod{7}.$$

Thus, this polynomial must be quite special—it is indeed the 6th cyclotomic polynomial.

4.1 Definition

The n -th cyclotomic polynomial $\Phi_n(x)$ is the monic polynomial whose roots are precisely the primitive n -th roots of unity. That is,

$$\Phi_n(x) = \prod_{\substack{0 \leq k \leq n-1 \\ \gcd(k,n)=1}} (x - \zeta_n^k),$$

where ζ_n denotes a primitive n -th root of unity.

The first few cyclotomic polynomials can be computed as follows:

$$\Phi_1(x) = x - 1,$$

$$\Phi_2(x) = x + 1,$$

$$\Phi_3(x) = x^2 + x + 1,$$

$$\Phi_4(x) = x^2 + 1,$$

$$\Phi_5(x) = x^4 + x^3 + x^2 + x + 1,$$

$$\Phi_6(x) = x^2 - x + 1,$$

$$\Phi_7(x) = x^6 + x^5 + x^4 + x^3 + x^2 + x + 1,$$

$$\Phi_8(x) = x^4 + 1,$$

$$\Phi_9(x) = x^6 + x^3 + 1.$$

[1]

4.2 Lemma

Lemma 4.1 — For any positive integer n , the degree of $\Phi_n(x)$ is $\phi(n)$.

Proof. This follows directly from the definition of $\Phi_n(x)$, using values of k that are coprime to n . □

Actually, computing $\Phi_9(x)$ by hand is not straightforward. Instead, a simpler method to determine the n -th cyclotomic polynomial is given by the following theorem:

Theorem 4.2 (Cyclotomic Polynomials Divide $x^n - 1$) — For any integer n , we have

$$x^n - 1 = \prod_{d|n} \Phi_d(x).$$

Thus, for a prime p , we have

$$x^{p-1} - 1 = \Phi_1(x)\Phi_p(x),$$

which gives

$$\Phi_p(x) = \frac{x^{p-1} - 1}{x - 1} = x^{p-2} + x^{p-3} + \cdots + x + 1.$$

This result comes from the fact that

$$x^n - 1 = (x - \zeta_n)(x - \zeta_n^2) \cdots (x - \zeta_n^n),$$

with terms grouped according to primitive roots of divisors of n .

Lemma 4.3 — For any positive integer n , the sum of $\phi(d)$ over all divisors d of n is n .

There is, however, a problem: most roots of cyclotomic polynomials are complex. Yet, as shown in smaller cases, $\Phi_n(x)$ surprisingly has integer coefficients. Is this merely coincidence? No, it is a general result, but it requires proof.

Theorem 4.4 — For all n , the n -th cyclotomic polynomial $\Phi_n(x)$ is an integer polynomial of degree $\phi(n)$.

Proof. We use induction. Base cases were previously established. Assume $\Phi_k(x)$ are integer polynomials for all $k < n$. Define

$$P_n(x) = \prod_{d|n, d \neq n} \Phi_d(x).$$

By the inductive hypothesis, $P_n(x)$ is an integer polynomial. Suppose $\Phi_n(x)$ is not an integer polynomial. Then there exist integer polynomials $Q(x)$ and $R(x)$ with $\deg R < \deg P_n = n - \phi(n)$, such that

$$x^n - 1 = P_n(x)Q(x) + R(x).$$

We need to show $R(x) = 0$. The polynomial $R(x)$ has degree $n - \phi(n)$, so it has exactly $n - \phi(n)$ roots $z_1, z_2, \dots, z_{n-\phi(n)}$. Since $R(z_i) = 0$, $R(x)$ must have these roots, leading to a contradiction because $R(x)$ cannot have more roots than its degree. Therefore, $R(x) = 0$ and $Q(x) = \Phi_n(x)$, which is an integer polynomial. \square

Another proof uses the following lemma, which is left as an exercise:

Lemma 4.5 — If f and g are polynomials with rational coefficients, and if all coefficients of $f \cdot g$ are integers, then the coefficients of f and g are also integers.

Using similar arguments, we can prove:

Lemma 4.6 — For all $n > 1$, $\Phi_n(0) = 1$.

Also, if ζ is a primitive n -th root of unity, then $\frac{1}{\zeta}$ is also a primitive n -th root of unity.

This leads to the following result:

Lemma 4.7 (Cyclotomic Polynomials are Symmetric) — For all $n \neq 1$,

$$\Phi_n(x) = x^{\phi(n)} \Phi_n\left(\frac{1}{x}\right).$$

The coefficients of $\Phi_n(x)$ are symmetric. These types of polynomials are known as reciprocals. Note that $\Phi_1(x)$ is not a reciprocal.

It can also be shown that all coefficients of $\Phi(x)$ are -1 , 0 , or 1 , but this is beyond the scope of this discussion.

Having defined cyclotomic polynomials, the relationship between primitive roots of unity and primitive roots modulo n should be quite intuitive: the primitive roots modulo n are solutions to the congruence

$$\Phi_{\phi(n)}(x) \equiv 0 \pmod{n}.$$

Expanding on this intuition, we have the following useful theorem:

Theorem 4.8

Let n be an integer and p be a prime number. If for any integer a , the condition

$$\Phi_n(a) \equiv 0 \pmod{p}$$

holds, then either a has order n modulo p (implying $n \mid (p-1)$), or $p \mid n$.

Proof. The condition

$$\Phi_n(a) \equiv 0 \pmod{p}$$

implies that

$$a^n - 1 \equiv 0 \pmod{p}.$$

Let m be the order of a modulo p . We consider two cases:

- If $m = n$, then a has order n modulo p , so $n \mid (p-1)$.
- If $m < n$, then

$$a^{m-1} \equiv 0 \pmod{p}.$$

Given that

$$a^{m-1} = \prod_{d \mid m} \Phi_d(a),$$

there exists a divisor $t \mid m$ for which

$$\Phi_t(a) \equiv 0 \pmod{p}.$$

Consequently, a is a double root of the polynomial

$$a^n - 1 = \prod_{d \mid n} \Phi_d(x)$$

modulo p , as $t \mid n$ as well. However, if a is a double root, it must also satisfy the first derivative $P'(x)$ of $P(x)$. Thus,

$$na^{n-1} \equiv 0 \pmod{p}.$$

Since $p \nmid a$, it follows that $p \mid n$.

□

Theorem 4.9

For any integer $n > 1$, the n -th cyclotomic polynomial can be expressed as:

$$\Phi_n(x) = \prod_{d|n} (x^d - 1)^{\mu\left(\frac{n}{d}\right)},$$

where μ is the Möbius function.

Corollary (Theorem 4.10)

Let p be a prime and n, k be positive integers. Then:

$$\Phi_{p^k n}(x) = \begin{cases} \Phi_n(x^{p^k}) & \text{if } p \mid n, \\ \frac{\Phi_n(x^{p^k})}{\Phi_n(x^{p^k-1})} & \text{if } p \nmid n. \end{cases}$$

Proof. If $p \mid n$:

$$\Phi_{p^k n}(x) = \prod_{d|p^k n} (x^d - 1)^{\mu\left(\frac{p^k n}{d}\right)}.$$

When $p^k \mid d$, we have:

$$\prod_{d|p^k n, p^k \mid d} (x^d - 1)^{\mu\left(\frac{p^k n}{d}\right)} = \prod_{d|n} (x^{p^k d} - 1)^{\mu\left(\frac{n}{d}\right)} = \Phi_n(x^{p^k}).$$

When $p^k \nmid d$, $\mu\left(\frac{p^k n}{d}\right) = 0$, which contributes nothing to the product.

If $p \nmid n$: When $p^k \mid d$:

$$\prod_{d|p^k n, p^k \mid d} (x^d - 1)^{\mu\left(\frac{p^k n}{d}\right)} = \prod_{d|n} (x^{p^k d} - 1)^{\mu\left(\frac{n}{d}\right)} = \Phi_n(x^{p^k}).$$

When $p^{k-1} \mid d$ but $p^k \nmid d$:

$$\prod_{d|p^k n, p^{k-1} \mid d, p^k \nmid d} (x^d - 1)^{\mu\left(\frac{p^k n}{d}\right)} = \prod_{d|n} (x^{p^{k-1} d} - 1)^{-\mu\left(\frac{p^{k-1} n}{d}\right)} = x^{p^{k-1}} \Phi_n(x^{p^{k-1}}).$$

For the remaining cases, $\mu\left(\frac{p^k n}{d}\right) = 0$, so these terms do not contribute.

Combining these results:

$$\Phi_{p^k n}(x) = \frac{\Phi_n(x^{p^k})}{\Phi_n(x^{p^k-1})}.$$

□

5 LTE and Other Important lemmas

Before diving into the proof of Zsigmondy's theorem, it's essential to understand some key facts.

Theorem 5.1: Let p be a prime number. If the polynomial $x^n - 1$ has a double root modulo p , meaning there exists an integer a and a polynomial $f(x) \in \mathbb{Z}[x]$ such that:

$$x^n - 1 \equiv (x - a)^2 f(x) \pmod{p},$$

then p divides n .

Proof: Observe that if f and g are polynomials where $f(x) \equiv g(x) \pmod{p}$ for all x , then $f'(x) \equiv g'(x) \pmod{p}$ for all x . This follows from comparing the coefficients of the two polynomials.

Taking the derivative of both sides of the given congruence, we get:

$$\frac{d}{dx}(x^n - 1) = nx^{n-1} \equiv 2(x - a)f(x) + (x - a)^2 f'(x) \pmod{p}.$$

Setting $x = a$, we have:

$$na^{n-1} \equiv 0 \pmod{p}.$$

Since $\gcd(a, p) = 1$ (as a satisfies $a^n - 1 \equiv 0 \pmod{p}$), it follows that p divides n .

Theorem 5.2 (Lifting The Exponent Lemma): Let x and y be integers, n a positive integer, and p an odd prime such that p divides $x - y$ and neither x nor y is divisible by p . Then:

$$v_p(x^n - y^n) = v_p(x - y) + v_p(n),$$

where $v_p(a)$ denotes the highest exponent k such that p^k divides a , often written as $p^k \parallel a$.

Proof. We will use induction on $v_p(n)$. First, let's establish the following claim:

$$v_p(x^p - y^p) = v_p(x - y) + 1.$$

To prove this, we need to show two things:

- p divides $x^{p-1} + x^{p-2}y + \dots + xy^{p-2} + y^{p-1}$.
- p^2 does not divide $x^{p-1} + x^{p-2}y + \dots + xy^{p-2} + y^{p-1}$.

Proof of (1): We start by noting that

$$x^{p-1} + x^{p-2}y + \dots + xy^{p-2} + y^{p-1} \equiv px^{p-1} \pmod{p}.$$

Since p divides px^{p-1} , it follows that p divides the entire sum.

Proof of (2): Let $y = x + kp$, where k is an integer. For $1 \leq t < p$, we have

$$y^t x^{p-1-t} \equiv (x + kp)^t x^{p-1-t}.$$

Expanding the right-hand side using the binomial theorem:

$$(x+kp)^t x^{p-1-t} \equiv x^{p-1-t} + t(kp)x^{p-1-t-1} + \text{higher order terms of } p \equiv x^{p-1} + tkpx^{p-2} \pmod{p^2}.$$

Thus,

$$y^t x^{p-1-t} \equiv x^{p-1} + tkpx^{p-2} \pmod{p^2}.$$

Using this result, we can compute:

$$x^{p-1} + x^{p-2}y + \cdots + xy^{p-2} + y^{p-1}.$$

Substituting in the values for $y^t x^{p-1-t}$ yields:

$$x^{p-1} + (x^{p-1} + kpx^{p-2}) + (x^{p-1} + 2kpx^{p-2}) + \cdots + (x^{p-1} + (p-1)kpx^{p-2}).$$

Combining these terms:

$$= px^{p-1} + (1 + 2 + \cdots + (p-1))kpx^{p-2}.$$

The sum of the integers from 1 to $(p-1)$ is $\frac{p(p-1)}{2}$, so:

$$= px^{p-1} + \frac{p(p-1)}{2}kpx^{p-2}.$$

Since $\frac{p(p-1)}{2}$ is divisible by p , but not necessarily by p^2 , we find:

$$x^{p-1} + x^{p-2}y + \cdots + y^{p-1} \equiv px^{p-1} \pmod{p^2}.$$

Thus, p^2 does not divide this sum.

Having established both (1) and (2), the proof of $v_p(x^p - y^p) = v_p(x - y) + 1$ is complete.

Returning to the original problem: We need to show:

$$v_p(x^n - y^n) = v_p(x - y) + v_p(n).$$

Assume $n = p^b$ where $\gcd(p, b) = 1$. Then:

$$v_p(x^n - y^n) = v_p((x^{p^b}) - (y^{p^b})).$$

Applying the result from the base case:

$$v_p((x^{p^b}) - (y^{p^b})) = v_p(x^p - y^p) = v_p(x - y) + 1.$$

Repeating this argument for b steps yields:

$$v_p(x^n - y^n) = v_p(x - y) + b.$$

Since $b = v_p(n)$, we have:

$$v_p(x^n - y^n) = v_p(x - y) + v_p(n).$$

Thus, the proof is complete [5].

Lemma 5.3: Let p be a prime, and $n = p^\alpha m$ where $p \nmid m$. The integer roots of Φ_n modulo p have order m modulo p .

Proof: Suppose a is an integer root of Φ_n modulo p . Then:

$$\Phi_n(a) \equiv 0 \pmod{p} \implies a^n \equiv 1 \pmod{p} \implies a^m \equiv 1 \pmod{p}.$$

Let $k = \text{ord}_p(a)$, so $k \mid m$. If $k = m$, the result holds. If $k < m$, since $p \nmid m$, from Theorem 4.10, we have:

$$\Phi_{p^\alpha m}(x) = \Phi_m(x^{p^\alpha}).$$

Given $p \mid \Phi_n(a)$, it follows that $p \mid \Phi_m(a)$. Since

$$x^k - 1 = \prod_{d \mid k} \Phi_d(x),$$

there exists some d such that $d \mid k$ and $p \mid \Phi_d(a)$. Therefore, $p \mid \Phi_m(a)$, making a a double root of $x^m - 1 \pmod{p}$. By Theorem 5.1, $p \mid q$, which is a contradiction.

Lemma 5.4: For any positive integer n and real number $x > 1$, we have:

$$(x - 1)^{\varphi(n)} \leq \Phi_n(x) < (x + 1)^{\varphi(n)},$$

with equality holding only if $n = 2$.

Proof: Using the triangle inequality for complex numbers, $x - 1 \leq |x - \zeta| < x + 1$ for any complex number ζ with $|\zeta| = 1$. The first inequality is strict unless $\zeta = 1$, and the second is strict unless $\zeta = -1$. Applying this, we get:

$$(x - 1)^{\varphi(n)} \leq \prod_{\zeta \text{ primitive}} |x - \zeta| < (x + 1)^{\varphi(n)},$$

with equality only if $\varphi(n) = 1$, which means $n = 2$. Note that the second inequality is always strict because $|x - 1| < |x + 1|$. The product in the middle is $|\Phi_n(x)|$. Since $x > 1$, by Theorem 4.9, $\Phi_n(x) = |\Phi_n(x)|$, thus the relation follows.

6 Combining the Results and the Proof

Let's now prove Zsigmondy's Theorem. Recall its statement:

Theorem 6.1 (Zsigmondy's Theorem): Let a and b be two co prime integers, and let $n > 1$ be another integer. There is a prime divisor of $a^n - b^n$ that does not divide $a^k - b^k$ for any $k \in \{1, 2, \dots, n - 1\}$, except in the following cases:

- $a = 2$ and $b = 6$
- $n = 2$ and $a + b$ is a power of 2.

Such a prime divisor is called a primitive prime divisor of $a^n - b^n$.

To prove this, it suffices to show that $a^n - b^n$ has a prime divisor that does not divide $a^k - b^k$ for $k \mid n$. If p divides both $a^k - b^k$ and $a^n - b^n$ and $k < n$, then it can be shown that k must divide n .

Consider the case of $x^n - 1$. By relating cyclotomic polynomials to Zsigmondy's theorem, we note that cyclotomic polynomials state that for $x^n - 1$, there is a root that is not a primitive root of unity for any $k < n$. We aim to establish a connection between $x^n - 1$ and $a^n - b^n$. Define:

$$z_n = a^n - b^n = b^n \left(\left(\frac{a}{b} \right)^n - 1 \right).$$

Substituting $\frac{a}{b}$ into the equation, we obtain:

$$\left(\frac{a}{b} \right)^n - 1 = \prod_{d \mid n} \Phi_d \left(\frac{a}{b} \right),$$

where Φ_d denotes the d -th cyclotomic polynomial. Multiplying both sides by b^n , we have:

$$z_n = b^n \left(\left(\frac{a}{b} \right)^n - 1 \right) = \left(\frac{a}{b} \right)^n - 1 \prod_{d \mid n} b^{\phi(d)} = \prod_{d \mid n} b^{\phi(d)} \Phi_d \left(\frac{a}{b} \right).$$

Define:

$$\Psi_n = b^{\phi(n)} \Phi_n \left(\frac{a}{b} \right),$$

so:

$$z_n = \prod_{d \mid n} \Psi_d.$$

By Möbius inversion, we can compute Ψ_n :

$$\Psi_n = \prod_{d \mid n} z_{\mu(n/d)}.$$

Here, Ψ_n divides z_n .

Consider $z_n = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$. Let $p_{b_1}, p_{b_2}, \dots, p_{b_t}$ be the primitive prime divisors of z_n , i.e., p_{b_i} does not divide any z_k for $k < n$. Define:

$$P_n = p_{b_1}^{b_1} p_{b_2}^{b_2} \cdots p_{b_t}^{b_t},$$

which consists of all powers of the primitive primes of z_n and P_n divides Ψ_n . Let $\Psi_n = P_n \lambda_n$. Note that $\gcd(\lambda_n, P_n) = 1$ because all prime divisors of P_n are primitive prime divisors of z_n and do not divide any z_d for $d < n$. We will show that $P_n > 1$ in cases where Zsigmondy's theorem does not apply.

6.2 An Upper Bound on λ_n

We need to show that $\lambda_n < \Psi_n$. To do this, we first show that if p divides λ_n , then p divides n , and then find a bound for the powers of these primes.

If $p = 2$, then $2 \mid \Psi_n \Phi_n \left(\frac{a}{b}\right)$, and by Theorem 4.8, $p \mid n$. For odd primes p , since p is not a primitive prime divisor of z_n , there exists $d < n$ such that $p \mid z_d$. Since $p \mid \Psi_n/z_d$, we need:

$$v_p \left(\frac{z_n}{z_d} \right) = v_p(z_n) - v_p(z_d) = v_p \left(\frac{n}{d} \right) > 0$$

by the Lifting The Exponent Lemma, which implies $p \mid n$.

Assuming $\lambda_n > 1$, we show that λ_n is a prime for $n > 2$. We first prove that λ_n can have only one prime divisor and then show that the power of that prime must be 1.

Write $n = p^s m$. From the proof of Lemma 5.3, we have:

$$\Psi_n \equiv \Phi_n \left(\frac{a}{b} \right) = \Phi_m \left(\left(\frac{a}{b} \right)^{p^s} \right) \equiv \Phi_m \left(\frac{a}{b} \right) \equiv \Psi_m \equiv 0 \pmod{p}.$$

Theorem 4.8 also gives that $p \equiv 1 \pmod{m}$, implying $p > m = \frac{n}{p^s}$. If q is another prime divisor of λ_n , then $q < \frac{n}{p^s}$, making $q^t \mid n$ where t must be less than p^s , which does not hold. Thus, λ_n has only one prime divisor, which is the largest prime divisor of n .

Now, if $\lambda_n = p^t$, consider two cases: $p = 2$ and $p > 2$. If $p = 2$, then n must be a power of 2, and $\Psi_n = a^2 + b^2$. Since 4 does not divide $x^2 + y^2$ for odd x and y , it follows that $2 \mid \Psi_n$. For $p > 2$, let $d < n$ and $p \mid z_d$. Write $n = p^t m$. By Lemma 5.3, m is the order of $\frac{a}{b}$ modulo p . Thus, p divides $z_m, z_{mp}, z_{mp^2}, \dots, z_{mp^t}$. The only contributions to p in Ψ_m come from z_{mp^t} and $z_{mp^{t-1}}$. Thus:

$$v_p(\Psi_n) = v_p(z_{mp^t}) - v_p(z_{mp^{t-1}}) = v_p(p) = 1$$

implying $p \mid \lambda_m$.

For $n = 2$, $a^2 - b^2 = (a - b)(a + b)$ has a primitive prime divisor unless $a + b$ is a power of 2, which is excluded in the theorem statement.

6.3 A Lower Bound on P_n

Consider the following cases:

1. **If $\lambda_n = 1$:** Then $P_n = \Psi_n = b^{\phi(n)} \Phi_n \left(\frac{a}{b} \right) \geq b^{\phi(n)} (a - b)^{\phi(n)} = (a - b)^{\phi(n)}$ by Lemma 5.4. This inequality is strict unless $n = 2$ and $a - b = 1$. Otherwise, since $a > b$, $P_n > 1$.
2. **If $\lambda_n = p$ and $a - b > 1$:** In this scenario:

$$P_n = \frac{1}{p} \Psi_n \geq \frac{1}{p} (a - b)^{\phi(n)} \geq \frac{2^{p-1}}{p} \geq 1.$$

This inequality is strict unless $p = 2$, $a - b = 2$, and $n = 2$, which has been previously addressed.

3. **If $\lambda_n = p$ and $a - b = 1$:** Since the previous inequality does not hold, assume $P_n = 1$ and $\Psi_n = p$. This leads to the remaining counterexample where $n = 6$ and $a = 2$. For $a - b = 1$, z_n is odd, so p must be odd. Let $n = p^t m$. If $t > 1$, by Theorem 4.10:

$$p = \Psi_n = b^{\phi(n)} \Phi_n \left(\frac{a^{p^{t-1}}}{b^{p^{t-1}}} \right).$$

However:

$$b^{\phi(n)} \Phi_{pm} \left(\frac{a^{p^{t-1}}}{b^{p^{t-1}}} \right) \geq b^{\phi(n)} \left(\frac{a^{p^{t-1}}}{b^{p^{t-1}}} - 1 \right)^{\phi(pm)} \geq a^p - b^p.$$

Since:

$$a^p - b^p = (b + 1)^p - b^p = \sum_{k=0}^{p-1} \binom{p}{k} b^k > p,$$

it follows that $\Psi_n > p$, which is not possible. Hence, $p = 3$, and the valid cases are $n = 3$ and $n = 6$. For $n = 3$, the theorem holds because $a^3 - b^3 = (a - b)(a^2 + ab + b^2)$ with $a - b = 1$. For $n = 6$, Zsigmondy's theorem fails, with the counterexample $n = 6$, $a = 2$, $b = 1$.

This completes the proof.

References

- [1] Andrew Arnold and Michael Monagan. "Calculating cyclotomic polynomials". In: *Mathematics of Computation* 80.276 (2011), pp. 2359–2379.
- [2] Walter Ledermann. *Complex Numbers*. Springer Science & Business Media, 2013.
- [3] Carl Lienert. "The Mobius Function and Mobius Inversion". In: (2020).
- [4] Bart Michels. *Zsigmondy's Theorem*. 2014.
- [5] Amir Hossein Parvardi. "Lifting the exponent lemma (LTE)". In: (2011).

Educational Disparities in Urban and Rural India and Its Impact on Economic Inequality By Akaanksha Chintakindi

Abstract

In the 21st century, India is growing and developing faster than ever, but the gap between urban and rural areas continues to widen. Differences in jobs, development, access to resources, and education are widening, and understanding how to reduce these disparities can have a big impact on improving sustainability and economic fairness. Since various factors contribute to the urban-rural divide, this paper focuses on the role educational disparities play in affecting economic inequality. The paper uses different measures of educational disparity, like gender enrolment ratios, student-teacher ratios, dropout rates, and income growth, using data from government sources. The study also examines factors that are harder to measure, such as the quality of school infrastructure and access to resources like books and technology, by using survey data to compare conditions in urban and rural areas.

Although the research does not consider all the quantifiers of economic inequality as the data available might be limited or not updated, it does address the most prominent indexes that have an impact on economic inequality due to educational disparities. To understand further on how these disparities can be curbed, the research focuses on examining what causes the education disparities: population, income level, migration pattern, gender discrimination and stereotypes, and life expectancy.

Keywords: Education, Economic Inequality, Sustainability, Educational disparities, Urban literacy, Rural literacy, Urban-rural differential literacy index.

Introduction

The rural urban disparity is an estimation of standard of living calculated per capita consumption or use of non-durable, durable consumer goods, consumption of electricity, healthcare, and other living facilities by the population of rural and urban areas of the country¹. While there are many types of disparities underlying as mentioned above, this study focuses on one aspect of the urban rural disparity: **educational disparities**.

Educational disparities refer to the unequal distribution of resources and opportunities among the students². Educational disparities occur in various forms throughout the world: the students' origin, their family background and language, the students' social class, and in some countries even based on their gender. The illiteracy rate is much higher in the underdeveloped and developing countries, like 31 percent in Sub-Saharan Africa³. Although some policies have been introduced to reduce the gap in access to education between rich and poor, the differences still exist, and the inequality for educational access between rural and urban areas has not diminished³.

The disparity in access to education between urban and rural areas worsened during the pandemic, as schools and colleges switched to online learning³. Many students in rural areas didn't have access to the necessary technology, which made it

difficult for them to continue their education, unlike students in urban areas who could participate in online classes. India, like many other developing countries, also experienced a significant increase in educational disparities during the pandemic. While urban schools adapted to online classes, rural school attendance dropped, with the dropout rate rising from 1.8% in 2018 to 5.3% in 2020⁴. This lack of technology for learning is only one aspect of a much larger issue driving the differences in access to education between urban and rural areas.

Even though the rural economy in India makes up 46% of the national income⁵, access to education in these areas still lags behind. One reason for this is that many girls in rural areas get married young, which stops them from pursuing higher education. The literacy rate for women in rural areas is just 70%, compared to 84.7%⁶ for men. Poor infrastructure in rural and government schools, lower household incomes, and health issues also affect the quality of education for students in these areas.

Understanding how education contributes to a country's economic growth is key to addressing issues like economic inequality and promoting sustainability. In India, there is a significant gap in wealth between urban and rural areas. The top 10% of the population holds 77% of the total national wealth, and 73% of all new wealth generated has gone to the richest individuals, while the remaining millions of people have only seen around a 1% increase in their wealth⁷. To reduce this gap and promote income equality, it's important to address the differences in education between rural and urban areas. Reducing these differences won't just help balance income inequality; it will also boost economic growth, as education is an investment in the development of human potential⁸. As of 2022, India has an overall literacy rate of 77.70%⁶, with urban India having 87.7% greater than that of rural India's 73.5%⁹. According to the world bank, increasing the average year of schooling by just one year can boost GDP growth by 0.37%, and raising literacy rates by 1% can lead to 0.3%⁸ increase in GDP growth rate.

Focusing only on educating urban areas, while neglecting rural regions, can do more harm than good. As people in rural areas are left behind, they tend to migrate to cities¹⁰ in search of better opportunities, which can eventually leave rural areas underdeveloped and desolate. High-quality education is just as essential for rural areas because it helps break the cycle of poverty, equips local businesses with skilled workers¹¹, and enhances agricultural practices through knowledge of sustainable and modern farming techniques. Improving education in rural regions boosts local economies and allows students to pursue higher education without needing to leave their communities. Curbing educational disparities in rural areas would not only foster development but also reduce economic inequality. While access to education in urban areas drives economic growth, the lack of access for most people living in rural areas results in low incomes, poor healthcare, and inadequate infrastructure. Addressing the urban-rural education divide is crucial for reducing economic inequality and ensuring balanced growth across the country.

Literature Review

This research paper explores one key aspect of the urban-rural divide: educational disparity and its significant impact on economic inequality in India. As part of this research, four key studies have been reviewed: *IFAD Research Series 27: Asia's Rural-Urban Disparity in the Context of Growing Inequality* (S.K. Imai, 2018), *Exploring Rural–Urban Education Divide in India* (Somprakash Bandyopadhyay, 2021), *Urban Rural Disparity in Education in India* (Sarthak Chakrabarti, 2019), and *Inequality and Imbalances in the Education System of India: Insights from NSSO 71st Round* (Susmita Mitra, 2020).

The *IFAD Research Series 27* (S.K. Imai, 2018) focuses on the extreme rural-urban disparities in Asia, particularly in China and India, which contribute to regional inequality. In contrast, countries like Thailand and Vietnam have made progress in narrowing their income gap. The study highlights how limited access to education in rural areas slows down their development and increases economic inequality.

Exploring Rural–Urban Education Divide in India (Somprakash Bandyopadhyay, 2021) looks specifically at India and explores the causes and effects of unequal access to education between urban and rural areas. The study emphasizes the socio-economic and political barriers in rural areas that prevent equal access to education. It also discusses the efforts being made to address these gaps and provide equal educational opportunities across the country.

The study *Urban Rural Disparity in Education in India* (Sarthak Chakrabarti, 2019) investigates how the educational divide between urban and rural areas affects economic outcomes and social mobility. It examines how factors like infant mortality rates and life expectancy influence educational standards in both regions.

Finally, *Inequality and Imbalances in the Education System of India* (Susmita Mitra, 2020) analyses data from the National Sample Survey Organization's 71st round. It explores broader issues of inequality in education, including gender disparities and the urban-rural divide. The study provides valuable insights into the complex nature of educational inequality in India and highlights the major challenges involved.

While these studies provide a broad understanding of educational disparities, they often focus on multiple aspects of inequality, which can lead to generalized conclusions about issues like economic inequality. Although most of these studies focus specifically on education disparities within India, they also consider other factors such as mortality rates and gender inequality, which broadens their scope. For instance, the *IFAD Research Series* addresses extreme rural-urban disparities across Asia, including in China, making its findings more applicable to global trends.

However, this research provides specific insights into the urban-rural divide in education, concentrating on a few key indicators such as income growth, economic inequality, and literacy rates. It aims to offer a well-defined analysis of how

educational disparities affect economic outcomes in India and how the gap between urban and rural education can be addressed.

Methodology

The research was conducted using both primary and secondary data sources. The educational disparities index was developed based on data from updated Indian government portals. Key indicators such as literacy rates in urban and rural areas, dropout and enrolment rates, income levels, and employment rates were obtained from official government websites. To further investigate the extent of economic inequality caused by disparities in educational access, additional data were collected through a nationwide survey.

Primary data, along with economic indexes from government sources, were analyzed using statistical tools to identify correlations between economic inequality indicators—such as unemployment rates and average household incomes in urban versus rural areas—and educational disparities, including gender enrolment gaps and educational levels in both regions.

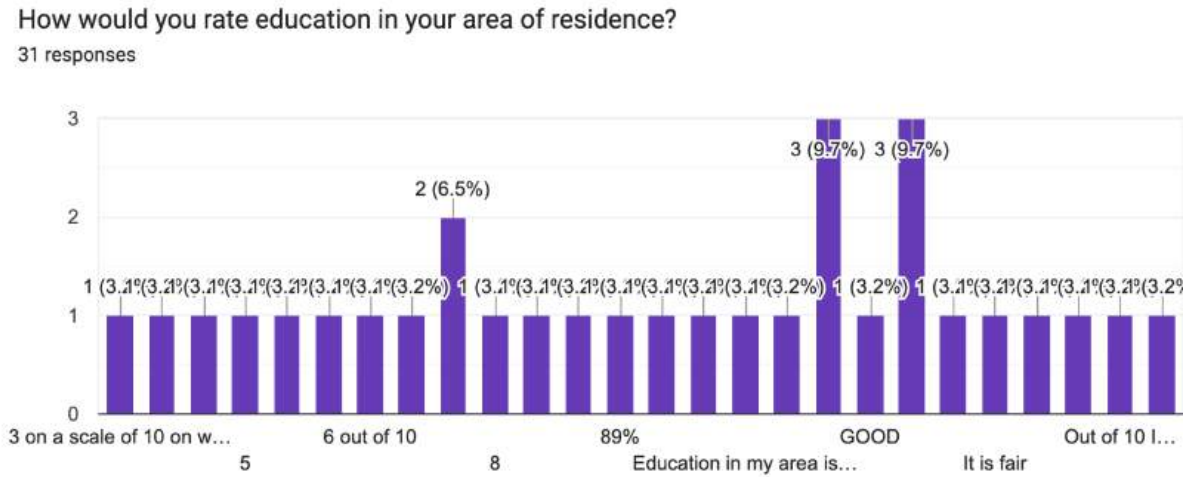
In addition, a thorough literature review was conducted to explore the role of educational disparities in sustainability and economic inequality. This included analyzing public survey data, as well as examining educational programs and initiatives implemented by the government and other institutions.

The findings were presented through visual aids, such as graphs, to provide a clear representation of the data collected. Both quantitative and qualitative data were employed to draw conclusions about the impact of educational disparities on the Indian economy and progress toward the Sustainable Development Goals (SDGs). Conclusions were derived from statistical and comparative analyses of educational access, alongside global findings, to ensure the validity of the study. Limitations due to constraints in some economic indicators were also acknowledged.

Data and Results

Data collected from a survey of 31 participants in India were analyzed to explore the correlation between education and economic equity. Of the participants, 87.1% were from urban areas and 12.9% from rural areas. The majority of respondents, 58.1%, held an undergraduate degree, while 41.9% had a postgraduate degree.

Figure 1

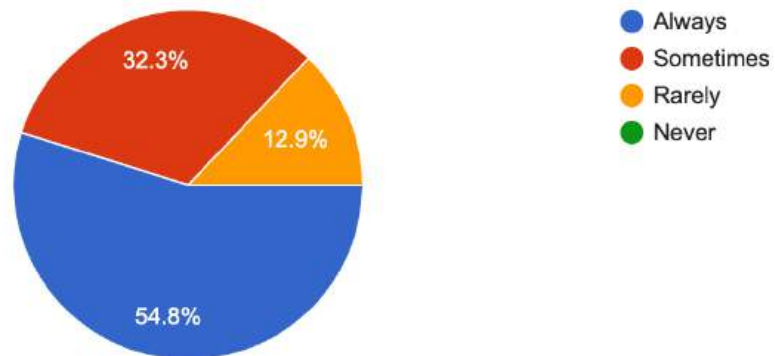


The bar chart above displays respondents' opinions on the quality of education in their area of residence. Out of 31 people, 13 rated education on a scale of 10, resulting in an average score of 6.04. A majority of respondents, around 70.9%, rated education as either average or poor, particularly with regard to access to resources. Even when education quality was considered good, many respondents noted that high fees prevented lower-income families from pursuing it, highlighting a major sustainability issue—economic inequality.

Figure 2

Do the schools in your area have sufficient access to education?

31 responses



The pie chart above highlights another key factor affecting educational disparities—sufficient access to education and resources. While a majority of respondents indicated that they “always” have access, a significant portion, 45.2%, reported that they “rarely” or only “sometimes” have sufficient access to education. This is concerning, as it suggests that even in urban areas, inequalities in access to education persist, and the issue is even more pronounced in rural regions.

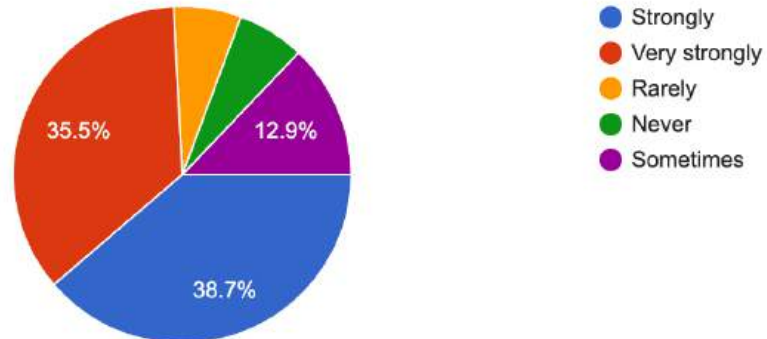
When asked what schools were lacking for proper access to education, most respondents pointed to inadequate infrastructure, poorly trained teachers, outdated technology, limited resources, and hygiene issues. Some respondents from rural areas raised concerns about the lack of knowledgeable teachers and ineffective teaching methods for introducing new concepts. This can be attributed to the fact that teachers in rural India often have limited opportunities for professional development and are frequently underpaid, which can reduce their motivation as they struggle to meet their daily expenses¹². India faces a shortage of nearly 11.16 lakh teachers to meet current educational needs¹³. According to the survey, 35.5% of respondents felt that the lack of qualified teachers was a major issue in their area.

Additionally, several participants noted that financial difficulties prevent some students from pursuing higher education. This was further supported by the survey results, which showed that 58.1% of respondents knew people who had to drop out of school due to financial constraints. Even when quality education is available, it remains out of reach for lower-income households, which, in turn, affects an individual’s earning potential.

Figure 3

In your opinion, how does the level of education in your area affect employment opportunities?

31 responses



The figure above shows that 74.2% of respondents believe that their level of education "strongly" or "very strongly" influences their job opportunities, with only 6.5% feeling otherwise. Additionally, 74.2% of respondents expressed that receiving education beyond the bare minimum would have significantly increased their earning potential. Respondents noted that better education would have provided them with improved guidance, stronger validation of their knowledge, and the ability to overcome societal barriers and expectations. As one respondent put it, "It was 10 times better than what I had earlier with only a graduation for employment opportunities."

Urban Rural differential literacy Index (URDLI) analysis

The Urban-Rural Differential Literacy Index (URDLI) measures the gap in literacy rates between rural and urban areas. A higher URDLI indicates greater disparities, while a lower index reflects fewer disparities between the two regions. The formula used to calculate the URDLI is:

$$\text{URDLI} = (\text{LU} - \text{LR}) / \text{LU} * 100$$

- LU- represents urban literacy
- LR- represents rural literacy

According to the Ministry of Education in India, the rural literacy rate was 67.77%, and the urban literacy rate was 84.11% in 2023¹⁴. Using these figures, the URDLI is calculated as follows:

$$(84.11 - 67.77) / 84.11 * 100 = 19.43\%$$

This result shows that literacy rates in rural areas of India are 19.43% lower than those in urban areas.

Discussion

The educational disparities in India, as identified through both primary and secondary data, are influenced by several key factors. These factors—such as infrastructure, access to resources, migration, and their impact on economic inequality—play a critical role in determining the quality of education and the resulting socio-economic outcomes.

Infrastructure

School infrastructure plays a crucial role in providing quality education. Many rural schools in India lack essential facilities such as proper buildings, desks, boards, and clean drinking water. Basic necessities like bathroom sanitation, infirmaries for emergencies, and electricity are often absent. Even when infrastructure is available, it may not be functional. According to research, only 80% of schools meet the infrastructural standards under the Right to Education (RTE) Act. Poor sanitation, particularly the lack of girls' washrooms, is a major factor contributing to dropout rates among young girls¹⁵. These issues are not limited to rural areas; some government schools in urban areas face similar challenges. While private and public schools in urban areas benefit from better infrastructure, funding, and curriculum, many families cannot afford to send their children to these schools, relying instead on government schools that lack these basic resources.

The inadequate infrastructure in government and rural schools negatively impacts students' learning outcomes. Students spend a significant portion of their day in school, and without proper resources, they are likely to feel demotivated, leading to decreased attendance. This is especially true for girls, who may miss school during menstruation due to the lack of hygiene products and proper bathroom facilities, causing them to miss out on valuable learning opportunities. In the survey, 11 out of 31 respondents identified poor infrastructure and hygiene as key barriers to accessing education. Even when facilities exist, they are often non-functional.

India's diverse climate further exacerbates these challenges. Schools should be equipped to operate during extreme weather conditions, but inadequate infrastructure often results in school closures during heatwaves or heavy rains, disrupting students' productivity. Attending school during these periods can also expose students to illness, leading to medical conditions that their families may struggle to afford.

In contrast, private and well-funded public schools, typically located in urban areas, do not face these challenges. Students in these schools can focus solely on their academics, as they have access to better facilities and resources. This disparity in infrastructure between urban and rural schools creates a significant gap in the educational experiences of students from these regions.

Access to resources & Migration

Access to resources is a crucial factor in shaping a student's educational experience. Forbes India reports that around 4.5-5% of the population lives in poverty¹⁶, with 31% in the middle class, projected to reach 60% by 2047¹⁷. Education in India is often expensive, and many families cannot afford the necessary resources

for their children's education. Rural and government schools lack funding to provide textbooks, technological access, and qualified teachers. Teachers often prefer to work in private or well-funded public schools, where salaries are higher, leaving rural schools with fewer, less-qualified teachers.

The COVID-19 pandemic highlighted existing educational inequalities in India¹⁸. During the lockdown, while urban students adapted to online learning, students in rural areas fell behind. Enrollments dropped by 3.55%, and a cross-country analysis showed that 29.6 million children lacked access to digital devices¹⁸. The pandemic also exacerbated gender disparities, with girls being given more household chores, further limiting their access to online education.

In the survey, 19 out of 31 respondents expressed concerns about gaps in digitalization, teacher shortages, and lack of access to study materials. Study materials are critical for understanding subjects and performing well on exams, but they are often paid resources that schools or families cannot afford, resulting in students falling behind academically.

The unequal access to education drives migration patterns, whether within a state, across states, or even out of the country, further worsening conditions in rural areas. Rural India is home to the majority of the population, but these areas are underdeveloped, leading to migration to urban areas for better living conditions, such as healthcare, employment opportunities, education, and overall economic well-being. In an effort to escape poverty, many people move permanently to urban areas. Every minute, around 25-30 people migrate from rural areas to urban centers, and if this trend continues, the population in urban India is expected to reach 600 million by 2030. According to the Economic Survey 2016-17, an average of 90 million people migrated annually between Indian states via railways over the last five years¹⁹. While migration may provide access to better facilities, it leaves rural areas increasingly desolate, with little room for improvement, while urban areas become overcrowded, creating unsustainable growth and increasing inequality.

The government has launched various programs to ensure that the situation does not worsen, with different states implementing their own approaches to improve education. For instance, in Andhra Pradesh, former Chief Minister Jagan Mohan Reddy introduced a 9-tier plan, one of which focuses on increasing access to education through the "Jaganna Amma Vodi" scheme. Under this scheme, eligible families receive 15,000 rupees annually to send their children to school up to grade 12²⁰. Many respondents in the study also emphasized that reducing educational disparities requires more government support, with a stronger focus on education funding and providing skill development training for teachers. This highlights the critical role of government intervention in addressing these disparities.

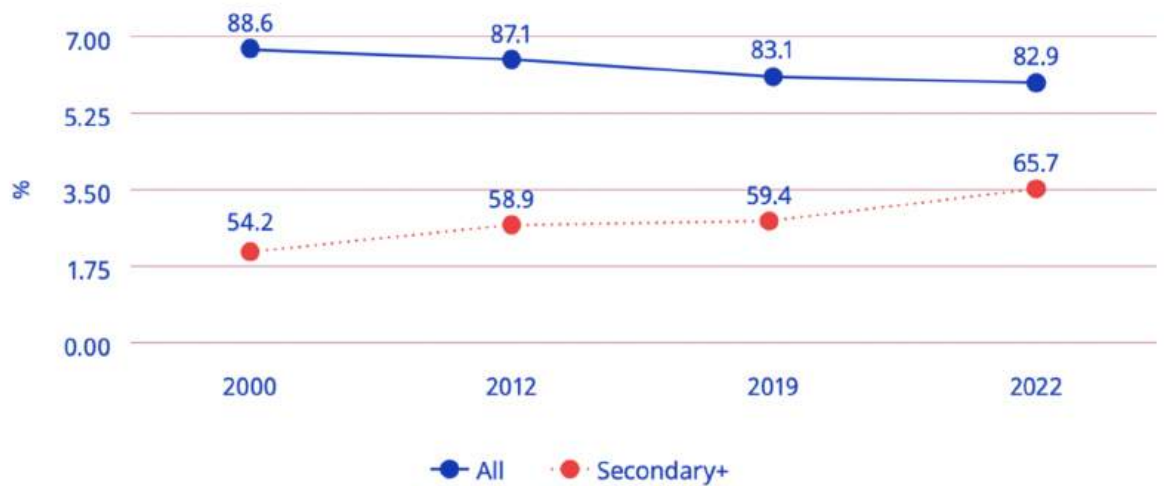
Impact of Educational Disparities on Economic inequality

Educational disparities have long-term, multifaceted impacts on a country's economy. Education is often seen as a pathway to a better life, offering a chance to break the cycle of poverty. For people in marginalized and low-income areas, education represents hope for a better future, equity, and economic well-being.

Without access to education, individuals often face social exclusion and are marginalized within their communities.

Stable employment is crucial for leading a secure life. Jobs that offer decent wages typically require specific skill sets, and without those skills, people are forced into informal sector jobs. In fact, around 90% of workers in India remain in informal employment²¹, taking jobs that keep them in poverty, unable to adequately support themselves or their families. Educational disparities contribute to this skill gap. A report noted that many employees lack basic skills, with 75% of youth unable to send emails with attachments, 60% unable to copy and paste files, and 90% unable to input a mathematical formula into a spreadsheet²¹.

Figure 5



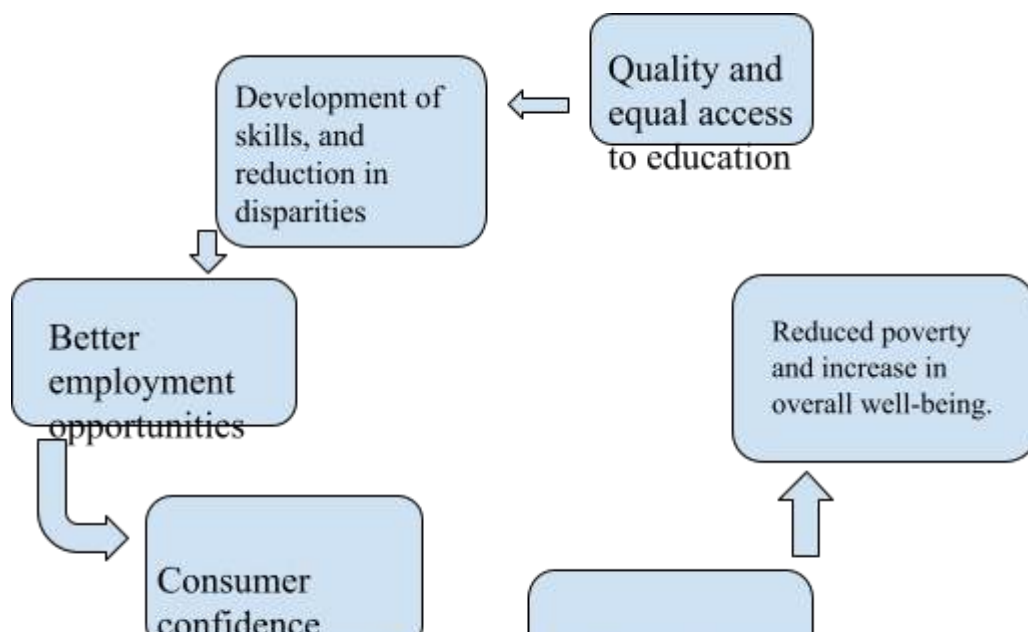
Source: Computed from Employment and Unemployment Survey data, Periodic Labour Force Survey data and Central Statistical Office data.

The figure above illustrates the high rates of youth unemployment in India in relation to secondary education levels. According to 2022 reports, youth unemployment rose significantly, reaching around 82.9%²².

This growing unemployment highlights the importance of developing the right skills, which can only be achieved through proper education. Without adequate education, an economy faces structural unemployment—where individuals’ skills are not transferable to new jobs or do not match the needs of employers. The longer people experience structural unemployment, the more discouraged they become, making it harder for them to re-enter the job market. In the current era of artificial intelligence, firms increasingly seek employees with technical skills, management expertise, and specialized knowledge. Due to educational disparities, individuals who lack access to quality education and practical experience with technology are more likely to remain unemployed. In fact, the International Labour Organization (ILO) reports that 53% of businesses in India struggle to recruit workers because of the skill gap²³.

This unemployment caused by the skill gap leads to a range of social, personal, and economic consequences. People who are unable to find employment often experience stress, a decline in their well-being, and a lower standard of living, as well as social exclusion²⁴. The government is also impacted, as reduced employment means lower tax revenues and increased spending on welfare, which can result in borrowing and budget deficits. Additionally, unemployment reduces consumption in the economy, lowering consumer confidence and decreasing aggregate demand, which negatively affects the GDP—Gross Domestic Product—the total value of goods and services produced by the country. This, in turn, slows economic growth. Furthermore, unemployment exacerbates economic inequality and poverty, both of which act as barriers to economic development.

Figure 6



As shown in figure 6 above, without equal access to education, achieving most of the Sustainable Development Goals (SDGs)—such as quality education, no poverty, decent work and economic growth, gender equality, and reduced inequalities—would be nearly impossible in the Indian economy. These factors are all interconnected. Without the skills gained from education, it becomes difficult to find decent-paying jobs, which in turn leads to increased poverty and inequality, ultimately hindering economic growth.

When survey respondents were asked, “How do you think improved education would impact economic sustainability in your region, and in what aspects of sustainability?” the majority indicated that improved education would have a profound impact on economic growth. They emphasized that it would promote sustainable development by raising awareness and, most importantly, significantly improve employment rates.

Conclusion & Limitations

Educational disparities play a key role in contributing to economic inequality in India. Efforts by the government and support from NGOs can help improve access to education for children. This paper analyses the role of education in India by collecting data from citizens and identifying the causes of these disparities. The findings suggest that increasing the skill set provided by education leads to a rise in the Labor Force Participation Rate (LFPR) and positively impacts the country’s economic development. The discussion section identifies several factors that contribute to educational disparities and their effects on economic inequality.

Although educational disparities are a global issue, the results of this study are specific to India. While the connection between education, economic activity, and employment is generally applicable, the statistics and data in this study are focused only on India and cannot be generalized for other countries. Since the factors contributing to educational disparities in India are broad and difficult to measure comprehensively, this research focused on a few key areas—such as infrastructure, resources, and migration—to provide an overview of how these specific factors contribute to educational disparities and their impact on economic inequality.

Due to limitations in the available data, only some causes of educational disparities were considered in this study. While there are many possible factors linking educational disparities to economic inequality, this research specifically examined how poor infrastructure, lack of resources, funding issues, and rural-to-urban migration contributed to these disparities in India. However, it is important to recognize that there may be other factors contributing to educational disparities that were not addressed in this study. Since addressing all these factors would have been too broad, this research focused on a specific set of issues to better understand the causes of educational disparities.

Additionally, the study demonstrates that educational disparities directly contribute to economic inequality. Limited access to quality education reduces

opportunities for skill development, leading to lower employment prospects. This, in turn, deepens economic inequality, which further restricts access to education, creating a vicious cycle that is difficult to break.

CITATIONS

1. Economic and Political Weekly, H S Shergill, **Rural–Urban Disparity in the Standard of Living across States of India** | **Economic and Political Weekly**, Vol 56, Issue No 45-46, 06 November 2021, Published January 17,2024.
<https://www.epw.in/journal/2021/45-46/special-articles/rural-urban-disparity-standard-living-across.html>
2. Study.com, Yuanxin(Amy) Yang Alcocer **Educational Inequality | Types & Examples**
<https://study.com/academy/lesson/educational-inequality-types-examples.html>
3. United Nations Chronicle, Sylvia Schmelkes **Recognizing and Overcoming Inequity in Education**, January 22,2020.
<https://www.un.org/en/un-chronicle/recognizing-and-overcoming-inequity-education>
4. Times of India, Shatana Sinha, **How Covid-19 changed India’s education system**
<https://timesofindia.indiatimes.com/blogs/developing-contemporary-india/how-covid-19-changed-indias-education-system/>
5. The Samiksha, **Educational Inequality in Rural and Urban India**, July 21,2023.
<https://thesamiksha.com/educational-inequality-in-rural-and-urban-india/>
6. Kritish Swargiary, Kavita Roy, **Literacy Rate in India In 2022 , July 2024**
https://www.researchgate.net/publication/363173220_LITERACY_RATE_IN_INDIA_IN_2022
7. Oxfam International, **India: extreme inequality in numbers**
<https://www.oxfam.org/en/india-extreme-inequality-numbers>
8. Nilangini Bansal, Invest India, **Impact of Education Policy on India’s GDP growth: Strengthening the Foundation for a Thriving Nation** October 26,2023
<https://www.investindia.gov.in/team-india-blogs/impact-education-policy-indias-gdp-growth-strengthening-foundation-thriving-nation>
9. IJFMR, Dipanjolly Nath, **Literacy Rate in India -2022**
<https://www.ijfmr.com/papers/2023/1/1409.pdf>
10. Times of India, **Education is intangible for rural development December 11, 2022**
<https://timesofindia.indiatimes.com/readersblog/analyzing/education-is-intangible-for-rural-development-47892/>
11. Elysium Foundation, **Importance of education in rural development**
<https://elysiumfoundation.org.in/importance-of-education-in-rural-development/>
12. Daily Excelsior **Education Inequality in Rural India: Challenges and Solutions**, July 20,2024
<https://www.dailyexcelsior.com/education-inequality-in-rural-india-challenges-and-solutions/>

13. Indian Express **India needs 11.16 lakh extra teachers to meet shortfall: report**, October 6, 2021
<https://indianexpress.com/article/india/india-needs-11-16-lakh-extra-teachers-to-meet-shortfall-report-7554560/>
14. NDTV Education, **International Literacy Day 2023: Know About Remarkable Progress Made By India** September 08, 2023
<https://www.ndtv.com/education/international-literacy-day-2023-know-about-remarkable-progress-made-by-india-4370870>
15. Center for Budget and Governance Accountability (CBGA) India, **Major dimensions of Inequalities in India : Education**
<https://www.cbgaindia.org/wp-content/uploads/2016/04/Inequalities-in-Education.pdf>
16. Forbes India, **Poverty Rate in India: Trend over the years and causes** July 5, 2024
<https://www.forbesindia.com/article/explainers/poverty-rate-in-india/90117/1>
17. Skift - **Is the World Ready to Meet the Indian Middle Class?**
<https://skift.com/2024/01/09/is-the-world-ready-to-meet-the-indian-middle-class-megatrends-2024/>
18. India Today, **Explained|How children with limited resources face inequality in education**
<https://www.indiatoday.in/education-today/featurephilia/story/explainedhow-children-with-limited-resources-face-inequality-in-education-2309943-2022-12-16>
19. Invest India, **India preparing for the biggest human migration on the planet**
<https://www.investindia.gov.in/team-india-blogs/india-preparing-biggest-human-migration-planet>
20. Andhra Pradesh, **Jagananna Amma Vodi**
<https://www.myscheme.gov.in/schemes/jav>
21. The Hindu, **Employment Scenario in India grim, ILO Report**, March 26, 2024
<https://www.thehindu.com/business/Economy/employment-scenario-in-india-grim-says-ilo-report/article67995387.ece>
22. The Wire, **Youth Made up 82.9% of India's Unemployed Population in 2022**, March 29, 2024
<https://thewire.in/labour/ilo-report-work-unemployment-india>
23. Vajiram and Ravi – **Most of the unemployment in India is structural in nature.**
<https://vajiramandravi.com/questions/most-of-the-unemployment-in-india-is-structural-in-nature-examine-the-methodology-adopted-to-compute-unemployment-in-the-country-and-suggest-improvements/>

The Impact of Brand Rivalries on Consumer Behavior in the Indian Smartphone Market

Research Question: To what extent did brand rivalries impact consumer behavior in the Indian Smartphone Market?

Author: Swakshi Jain¹

Abstract

As the smartphone market becomes more competitive than ever, companies are resorting to distinct promotion strategies, avoiding the traditional methods; with comparative advertising as one of them. The study employs a mixed-method approach, combining a cross-sectional survey with statistical analysis, both descriptive statistics (central tendencies and standard deviation) and inferential statistics (t-tests). A sample of 113 Indian respondents, primarily in the Northern region, presented their opinions on different advertisements, rating the responses from 1 to 5, by answering a survey that included different comparative advertisements ranging from videos to billboards to tweets by major smartphone brands such as Samsung, Apple, and OnePlus. The study investigates the impact of brand rivalries on consumer behavior in the Indian smartphone market, mainly on two aspects: perception and brand loyalty. The purpose of the study was to investigate how smartphone brand rivalries influence Indian consumers' decision-making, brand loyalty, and product preferences. Results indicate that most of the advertisements had a significant impact on consumer perception and brand loyalty. However, consumers' perception was more adversely affected than brand loyalty. Their responses were primarily affected by the priority of the feature, like choosing privacy over a C-type charger. An astonishing result was that no matter how prestigious or famous a brand was, advertisements about innovative and unique features, work the best in the market and are the most influential, resulting in changing perspectives and brand loyalties. This implies that companies should focus more on developing unique selling points for their products in order to differentiate from their competitors. This research contributes to the understanding of comparative advertising in the highly competitive Indian smartphone context, and it adds to the discussion of the legality of comparative advertising in India.

Keywords: comparative advertising, brand loyalty, consumer perception, product preferences, product differentiation, consumer judgment

¹ Under the mentorship of PhD Professor Dr. Shreshti Rawat

Index

Abstract.....	1
Introduction.....	3
Literature Review.....	4
Methodology.....	6
<i>The Questionnaire</i>	7
<i>Data Sampling</i>	8
<i>Data Analysis</i>	8
Results.....	9
<i>Descriptive Statistics</i>	9
<i>Age-Wise Analysis</i>	11
<i>Gender-Wise Analysis</i>	12
<i>Inferential Statistics: T-Test</i>	12
Discussion And Conclusion.....	16
References.....	18
Appendix.....	19

Introduction

Comparative advertising is a brand's way to fetch the audience's attention by drawing comparisons between different attributes of a product in an attempt to showcase their product as being better than other competitors. Several ways to achieve product differentiation are through physical or quality disparities, differing locations (some brands locate themselves in areas easily accessible to the public), services (some brands offer special services to improve customer experience), and product images, which are created by varying advertising techniques and packaging.² Comparative advertising can be bifurcated into two categories- direct and indirect. While direct comparative advertising involves mentioning the rival brand's name publicly, indirect comparative advertising discreetly highlights the rival brand through signs, symbols, or cynical remarks. As direct comparative advertising might lead to legal and ethical risks or a negative brand image, brands prefer indirect ways of comparisons. The primary goal is to influence consumer preferences and mold their decision-making in favor of the brand. An additional benefit is that it helps to educate the consumers effectively. However, sometimes the message delivered can be perceived as misleading ending up in a negative brand image. An evident example of comparative advertising is McDonald's and Burger King wherein they compete openly with each other constantly comparing the different attributes like size and flavor of the food items offered. One of the examples of comparative advertising in practice in real life is given below.

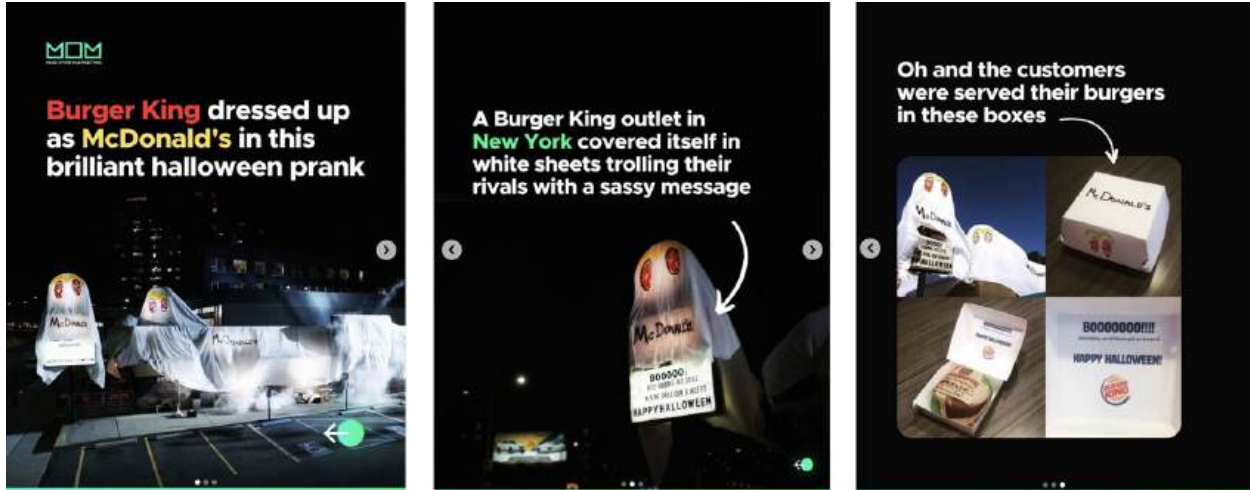


Figure 1. Source: https://www.instagram.com/p/CzCACJ_vlv8/?utm_source=ig_web_copy_link&igsh=MzRIODBiNWFIZA==

The above advertisement epitomizes the famous rivalry between the brands Burger King and McDonalds, with the former taking the lead here. In 2023, Burger King enacted this intending to dominate the market with sarcasm and potentially attract the customer base of the rival. With this

² Tragakes, E. (2011). Economics for IB Diploma, Second Edition. CD-ROM-Cambridge University Press.

example, Burger King propelled his performance into a higher tier, with the blend of sarcasm and festive vibes. So far, there hasn't been any news of McDonald's retaliation to this move.

Such rivalries are also evident in the smartphone market, like between Samsung and Apple. They try to dominate the market by showing the product of the opponent as inferior to their own product. Moreover, considering how competitive the smartphone market is, each company is trying to gain as much market share as possible through various strategies, and comparative advertising is one of them. What makes comparative advertising harder in this market is that the products are not highly differentiated. The basic features are the same, hence the companies must advertise unique features to compare the companies and gather the attention of the audience. The comparison between comparative and non-comparative advertising has never been explored before in the Indian smartphone market, thus it is very interesting to analyze and contributes to the discussion on comparative advertising in the Indian context.

Therefore, this research paper analyzes the impact of brand rivalries on consumer behavior in the smartphone market. It investigates how brand rivalries influence consumer decision-making, brand loyalty, and product preferences and explores the emotional and psychological factors that drive consumers to favor one brand over its rival in highly competitive markets.

Literature Review

In this section, the differences between comparative and non-comparative advertising are explored, the reasons the research focuses on comparative advertising in the Indian context are listed, and specific research questions that the study aims to answer are mentioned.

While comparative advertising has shown its benefits over the years (Muehling 1990; Advertising Age 1976b; Diamond 1978; Krakowiecki 1977; Levine 1976; Neiman 1987; Phillips 1980), non-comparative advertising has also turned to be profitable in certain cases (Thompson 2006; Gorn and Weinberg 1984). Specifically in the context of the Indian market, where there is a large number of brands competing against each other and all kinds of population- savvy and less informed, it can be advantageous to compare different aspects of products through comparative advertising to show product's superiority and educate the general public about the brand and its products. However, in case brands get carried away with their rivalry, it can have serious implications like legal risks and low brand image. As a result, most brands generally resort to non-comparative advertising. It is low risk and helps to build an emotional connection with the audience. In a country like India where brand loyalty plays a significant role, non-comparative advertising helps to develop long-term brand building and leaves a positive impact on the people. In contrast, it can lead to the advertisement being less effective, thus ending up being lost amongst other advertisements. Moreover, consumers might also get confused because of the variety available and not able to understand the product differences.

Indians are known worldwide to appreciate their cultural values a lot, thus they are likely to make more intuitive decisions, trusting their gut, contrasting to pure rational thinking where one takes a more logical step-by-step approach to reach to a final decision. The research by Batra et al., 2000 indicates how Indian consumers purchase products from companies that resonate better with their ideals and cultural identities. Moreover, even Steve Jobs gave his impression of India in his biography by Walter Isaacson,

*“Coming back to America was, for me, much more of a cultural shock than going to India. The people in the Indian countryside don’t use their intellect like we do, they use their intuition instead, and their intuition is far more developed than in the rest of the world. Intuition is a very powerful thing, more powerful than intellect, in my opinion. That’s had a big impact on my work.”*³

Thus, non-comparative advertising has worked great in India as it increased the brand’s reputation in the eyes of the consumers.

Comparative advertising is, legally, very complex in India. It is unethical and illegal to defame another rival brand to market one’s products resulting in brands adopting non-comparative advertising strategies to play safe. Hence, since companies still want to target a specific brand’s market share, they resort to non-explicit means of comparative advertising by discreetly condescending to another brand’s product.

However, what if comparative advertising was legal in India and companies were free to do it? How would the consumers react to it? Would the consumers react positively, or would it backfire and decrease their sales? These questions are further pertinent to discuss in the context of the Indian smartphone market because of the increasing availability of smartphones and the increase of social media use.

With the rise of social media users exponentially, a trend of e-commerce and digital advertising has been seen over the past few years. By January 2023, there were 492.70 million active internet users in India, out of whom 67.5% of all internet users utilize at least one platform of social media. 40.2% of the country’s entire population consists of youth, aged 18 years or more, using social media applications. Thus, companies have started utilizing social media apps vividly to promote their products and lure younger audiences into knowing about and buying their products. However, they can’t only rely on social media. India’s population was recorded at approximately 1.438 billion, therefore there are still 945.3 million citizens who are not on the Internet and don’t know anything about the technological happenings on social media ⁴. To

³ Krishnan, H. (2014, September 7). Intuition over Intellect: What Steve Jobs Learned in India . Retrieved from LinkedIn: <https://www.linkedin.com/pulse/20140907032905-89270668-intuition-over-intellect-what-steve-jobs-learned-in-india/>

⁴ Statistics mentioned in this paragraph regarding the use of social media have been taken from <https://www.forbes.com/advisor/in/business/social-media-statistics/>

address that sector of the population, it is pivotal that the companies use traditional media as well such as billboards, newspaper advertisements, etc. Hence, this research contains both online advertisements and billboards so that the impact can be analyzed on both traditional and digital media.

Since it is established that the research looks at all forms of advertising, it is also critical to address the research questions this research aims to answer: Would the brand loyalty of consumers change after seeing their favorite brand be discreetly demeaned in advertisements? If they knew that the rival brand had more features, would they change their product preferences? Most importantly, do they think that the advertisements are justified?

Methodology

In this section, the entire process of the research is explained in detail from choosing the right questions for the survey to the data sample to statistical analysis (both descriptive and inferential).

A cross-sectional survey was conducted to explore the impact of brand rivalries on consumer behavior in the Indian smartphone market, which would later be analyzed through statistical tests such as descriptive analysis and t-tests, making the study mixed-method research. In a cross-sectional study, the data is analyzed at a specific point in time and mixed-method research includes both qualitative and quantitative analyses. The purpose of the study was to investigate how smartphone brand rivalries influence Indian consumers' decision-making, brand loyalty, and product preferences. In order to evaluate the questions mentioned in the literature review, the respondents were asked to provide their perceptions of different advertisements by renowned brands such as Samsung, Apple, and OnePlus.

As mentioned above in the literature review section, since comparative advertising is considered illegal in India, this research assumes a scenario where it is legal with brands utilizing it to their benefit and analyzes Indian consumers' responses to it.

The survey was designed to cater to prove or nullify the following hypotheses:

Hypothesis 1 (H1): Comparative advertising within the smartphone market affects consumer perceptions of technological superiority, contingent upon the showcased attributes and pricing, as influenced by consumers' evaluations of the fairness and credibility of the comparison.

Hypothesis 2 (H2): Comparative advertising influences consumer's brand loyalty, and its effectiveness is determined by how consumers perceive the fairness and credibility of the comparison.

While H1 addresses the fairness of product comparison and consumer perception, H2 addresses the fairness of comparison and brand loyalty.

The Questionnaire

The questionnaire for the survey contains official YouTube video advertisements, billboard ads, and tweets. For each advertisement or tweet, the questions ranged from two to three, pertaining to the component of either brand loyalty, perception, or judgment.⁵

An example of a tweet with three questions from the actual questionnaire is presented below:

Tweets:

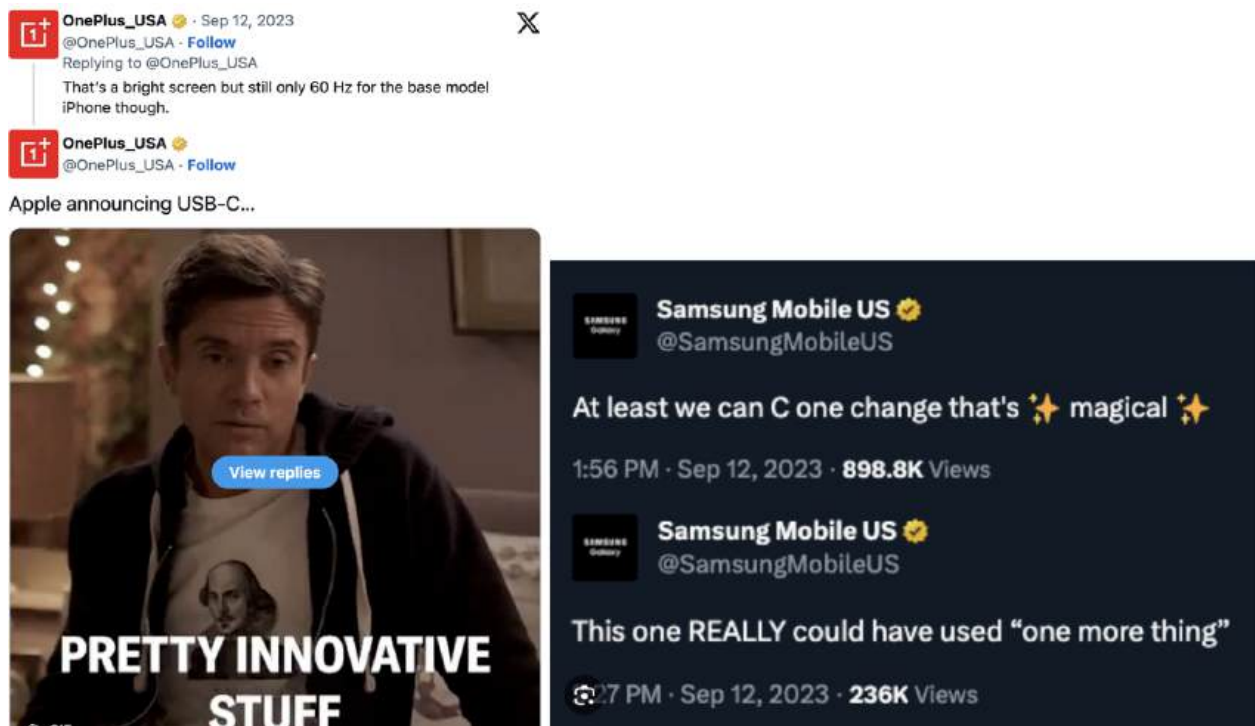


Figure 2. Source: <https://www.marca.com/en/technology/2023/09/13/6501057e22601de0128b45ee.html>

Question 1: *To what extent does this ad influence your decision to buy a specific phone based on its charging features? (Scale: 1 -5)*

This aimed to evaluate if comparative advertising in the smartphone market shaped consumer perceptions of technological superiority, influencing brand loyalty based on featured attributes.

Question 2: *In your judgment, how fair is this comparison between the brands? (Scale: 1 -5, from very unfair to very fair)*

⁵ **Note:** An unfilled copy of the questionnaire is attached in the Appendix.

Question 3: *How likely is it that an ad like this would impact your brand loyalty toward a product?* (Scale: 1 -5, from very unlikely to very likely)

These questions aimed to evaluate if comparative advertising either strengthened or weakened consumer loyalty, depending on how consumers perceived the fairness and credibility of the comparison.

Data Sampling

Data were collected from Indian residents, specifically in the Northern region, ranging across various ages and genders, in February 2024. A convenience sampling method was utilized, wherein the inclusion criteria included people who were aware of the famous smartphone brands, specifically Samsung, Apple, and OnePlus. The survey was created on Google Forms and shared with the researcher and association's network across different social media and messaging apps to recruit respondents.

The sample consisted of 113 respondents whose age and gender division are listed below:

Age

Youth (13 to 24): 45.13% (n=51)

Adults (25 to 60): 54.87% (n=62)

Gender

Male: 38.94% (n=44)

Female: 60.18% (n=68)

Prefer Not to Specify: 0.01% (n=1)

Data Analysis

The results are evaluated through descriptive analysis and t-tests. Descriptive statistics includes a set of methods used in the summarization of the collected data through measures of central tendency, standard deviation, and frequency. Specifically, this research used mean, median, mode, and standard deviation to create a general overview of the demographic characteristics of the sample and address any unusual similarities or differences visible throughout the responses. Along with the common analysis, age- and gender-wise statistics are generated using Google Sheets to discover key variables associated with product preferences and brand perceptions.

On the other hand, a t-test, a component of inferential statistics, is a ratio that quantifies how significant the difference is between the 'means' of two groups while taking their variance or distribution into account. Using Google Sheets, the results are assessed for each advertisement

with a judgment question, through a two-sample unequal variance (heteroscedastic) test using a two-tailed distribution.

Firstly, the respondents were divided into two categories: *Fair* and *Unfair*. The *Fair* community represents individuals that considered the comparison between the two brands expressed in the advertisement was justified. However, the individuals in the *Unfair* community did not believe the comparison to be fair, thereby deeming the advertisement unfair from their perspective. People who chose 4 and 5 on the question, “In your judgment, how fair is this comparison between the two brands?” were categorized into the *Fair* community, and the ones who chose 1 and 2 were categorized into the *Unfair* category. After the segregation, these became the two sample groups to be considered for the test. The test was conducted first to evaluate whether there is a significant difference in brand loyalty between respondents in the two groups, and subsequently, to determine if there is a significant difference in their perception of the product after viewing the advertisement. Results are tabulated in the next section.

Results

In this section, the results of the survey are analyzed and interpreted. Both descriptive statistics and t-test results are tabulated and then explored further in greater detail.

Descriptive Statistics

The table below presents the consumers’ judgment on whether the comparison in the advertisement was fair or not for each question, highlighting the measures of central tendencies (mean, median, mode), and standard deviation. Each advertisement had a similar question: “In your judgment, how fair is this comparison between the two brands?”

Judgment				
<i>Sample Size: 113</i>				
Advertisement	Mean	Median	Mode	S.D.
Video 1	2.841	3	3	1.424
Video 2	3.177	3	4	1.167
Image 1 + Image 2: C type charger	3.124	3	3	1.158
Image 3: Camera	3.451	4	4	1.275

Image 4 +Image 5: Privacy	3.540	4	4	1.303
Image 7 + Image 8: Flip Feature	3.283	4	4	1.271
Image 9: Photos & Videos Quality	3.266	3	3	1.239
Text 10 & Text 11: Accessories	3.540	4	4	1.343

Table 1. Descriptive Statistics: Judgment

The table below presents the effects of the advertisements on consumers' brand loyalty for each question, highlighting the measures of central tendencies (mean, median, mode), and standard deviation. Each advertisement only had a similar question: "How likely is it that an ad like this would impact your brand loyalty toward a product?"

Brand Loyalty				
<i>Sample Size: 113</i>				
Advertisement	Mean	Median	Mode	S.D.
Video 1	2.708	3	1	1.443
Video 2	2.805	3	4	1.487
Image 1 + Image 2: C type charger	2.628	3	3	1.219
Image 6: Price	2.584	3	3	1.272
Image 7 + Image 8: Flip Feature	2.593	3	1	1.334
Text 10 & Text 11: Accessories	3.080	3	4	1.377

Table 2. Descriptive Statistics: Brand Loyalty

The table below presents the impact of the advertisements on consumers' perception of the product for each question, highlighting the measures of central tendencies (mean, median,

mode), and standard deviation. Each advertisement only had a similar question: “To what extent does this ad influence your decision to buy a specific phone based on its **specific based on the ad** features?”

Perception				
Sample Size: 113				
Advertisement	Mean	Median	Mode	S.D.
Video 1	2.655	3	1	1.146
Image 1 + Image 2: C type charger	2.655	3	3	1.266
Image 3: Camera	2.743	3	1	1.406
Image 4 +Image 5: Privacy	3.513	4	4	1.376
Image 6: Price	2.433	2	1	1.281
Image 7 + Image 8: Flip Feature	2.575	3	1	1.400
Image 9: Photos & Videos Quality	2.743	3	3	1.178

Table 3. Descriptive Statistics: Perception

As evident from the tables above, most people were neutral towards the fairness of the advertisements or believed that they were fair. Yet, the advertisements didn’t waver their brand loyalty ($\bar{x} = 2.58$ to 3.08 , $\tilde{x} = 3$). A noticeable feature is that despite the mean and median being almost constant throughout the questionnaire, the Mo varies a lot from 1 to 3 to 4, implying that some ads were unsuccessful, yet some were partially successful for the same set of audiences.

Nor did the ads change their perception of the product ($\bar{x} = 2.43$ to 3.51). In this case, both \tilde{x} and Mo were varied, the former ranged from 2 to 3 to 4, and the latter ranged from 1 to 3 to 4. This implies that the data is not symmetrically distributed.

Age-wise Analysis

The dataset was divided into two categories: youth and adults. The youth was generally neutral to negative in judgment of fairness of advertisements, with the mean score being 2.31. In the

case of youth, specific features had a significant influence. For example, privacy played a crucial role and for that question, the mean score shot up to 3.58 reflecting how the youth values security features, and with the standard deviation of 1.41, the responses had a decent extent of variability. While the adults also had their concerns for privacy, the mean score was relatively less (3.45) and the standard deviation was 1.35, implying the responses were well distributed. In terms of brand loyalty, the mean score of responses by the youth was 2.82, implying they were completely neutral to the decision and unfazed by the advertisement. They are more open to explore a new brand than adults if an advertisement explores a compelling feature that catches their attention. The adults had even a lower readiness to switch brands, with their mean score of 2.17. This implies that they were less likely to change their brand preference by comparative advertising, most likely because of their past experiences and established preferences.

Gender-wise Analysis

Female respondents, though open to shifting their brand preference, were relatively less likely to shift their brand loyalty from the brand that they were familiar with, like Apple, which is indicated by a lower mean score of brand switching (2.77) as compared to men who were more flexible with their brand loyalty, with a higher mean score of 2.88. In terms of perception, both males and females had similar mean scores (2.43 and 2.28 respectively) showing both had neutral perceptions of the advertisements. Female respondents had a standard deviation of 1.14 (perception) and 1.30 (brand loyalty), implying a sense of unity amongst the responses. Almost all women perceived the advertisements in the same manner. On the other hand, male respondents, despite having unity amongst the perception of the advertisements, had a highly varied impact on their brand loyalty, denoted by a standard deviation of 1.58. Moreover, women judged the advertisements more critically than men which is highlighted by the difference in the mean scores, from 2.29 to 2.43. Though minimal, men were more tolerant of the advertisements.

Inferential Statistics: T-Test

In order to reach an outcome, it is necessary to test the hypotheses. Let the mean of the perception or brand loyalty value for the group that considers the ad to be fair be \bar{x}_F and the mean of the perception or brand loyalty value for the group that considers the ad to be unfair be \bar{x}_U .

To determine whether there is a significant difference in perception or brand loyalty between two groups—those who consider the advertisement to be fair and those who perceive it as unfair—it is necessary to test the following hypotheses.

Let:

- \bar{x}_F represent the mean perception or brand loyalty score for the group that views the ad as fair.
- \bar{x}_U represent the mean perception or brand loyalty score for the group that views the ad as unfair.

The hypotheses are set up as follows:

- H_0 (Null Hypothesis): $\bar{x}_F - \bar{x}_U = 0$

This null hypothesis states that there is no difference in the mean perception or brand loyalty between the Fair group and the Unfair group. In other words, the average perception or brand loyalty for both groups is equal.

- H_a (Alternate Hypothesis): $\bar{x}_F - \bar{x}_U \neq 0$

The alternative hypothesis suggests that there is a difference in the mean perception or brand loyalty between the two groups. This implies that the two groups do not have the same average perception or loyalty towards the brand.

Under the alternate hypothesis, two possibilities are considered:

- $\bar{x}_F - \bar{x}_U > 0$: This indicates that the mean perception or loyalty of the Fair group (\bar{x}_F) is higher than that of the Unfair group (\bar{x}_U), meaning those who find the ad fair have a more favorable perception or stronger brand loyalty.
- $\bar{x}_F - \bar{x}_U < 0$: This suggests that the mean perception or loyalty of the Fair group is lower than that of the Unfair group, implying that the Unfair group holds a more positive perception or stronger loyalty.

Testing these hypotheses allows to conclude whether there is a statistically significant difference between the two groups' perception or brand loyalty based on how they perceive the fairness of the advertisement. If the null hypothesis is rejected, it would indicate that there is indeed a meaningful difference in how the two groups view the brand or product after seeing the advertisement.

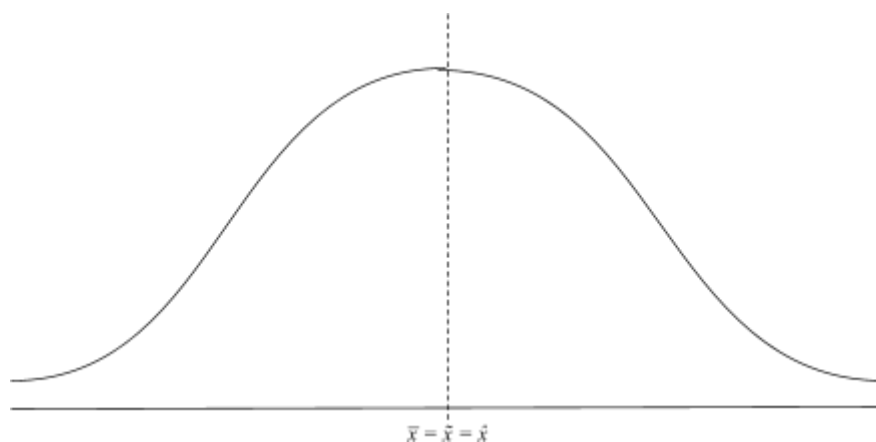


Figure 3. Normal Distribution Curve

Figure 1 above illustrates a normal distribution curve, wherein all central tendencies, mean, median, and mode are equal, and the standard deviation is constant. A t-test helps determine by what factor the dataset's values are positively or negatively skewed.

The significance level is set at 5% ($\alpha = 0.05$), which establishes the threshold for determining statistical significance. In this context, a p-value below 0.05 would indicate sufficient evidence to reject the null hypothesis, whereas a p-value above 0.05 would suggest that we fail to reject the null hypothesis.

To summarise the discussion above, failing to reject the null hypothesis implies that the mean values for brand loyalty or perception are not significantly different between the two groups (Fair and Unfair), suggesting no meaningful distinction between them in terms of these metrics. Conversely, rejecting the null hypothesis would indicate that the mean values of brand loyalty or perception differ significantly between the two groups, signifying a meaningful difference in how the Fair and Unfair groups respond to the product or advertisement.

This framework helps us assess whether the two groups exhibit comparable or divergent attitudes toward the brand or product after viewing the advertisement.

Table 4 below presents the t-test results showing the impact on brand loyalty of the respondents.

Brand Loyalty			
Advertisement	Mean value of those who believed the ad was unfair	Mean value of those who believed the ad was fair	p-value
Video 1	2.050	3.000	0.012
Video 2	2.200	2.960	0.027
Image 1 + Image 2: C type charger	2.545	2.605	0.863
Image 7 + Image 8: Flip Feature	2.605	2.364	0.00004
Text 10 & Text 11: Accessories	3.129	4.132	0.0005

Table 4. T-test: Brand Loyalty

As evident in Table 4, the p -value is greater than 0.05 in only one of the advertisements (0.863 > 0.05), which consisted of two images talking about Apple releasing a C-type charger as a “new” feature while the rest of the Android phones had already switched to this port. This means that we fail to reject the null hypothesis in the t-test for this ad. For those who viewed the ad as fair as compared to those who viewed it as unfair, there is no significant difference in the mean impact on brand loyalty between the two groups. This means that the average effect that the advertisement had was not the same for both groups. Collectively, the advertisement did not affect their brand loyalty. The respondents likely didn’t believe that this advertisement was persuasive enough for them to switch brand loyalties. They decided to remain loyal to Apple and buy its product even though other Android companies believed that the “new” feature didn’t make Apple worth buying.

However, in the rest of them, we reject the null hypothesis. In this case, there is a significant difference in the means of brand loyalty towards the iPhone between the two groups. This shows that for some features, comparative ads might affect the loyalty of individuals, pushing them towards a different brand. Another noticeable factor is that p -values are quite varied throughout the survey, ranging from 0.00004 to 0.027 to 0.863, implying that even though they had an impact, the variance of these advertisements on the consumers as compared to others was very high despite being presented on the same page. Another factor contributing to this could be different priorities consumers keep in mind while purchasing a phone. Since a C-type charger (p -value = 0.863) is a common feature in almost all phones in today’s era, it didn’t affect their loyalty between the brands much, but another innovative feature like a flip feature (p -value = 0.000041) is still limited to Android phones. Consumers likely prefer to see and learn about new and unique features in advertisements which then affect their perception of the brand rather than monotonous ones like a charging port for their loyalty to shift from one brand to another.

Table 5 below presents the t-test results, highlighting the significant role of consumer perception in shaping brand preferences.

Perception			
Advertisement	Mean value of those who believed the ad was unfair	Mean value of those who believed the ad was fair	p -value
Video 1	2.450	2.595	0.676
Image 1 + Image 2: C type charger	2.364	2.907	0.098
Image 3: Camera	2.143	3.164	0.005

Image 4 +Image 5: Privacy	2.857	4.042	0.005
Image 7 + Image 8: Flip Feature	1.960	2.983	0.001
Image 9: Photos & Videos Quality	2.250	3.115	0.012

Table 5. T-test: Perception

In the case of perception, the p -value is greater than 0.05 in two of the advertisements, Video 1 ($0.676 > 0.05$) talking about the launch of a new Samsung Galaxy phone and the images of a C-type charger ($0.098 > 0.05$). It means that we fail to reject the null hypothesis, implying that there is a significant difference in the mean perception between the two groups, those who considered the comparison to be fair contrary to those who considered it to be unfair. It further means that the difference in the mean values of perception equals zero, implying that these advertisements had no impact on consumer perception. Despite the advertisements being sarcastic or humorous, customers didn't agree with them or agreed to an extent where the good brand image of the company and its products didn't waver. They chose to view Apple in that positive light as they did previously.

On the other hand, for the remaining advertisements, the p -values are less than 0.05, rejecting the null hypothesis. It suggests that there is a significant difference in the mean perception between the two groups. For both groups, the advertisements molded the perception of the respondents. For example, image 3, in which Samsung taunted Apple for recently releasing 8K filming while they had been in the market for the last 2 years and 6 months. This means that the camera (p -value = 0.005) is one of the significant features that the consumers think about and prioritize before finalizing the brand that they're going to purchase. Similarly, privacy (p -value = 0.005) is also one of the prominent features that are prioritized by consumers. In this advertisement, Apple retorted to competitors' contemptible advertisements about its product. Consumers likely viewed it as an appropriate response to those ads and, thus were highly persuaded by it. With increasing malicious activities occurring around the world, privacy is a major feature that people look for now which could also contribute to this advertisement being more convincing than others.

Discussion and Conclusion

This paper attempted to test the extent to which comparative advertising impacted consumer perception of the product and brand loyalty of the smartphone market, specifically in the Indian context. It aimed to evaluate if the advertisements changed brand loyalty and product preferences negatively or positively, depending on whether an individual thought the advertisement was fair

or unfair. Since it is not allowed for the companies to explicitly address another brand in a negative light, this research assumes an “idealistic” scenario where it is legal, and companies have complete freedom to compare product features. A survey was conducted, and the results were analyzed for a data sample of 113 respondents through descriptive statistics and t-tests.

The t-test results highlight the influence of advertisement fairness on consumer perception and brand loyalty. For most advertisements, significant differences were observed between those who viewed the ads as fair and those who did not. This demonstrates the power of comparative advertising, particularly when focused on innovative and unique product features, in shaping consumer attitudes.

Regarding brand loyalty, significant differences were seen across most ads. For example, the C-type charger ad had no significant effect, indicating that routine features like charging ports are less likely to influence consumer loyalty. In contrast, advertisements for more innovative features, such as the flip feature and accessories, had a strong impact, demonstrating that unique product attributes are critical in driving changes in consumer loyalty. Interestingly, the flip feature ad had the smallest p-value (0.00004) for brand loyalty, showing that it had the most substantial effect on changing brand loyalty.

For perception, the p-values were generally less varied than those for brand loyalty, indicating a more consistent impact across ads. This reflects the idea that perception is easier to sway than brand loyalty, which typically requires longer-term engagement with the brand. Notably, the flip feature ad also had the strongest effect on perception (p-value of 0.001), further supporting the idea that innovative features drive consumer attitudes. In contrast, the C-type charger ad had no significant effect on either perception or brand loyalty, underscoring that standard features do not generate much consumer engagement.

Given the dataset only included the population in the North Indian region, it is not exactly representative of the entire Indian population. Moreover, convenience sampling was adopted, and an online survey was shared due to the lack of time and accessibility. This resulted in reduced in-person interactions, hence the behavioral change could not be assessed very efficiently. This research can be further enhanced by taking a larger dataset that is representative of the entire population and more personal interactions with the respondents. Perhaps, taking more brands in consideration, instead of just three, could also give a clearer picture of how the consumers respond to advertisements by the brands in this highly competitive smartphone market.

The key takeaways from this research are:

1. Innovative features drive both perception and loyalty: Ads highlighting unique features like the flip feature had the greatest impact on both brand loyalty and perception, with the lowest p-values for both metrics. This underscores the importance of innovation in competitive advertising.

2. Perception is more easily swayed than loyalty: The more consistent p-values in perception suggest that consumer perception can be influenced more readily than brand loyalty, which requires longer-term engagement and sustained brand experience.

3. Standard features have little impact: Ads focusing on routine features, such as the C-type charger, showed no significant effect on either brand loyalty or perception, indicating that these features do not generate meaningful consumer responses.

This research demonstrates that competitive advertisements focused on unique, innovative product features are the most influential, while standard or common features fail to sway consumer perceptions or loyalty. Additionally, Apple's relatively fewer advertisements compared to its competitors may reflect a strategic preference for non-comparative advertising methods, relying instead on more general, widely accepted advertising practices.

References

- Batra, R., Ramaswamy, V., Alden, D. L., Steenkamp, J.-B. E., & Ramachander, S. (2000). Effects of Brand Local and Nonlocal Origin on Consumer Attitudes in Developing Countries. *Journal of Consumer Psychology*, 9(2), 83-95.
- Muehling, D. D., Stoltman, J. J., & Grossbart, S. (1990). The Impact of Comparative Advertising on Levels of Message Involvement. *Journal of Advertising*, 41-50.
- Gorn, G. J. (1984, September). The Impact of Comparative Advertising on Perception and Attitude: Some Positive Findings. *Journal of Consumer Research*, 11.
- Tragakes, E. (2011). *Economics for IB Diploma, Second Edition*. CD-ROM-Cambridge University Press.
- Krishnan, H. (2014, September 7). *Intuition over Intellect :What Steve Jobs Learned in India* . Retrieved from LinkedIn: <https://www.linkedin.com/pulse/20140907032905-89270668-intuition-over-intellect-what-steve-jobs-learned-in-india/>
- Wong, B. (2024, February 6). *Top Social Media Statistics And Trends*. Retrieved from Forbes Advisor: <https://www.forbes.com/advisor/in/business/social-media-statistics/>
- Kaliyadan, F., & Kulkarni, V. (2019). Types of Variables, Descriptive Statistics, and Sample Size. *Indian Dermatology Online Journal*, 82–86. Retrieved from National Library of Medicine: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6362742/>
- Wadhwa, R. R., & Marappa-Ganeshan, R. (2023). *T Test* . StatPearls.

- Kershaw, A. G. (1976). For and against comparative advertising. *Advertising Age*.
- Diamond, S. A. (1978). Beware of naming the competitor. *Advertising Age*.
- Krakowiecki, M. (1977). What Else is Going Up? Ad Liability Insurance Cost. *Advertising Age* .
- Levine, P. (1976). Commercials That Name Competing Brands. *Journal of Advertising Research*.
- Neiman, J. (1987). The trouble with comparative ads. *Ad Week*.
- Phillips, G. H. (1980, October 27). A Review of Comparative Advertising. *Speech presented at AAAA Mid-Atlantic Conference*.
- Thompson, D. V., & Hamilton, R. W. (2006). The Effects of Information Processing Mode on Consumers' Responses to Comparative Advertising. *Journal of Consumer Research*, 530-540.

Appendix

Questionnaire:



Survey Responses:

