

|CURIEUX|

---

# ACADEMIC JOURNAL

**December Issue**

Part 2 Issue 45

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: Optimising Outcomes: The efficacy of Preoperative and Postoperative Care for Transposition of the Great Arteries By Chanwipha (Phoebe) Sirisaengtaksin**

**Page 27 : Reading Help for the Visually Impaired: the Case of Talking Ads By Stefan Maxim**

**Page 33: *Salmonella*'s Type 6 Secretion System: Pathogenesis, Host Cell Manipulation, and Response By Vana Scorza**

**Page 46: A Review on Colorectal Cancer in Young Vs Old Patients, Including Related Symptoms, Diet, and Gut Microbiomes By Anirudh R. Guddeti, David H. Parkinson**

**Page 57: Understanding Financial Literacy in Immigrants and Non-English Speakers in the United States By Jack Andreuccetti**

**Page 75: Natural Disasters Effect on Puerto Rican Educational Output By Matthew Cattaneo**

**Page 83: An Exploration of Gender Disparity in Business Classes at Public High School By Festina Hyseni**

**Page 97:Low Fertility Rates in Developed Nations: Causes, Consequences, and Policy Solutions By David Lubin**

**Page 106: Cancer in a Global Context: Addressing Risk, Access, and Policy Disparities Between High-Income Countries and Low and Middle-Income Countries By Mali Lord**

**Page 127: Empowering Small Businesses and Nonprofits: Effective Digital Marketing Strategies for Growth and Visibility By Aanya Mishra**

**Page 137: Truth is Marching On: The Evolution of Afro-American Spirituals and Their Impact on the American Civil Rights Movement By Matthew Lee**

**Page 142: Mediapipe Fingering Classification for Violin By Spencer Wang**

**Page 152: Body Mass Index Mediates the Association Between Physical Activities and Mental Health By Amay Munjal**

**Page 160: What Practices Optimize Educational Experiences for Youth in the Juvenile Justice System? By Eliana White**

**Page 176:The Impact of Hay Fever on Public Health in Japan By Sara Tiwari**

**Page 187: Analysing the Transformative Effects of Chinese Government Border Policies on Hong Kong's Tourism Industry: Insights from COVID and Post-COVID Analysis By Yu Wei Lin William**

**Page 202: Consumer Purchase Intention When Buying Athletic Shoes By Vedant Kothari**

**Page 209:The Decline of Shui Script and Its Cultural Implications: A Study on the Changing Attitudes Towards Cultural Heritage Among the Shui Tribe By Rongyu Zhou**

**Page 225: The Roles and Mechanisms of Burst Firing Neurons in the Lateral Habenula Circuit in Depression, and Implications of Targeted Inhibition By Antidepressants as Potential Treatment By Emily Chen**

**Page 250: The Influence of Digital Platforms on Hindu Spiritual Guidance: A Comparative Study of Online and Real-Life Gurus By Karan Shroff**

**Page 260: Study of Statistics of Solar Flares By Rania Shekhar**

**Page 272: What Shaped Jacob Riis Into an Urban Reformer By SeungJun Lee**

**Page 285: How does Dynamic Pricing Affect Airline Customers? By Diego Alvarez Gonzalez**

**Page 296: What can Social Media Tell Us About Jan 6th and the Future of the MAGA Movement By Matthew Sidun**

**Page 304: How Can Gene Therapy be Used to Treat PLA2G6 Associated Neurodegeneration? By Aadya Jain<sup>1</sup>, Jacob Odell<sup>2</sup>**

**Page 320: Behind the Glamour: Examining Gender Discrimination within K-pop By Riley Kim**

**Page 334: Applying Machine Learning Within Augmented Reality As A Tool For Various Industries By Karan Jain**

**Page 355: The Statistical Link Between ADD/ADHD and Substance Use, and Opportunities for Preventing Adolescent Overdose Fatalities within This Population By Liv MacMillan**

**Page 362: Association Between Social Media and Forensic Psychoanalysis: Investigative Methodology By Vinicius Meireles Aun Carteiro**

**Page 369: The Role of Anti-PD-1 as A Treatment for Cancer By Siena Lin**

**Page 374: Genetic and Parental Behavior's Effect on the Development of Childhood Anxiety By Aon Huang**

**Page 379: Athletes and Mathletes: How Non-contact Sports Improve Math Learning By Improving Students' Spatial Skills By William McAfee**

**Page 391: Sustainability of Bioplastics in Aquatic Environments: A Comparative Analysis into the Biodegradability and Environmental Impacts of PLA, PHA, Starch-based, and Cellulose-based Bioplastics By Sean Tey Yi-Shuen**

**Page 405: Bridging Cultures: Analyzing Phonological Challenges in Korean Language Acquisition for Filipino Learners By Dominic Hahm**

**Page 415: Hyperparameter Tuning Improves Computer Vision Tools for Wildlife Conservation By Achyut Venkatesh**

**Page 423: How Was Lili Boulanger's Music a Sign of the Rise of Early Feminism in France During the Early 20th Century? By Jenny Zhu**

**Page 438: The Credit Scoring System's Relationship with Financial Inclusion in Developing Countries By Khanh Nam Nguyen (student), Alan Zhang (mentor)**

**Page 450: Breast Cancer Survey By Arnav Rajadhyax**

**Page 455: A Cultural Awakening for South Korean Guide Dogs By Catrina Cho**

**Page 461: RNA Immunoprecipitation for RNA Fragments Containing the m7G Cap**

**By Sophie Kwon and Serene Kwon**

**Page 467: Household Readiness for Induction Cooktop Adoption in India By Evan Obbard**

**Page 485: Evaluating Plant and Bird Species Diversity and Abundance and Trash Pollution at Three Nature Reserves in Singapore By Yiqing Gong**

**Page 508: Lightweight Cryptography as A Tool to Meet Singapore's Carbon Emission Targets By Ananya Kharbanda**

**Page 520: The Impact of Cryptocurrencies and Central Bank Digital Currencies on Financial Stability: A Cross-Country Analysis By Mingrui Deng**

**Page 531: How Convolutional Neural Networks Can Help Predict NFT Prices by Analyzing Images By Shunsuke Nakatani**

**Page 539: The “Keys” to Success: Exploring the Impact of Music on the Abilities of Novice Surgeons By Rithika Jamisetty**

**Page 557: Contaminated Waters: Unveiling the Health Risks of Islamabad’s Water Quality Crisis By Abeeha Atif**

**Page 567: Music as a Tool in the World of Streaming Platforms By Michael L Wan**

**Page 584: Which Has a Bigger Effect on History: the Plans of the Powerful or Their Mistakes? Understanding Historical Impacts: Mistakes vs. Planning By Zelai Li**

**Page 593: Upper-Echelon Theory - The Impact of CEO Socio-ethical Backgrounds on Profits and Business Success By Aryan Agarwal**

**Page 640: The Impact of Brand Rivalries on Consumer Behavior in the Indian Smartphone Market By Swakshi Jain<sup>1</sup>**

**Page 651: Thank you for the guidance of Simon Alford, mentor from Cornell University in the development of this research paper By Kevin Hsieh**

**Page 661: Labour Laws in India and Canada: A Comparative Study of Legislative Changes and Their Driving Forces By Reyhaan Aryan**

---

<sup>1</sup>

## **Optimising Outcomes: The efficacy of Preoperative and Postoperative Care for Transposition of the Great Arteries By Chanwipha (Phoebe) Sirisaengtaksin**

### **Abstract**

Transposition of the Great Arteries (TGA) is a common congenital heart defect (CHD), which is fatal if not treated. The arterial switch operation (ASO) is typically performed due to its short- and long-term benefits. This review will outline the management before and after the ASO – preoperative and postoperative management–by assessing its risks and benefits and the standard protocol and technologies. Monitoring will be a major part of both, but the main focus for preoperative care is the administration of prostaglandin E1 (PGE1) and balloon atrial septostomy (BAS) to allow sufficient mixing of systemic and pulmonary circuits. Postoperative care will aim to stabilise the patient after surgery by careful monitoring, anaesthesia and mechanical support. Furthermore, treating the side effects and complications resulting from the ASO is crucial. Patients with TGA usually have associated conditions like ventricular septal defect (VSD). Therefore, cautious monitoring and drug administration are applied specifically to the patient. TGA remains intractable, so the approach towards efficient care for neonates with TGA will be based on the efficacy of pre and post-operative management.

### **Introduction**

Transposition of the Great Arteries (TGA) is a congenital heart defect (CHD) where the position of the aorta and pulmonary artery are switched, and the aorta connects to the right ventricle while the pulmonary artery connects to the left ventricle, meaning the two circuits are in parallel. (Haeffele and Lui, 2015) When ventricle systole occurs, deoxygenated blood is pumped from the right ventricles into the aorta and returns to the systemic circuit. In contrast, oxygenated blood leaves the pulmonary artery, flows through the pulmonary circuit, and returns to the left ventricles so oxygenated blood does not reach the body cells. It is common for babies to develop cyanosis within 30 days and tachypnea or hypoxemia in severe cases. Mixing can occur if the patient has a ventricular septal defect (VSD), another CHD where the baby has a hole in the septum, which is beneficial for TGA. It can also occur when the patent ductus arteriosus (PDA, a blood vessel that connects the aorta to the pulmonary artery) and foramen ovale (FO, a hole between the left and right atria) remains open. These structures are only present in a prenatal heart since they start closing after birth, so preoperative procedures are performed to keep them open. (Szymanski *et al.*, 2024)

TGA is responsible for 5-7% of all cardiac diseases and is more common in males than females. Due to the development of the arterial switch operation (ASO), patients have up to a 90% chance of survival within 20 years. In terms of recurrence risk, it is between 0.27 to 2% for siblings, and for CHD, it is 1.45 for parents and siblings overall. (Villafañe *et al.*, 2014) Even though numbers are relatively low, the main issue is the consequences of treatment for TGA, which may not be due to the surgery itself but the care received before and after the operation.

The desirable time for surgery is important since the preoperative period presents some risks for the patients. There is no fixed timing, but the use of mechanical support may provide safety for delayed surgery. Generally, ASO is performed for neonates aged 3-8 weeks, but to decide whether early surgery would be beneficial, monitoring using near-infrared spectroscopy (NIRS) and considering their clinical stability is of major importance. In cases of comorbidities like cerebral stroke with TGA, ASO is often delayed for 3-4 weeks to reduce chances of neurological damage. However, lengthening the preoperative time allows for further cerebral injuries, so determining which patient should receive early surgery is detrimental. Preoperative mortality with TGA is about 3.6% to 10.3% for neonates, but the combination of medical technologies and other primary care has helped to evaluate cardiac performance and risk of poor preoperative cerebral oxygenation. (Séguéla *et al.*, 2017)

Postoperative care is equally crucial since the concerns following ASO are low cardiac output syndrome (LCOS), myocardial dysfunctions, extensive bleeding, arrhythmias, etc., thus ensuring sufficient postoperative treatment like extracorporeal membrane oxygenation (ECMO) and hemodynamic/respiratory support is vital for the patient's survival. (Datt *et al.*, 2022) Nevertheless, the question of the efficacy of these procedures and risks should be explored to establish the best techniques for life-long patient care.

Specific pathways for patient care will ensure outcomes for patients are optimised both short-term during the period of their hospitalisation, and long-term to prevent risks of underlying conditions. Hence, by evaluating the management methods, we are able to identify gaps in knowledge whether that is the risks of each procedure or its unknown benefits for the patients that need to be filled.

## **Main Text**

### **Clinical protocols for Preoperative Care of Transposition of the Great Arteries**

After diagnosis for TGA, neonates are monitored in the ICU, which includes non-invasive procedures such as checking for peripheral perfusion, vital signs, urine output monitoring and cardiovascular examination. Here, stabilisation, management of hypoxia (insufficient oxygen level to maintain homeostasis), maintenance of systemic perfusion (delivery of blood through the circulatory system) and correction of acidosis. To ensure adequate mixing of oxygenated and deoxygenated blood, oxygen saturation should be 75-85% of the arterial blood gas and oxygen saturation of arterial blood ( $\text{SaO}_2$ ) depends on the levels of two sources of systemic circulation. Generally, neonates with extreme hypoxaemia ( $\text{SaO}_2 < 60\%$ ) would require more attention. (Sarris *et al.*, 2017) Hypoxaemia refers to low arterial oxygen tension ( $\text{PaO}_2$ ), whilst hypoxia is the failure of oxygenation at the tissue level, which can lead to dangerous hypoxia.

Although standard diagnostic tools like chest X-rays and electrocardiograms (ECG) allow anatomical structures of the heart to be seen and spot abnormal rhythms, their findings are not specific enough for CHDs like TGA. Monitoring tools like pulse oximetry, echocardiography and cardiac catheterisation are more reliable tools with high sensitivity and specificity, which

will be discussed. (Martins and Castela, 2008) Neonates with TGA have an immature brain (1-month delay in brain development), making neurological symptoms very likely. Therefore, precise monitoring of cerebral perfusion and oxygenation is necessary. (Séguéla *et al.*, 2017) Other interventions include an increased fraction of inspired oxygen ( $FiO_2$ ), transfusion to treat anaemia and increase oxygen-carrying capacity, increasing pulmonary blood flow, sedation to reduce oxygen consumption and inotropic agents (medicine to increase cardiac output.) Still, these will not be mentioned in detail in this section since they are specific to the postoperative section. (Sarris *et al.*, 2017)

## 1. Existing methods for monitoring of TGA

### Non-invasive approaches

A. The **pulse oximetry** measures the peripheral oxygen saturation ( $SpO_2$ ) in the blood. The device is placed in the patient's fingertips, and the light source passes through the skin; then, light in two wavelengths will illuminate oxygenated and deoxygenated haemoglobin in the blood. (Therapeutic Goods Administration, 2022) Compared to clinical examinations, pulse oximetry is more cost-effective, straightforward and accurate. With this, interventions like surgery for d-TGA can begin when appropriate, which reduces the risks of complications. (Amouzeshi *et al.*, 2021) On the other hand, pulse oximeters can over or under-estimate oxygen saturation due to multiple factors such as skin pigmentation and peripheral blood flow, which means hypoxia may not be acknowledged and treated. (Therapeutic Goods Administration, 2022)

B. **Near Infrared Spectroscopy (NIRS)** is a continuous non-invasive method used to evaluate regional cerebral oxygen saturation ( $rScO_2$ ) and can be used for preoperative risk assessment since the brain is seen as an index organ for perfusion—specifically the balance between oxygen delivery and consumption. (Pérez Moreno *et al.*, 2018) Typically, NIRS works by applying the Beer-Lambert Law (the correlation between concentration and absorbance of a solution.) Near-infrared light photons are placed on the skin from a source in which they will penetrate the tissue, and some of the photons will reflect and enter detectors. The absorption of the tissue is measured, and oxyhaemoglobin and deoxyhaemoglobin are calculated, giving the  $rScO_2$ , which corresponds to the venous oxygen saturation. During full-term birth,  $rScO_2$  increases to 70-80% in the first 15 minutes, and these values will help to determine the neurodevelopmental outcomes of TGA patients. Balloon atrial septostomy (BAS) increases  $rScO_2$  from an average of 42% to 64%; hence, NIRS values should be a factor in determining whether BAS should be performed. (refer to BAS section) It has also been proven that ASO increases  $rScO_2$  from 56% to 80% after surgery. Furthermore,  $rScO_2$  is related to the diameter of the patent ductus arteriosus since ductal steals can contribute to brain injuries like haemorrhage. (Séguéla *et al.*, 2017) Many studies show that the decisions made by NIRS monitoring have reduced mortality and improved clinical outcomes, resulting in shorter stays in the ICU and postoperative hospitalisation. (Cheung *et al.*, 2022) Consequently, evaluating  $rScO_2$



using NIRS will help to predict a suitable time for surgery, especially when rScO<sub>2</sub> values are low, and patients will benefit from early surgery.

Compared to pulse oximetry, NIRS calculates the percentage of oxygen in both arterial and venous blood, whilst the pulse oximeter only calculates arterial oxygen saturation. Additionally, NIRS can detect changes in oxygen saturation faster than the pulse oximeter. (Cheung *et al.*, 2022) However, things to consider about NIRS are that their samples are limited to the frontal lobe of the brain, and their maximum depth is 30-40 mm. There is also no definition of the baseline for rScO<sub>2</sub>, which suggests that the amount is abnormal. Plus, there is not a clear link between the decrease in rScO<sub>2</sub> and neurological complications. (Ghumman, Kumar and Kumar, 2019) When applied to TGA, other monitoring devices like pulse oximetry still need to be used alongside NIRS since its values may not be adequate when making treatment decisions like surgery.

**C. Echocardiography** provides an accurate understanding of the morphology of TGA alongside its associated lesions. Evaluating the coronary artery pattern will help to include or eliminate other malformations. Adequate mixing could also be monitored by observing the flow through the arterial duct and septal defects (if present). (Martins and Castela, 2008) In the presence of cyanosis, the size of atrial septal defects (ASD) will influence BAS. In cases of TGA with left ventricular outflow tract obstruction (LVOTO), echocardiography can identify this from different perspectives. Moreover, coronary arteries in the pulmonary artery, aorta and between the great arteries (intramural course) could be assessed. (Cohen and Mertens, 2019) This will allow treatment options like surgery to be explored, as well as the impacts of these treatments preoperatively. However, two-dimensional echocardiography may miss-diagnose some areas of the heart; for example, PDA can be missed in patients with persistent pulmonary hypertension (PPHN.) Therefore, echocardiography needs to be paired with additional imaging tools like MRI, CT, and cardiac catheterisation, which can be used in different cases. (Soongswang *et al.*, 2000)

### Invasive approaches

Even Though echocardiography is accepted as the most efficient diagnostic tool for CHDs, **cardiac catheterisation** is the gold standard for obtaining haemodynamic (blood flow) data. Cardiac catheterisation gives information on a computer monitor to make the heart visible. It works by a thin, flexible tube called a catheter being inserted into the blood vessel in the groin or arm and then moved up to the heart. The X-ray contrast agent is then injected to see the heart more clearly, ensuring that pressure in all heart chambers is stable. This is a minimally invasive procedure since there is no need for big incisions, and only 1% of examinations result in complications. However, risks to be considered are blood vessel and nerve damage, acute heart attack or stroke, allergic reaction to the contrast medium, bleeding at the insertion site and temporary arrhythmia. (InformedHealth.org, 2023)

## **2. Haemodynamic management to prepare for the surgery for TGA**

### Prostaglandin E1

Prostaglandins (PG) are lipids with functions similar to hormones since their role in the human body is defence and repair at the site of tissue damage or infection. Meanwhile, pharmacological PG is used in medicine as a short-term treatment before surgery. (Ashraf, 2020) In order to improve the mixing of systemic and pulmonary circulations, IV infusion of PGE1, a powerful vasodilator, is used to keep the ductus arteriosus open. All TGA patients will have ductal-dependent mixing if the intracardiac shunt does not work properly. To put simply, PGE1 stabilises the patient before surgery, like BAS, since it allows the rest of the patient's body to receive more oxygenated blood and is necessary for hypoxic neonates. (Séguéla *et al.*, 2017) In practice, PGE1 infusion should start with lower doses from 0.0125-0.05 mg/kg/min to maintain the opening, but higher doses up to 0.1 mg/kg/min could be used for the ductus to reopen. Afterwards, PGE1 can be discontinued once mixing is sufficient at the atrial level unless the patient has LVOTO. (Sarris *et al.*, 2017)

Even though this treatment is essential to prolonging the patient's life, PGE1 is a proinflammatory molecule, so there are many side effects, including oedema, hypotension, increase in body temperature and hypoventilation. In a case report with 492 infants with CHD treated with PGE1, around 12% of children had hypoventilation, mainly occurring in patients with low birth weight (less than 2 kg.) Even with low doses of PGE1, about 0.05-0.1 mg/kg/min, it could cause pulmonary vasodilation, which increases blood flow, leading to pulmonary hypertension. Typically, CNS issues make up 16% of the patients since PGE1 affects brainstem respiratory regions (most likely the microglial-dependent mechanism), with the most common side effects being seizure and apnoea. (Tabbutt, Helfaer and Nichols, 2006) In addition, having adequate oxygen saturation does not always mean there is sufficient oxygen because cerebral venous oxygen saturation is sometimes lower than predicted. The mean cerebral oxygen saturation in infants with normal saturation or with PDA is around 50 and is even lower for TGA patients. This illusion of security is dangerous when discussing the timing for surgery—prolonged waiting time subjects patients to risks of infection, paradoxical embolism, mechanical ventilation, delaying the surgery and increased costs of being hospitalised for longer, which is stressful for both the patient and their family members. (Villafañe *et al.*, 2014) Infusion is usually short-term, from 6 hours to 20 days, but in cases with infection, low birth weight, or limitations with specialised hospitalisation (in countries where some medical treatments are unavailable), this time span could be extended for months. (Ashraf, 2020) In an experiment to determine the effect of prolonged treatment using PGE1, seven neonates with TGA with no VSD received PGE1 infusion for around 39 days. The results showed that their arterial partial pressure of oxygen (pO<sub>2</sub>) has increased from a mean of 27 torr to 40 torr, which remains for the rest of the therapy. Thus, it is reasonable to conclude that PGE1 is effective in opening the ductus for many weeks. (Teixeira *et al.*, 1983) However, long-term treatment (>120 hours) is often avoided since it leads to more severe side effects such as gastric-outlet obstruction, cortical hyperostosis and pseudo-Bartter syndrome. This means that patients require lengthened preoperative mechanical ventilation and postoperative hospitalisation, which could be uncomfortable for young patients to handle. (Séguéla *et al.*, 2017) Cortical hyperostosis seems to

be directly linked to the dose or length of PGE1 infusion since there is an increase from 42% to 100% of patients who developed hyperostosis from 30 to 60 days, although some cases can be seen as early as 9-11 days. Fortunately, this condition is reversible. Self-recovering and discontinuation of PGE1 allow the condition to disappear (but maybe not instantly.) (Ashraf, 2020)

As mentioned, PGE1 is used to improve systemic oxygenation, but there are some cases where PGE1 is not recommended for TGA. For example, if the patient has a small gap in the patent foramen ovale, PGE1 might cause the foramen ovale to close since more blood returns to the left atrium resulting in less mixing of blood and causing the side effects mentioned above. (Beerman, 2023) Discontinuation of PGE1 could also pose a risk of rebound hypoxemia, so it is recommended to wean from PGE1 rather than completely stopping the treatment. (Sarris *et al.*, 2017)

Therefore, it is recommended that patients with PGE1 infusion are monitored cautiously, especially if the patient develops hypotension or hypoventilation. Decreasing the rate of infusion can also be done if the patient experiences fever and jitteriness, and antibiotics are advised to prevent infections. (Lewis *et al.*, 1981) For patients with severe hypoxemia and who have a slow response to PGE1 or a very restricted foramen ovale, cardiac catheterisation and BAS are vital, which will be discussed later in the review. (Beerman, 2023)

### Balloon Atrial Septostomy

Balloon Atrial Septostomy (BAS), also known as the Rashkind Procedure, is also a short-term treatment when the foramen ovale has a small opening which prevents sufficient systemic oxygenation. BAS was first done in 1966 by Rashkind, where an inflated catheter with a balloon tip is inserted from the left to the right atrium through the foramen ovale and then inflated in order to enlarge the opening of the atrial septum for oxygenated and deoxygenated blood to mix. Afterwards, the balloon is deflated and the catheter removed, and these steps can be repeated more than once. (Séguéla *et al.*, 2017) BAS can be performed up to the age of 6 weeks since, over 6 weeks, the atrial septum will thicken. In the past, BAS had to be done in a cardiac catheterisation laboratory with uniplane fluoroscopy, which was complicated. However, there is now an option for echocardiographic-assisted atrial septostomy using the femoral or the umbilical vein, which can be done in the ICU. The femoral approach is done by puncturing the vein and inserting the catheter and dilator, which can travel to the right atrium, whilst the umbilical approach is when the umbilical cord is cut at a tangent to the skin and placing the catheter into the umbilical vein. Both methods enable the procedure to be done with a smaller risk of injury to the atrioventricular valves, reducing the amount of time being exposed to ionising radiation, and the procedure can be performed faster with fewer expenses, too. (Cirstoveanu *et al.*, 2023)

Around 28% of TGA patients have undergone BAS since it helps to prolong the life of patients before ASO. If BAS is adequate, PGE1 can be discontinued, but it is usually continued for a short period of time after BAS to ensure oxygen saturation is sufficient and to limit side

effects. (Séguéla *et al.*, 2017) Generally, the success of BAS is classified by at least a 10% increase in systemic arterial saturation, interatrial communication shown in an echocardiogram (increase in the atrial septal defect (ASD) diameter of more than around 33% of the total septal diameter) and minimal blood pressure difference between two sides of the heart. BAS in patients with optimal oxygenation has a survival rate of 100% before surgery compared to 84% in non-BAS patients, but this is contradictory to patients with poor oxygenation (mortality is 36.3% compared to 0%, respectively). Therefore, this data suggests that BAS aids the survival of patients with sufficient oxygen saturation and minimal size of ASD when the surgery is delayed for more than 3 days of life. (Cirstoveanu *et al.*, 2023)

Research done in the USA revealed that 8681 patients with TGA were analysed, and 1742 patients underwent BAS. It reveals that patients who underwent the procedure had lower mortality (10% compared to 12% for non-BAS patients.) However, BAS increases the median length of stay and median comorbidities with almost double the risk of stroke. (Mukherjee *et al.*, 2010) Neurological injuries were suspected side effects of BAS caused by the displacement of pre-existing thrombi. This included white matter injury (WMI), periventricular leukomalacia (PVL) and intraventricular haemorrhage. Another research using MRI scans shows that 41% of patients with d-TGA who had undergone BAS out of 29 patients have acquired preoperative brain injury. (McQuillan *et al.*, 2006) Nevertheless, modern research has rebutted this assumption. In a meta-analysis of a total of 10,108 infants with TGA, there was no link between BAS being the cause of brain injuries. (Séguéla *et al.*, 2017) A prospective study also using MRI indicates stroke is present in 3% of patients with BAS and 9%. Overall, the risk for brain injury from BAS was 3%, with no association with the time or route in which BAS was performed. (Beca *et al.*, 2009) Studies show a correlation between surgery delay and hypoxemia, which leads to lower oxygen saturation, metabolic acidosis and brain vulnerability and complications. The initial assumption of BAS increasing brain complications could be because patients who require BAS have severe hypoxemia and haemodynamic instability, resulting in haemorrhages and ischemia. (Cirstoveanu *et al.*, 2023)

Aside from neurological injuries, other complications of BAS are around 34.3%. Arterial hypotension and metabolic acidosis made up the most part (21.8%), followed by ventricular fibrillation and supraventricular paroxysmal tachycardia (9.3%), and lastly, deep vein thrombosis (3.3%). (Cirstoveanu *et al.*, 2023) Another concern with BAS is its procedural complications. This can be classified into different types of complications: mechanical (e.g. ruptures, deflations), traumatic (e.g. mitral valve damage, rupture of left atrial appendage), embolic (e.g. ischemia) and electrophysical (e.g. transitory heart rhythm disorders.) Although there are minimal risks using the femoral or umbilical approach, both methods have chances of failure. The former could be due to the difficulty of penetrating the atrial septum, whilst the latter could be the impermeability of the ductus venosus. (Cirstoveanu *et al.*, 2023) Other than this, critical cases where there is a delay in the procedure due to transportation issues could have caused adverse outcomes for BAS. As mentioned, time is crucial for patients with TGA. (Doshi, Venugopal and MacArthur, 2012)

### **3. Nutritional suggestions for patients before surgery**

#### **Enteral feeding**

Enteral feeding is often recommended for neonates who cannot orally intake food which would be the case for most TGA patients. This method is where nutrition is directly delivered to the stomach or small intestine using a feeding tube. There are many advantages of this, which include lowering the risk of infection, preserving the barrier function of the gut, decreased cost, effectiveness and safety. (Adeyinka *et al.*, 2024) In a test reviewing 18 neonates with TGA, it is shown that they have gained lots of body weight after the operation which suggests that enteral feeding is effective for the growth of neonates. (Ebishima, Kurosaki and Shiraishi, 2015) Other research indicates that the lack of enteral feeding may be correlated with increased length of postoperative care in the hospital. (Séguéla *et al.*, 2017)

However, it is still controversial whether enteral feeding for neonates with unstable haemodynamics is more beneficial than risky, especially during PGE1 therapy. Some research has shown that PGE1 results in risks of necrotizing enterocolitis (NEC is an inflammation of the gastrointestinal tissue which can cause severe abdominal infections), which is suggested to be caused by the lack of blood flow due to the ductal steal. Additionally, the diastolic flow of blood in the heart from the descending aorta could cause mesenteric ischaemia (decreased blood flow to the small intestine.) In a retrospective study of the relationship between NEC and CHDs, 643 neonates with heart diseases were evaluated, and the results show that these neonates experience intestinal ischemia due to decreased systemic perfusion or pathophysiology of the disease whilst 97% of neonates who underwent PGE1 had normal enteral feeding tolerance. (Willis *et al.*, 2008) In a larger study with 6710 neonates with ductal-dependent CHD, only 0.3% of neonates had NEC, so enteral feeding was not statistically significant in increasing the chances of NEC. (Becker *et al.*, 2014) Both of these data suggest that the ductal flow pattern does not affect NEC, so with the limited evidence, it is still recommended for TGA patients to be fed enterally due to its feasibility during the perioperative course. (Sarris *et al.*, 2017)

### **4. Summary of preoperative care for TGA**

In summary, monitoring needs to be administered immediately after diagnosis to come to a decision on the timing of surgery but also to carefully evaluate the need for other interventions. A combination of monitoring devices should be used to ensure its accuracy, but invasive procedures like cardiac catheterisation should be considered when weighing up the benefits and risks. For haemodynamic management, PGE1 can be started for patients with restrictive foramen ovale but many side effects could arise so keeping in mind the length of infusion is important. BAS is another procedure shown to reduce mortality rates before surgery, but similarly to PGE1, it is associated with procedural risks and other disorders. However, more research still needs to be done on the correlation between BAS and brain injuries. Another consideration is nutrition, and enteral feeding is usually the most efficient way. Nevertheless, it remains debatable whether enteral feeds for patients who underwent PGE1 infusion would be recommended. Therefore, the

preoperative management methods mentioned still need to be researched further, perhaps on a larger scale than the sources mentioned, to prevent possible side effects early on.

### **Postoperative considerations for Transposition of the Great Arteries**

Postoperative care after ASO mainly focuses on ensuring a safe homeostatic environment for neonates without affecting their metabolism. (Gupta *et al.*, 2014) The concern often regards the ability of the left ventricle to adapt and accommodate the work and the effectiveness of coronary blood flow. (Cabrera *et al.*, 2024) The aim for a steady haemodynamic state is defined by left atrial pressure (LAP) 5-8 mmHg, mean arterial pressure (MAP)  $\geq 45$  mmHg, systolic blood pressure (SBP)  $> 65$  mmHg, venous oxygen saturation (SvO<sub>2</sub>)  $\geq 60\%$  and lactate  $< 3$  mmol/L. (Anders, 2013) Many of the management methods are the same as those in preoperative care or common ICU management, so they will not be mentioned in detail (e.g. nutrition, renal management).

Postoperative outcomes vary for TGA with other CHDs, including IVS and VSD which will be discussed separately. Other cases that also require more attention include cyanosis and pulmonary ventricular resistance (PVR.) Cyanosis could delay anaesthesia since children lack effective tissue perfusion, so anaesthesia could reduce oxygenation further. Moreover, in neonates with high PVR, there is a lot of resistance against the flow of the pulmonary vein to the left atrium, which could lead to pulmonary hypertension and hypoxia. (Sarris *et al.*, 2017) In addition, the postoperative route is often dependent on the surgery outcomes and complications: coronary artery issues, low cardiac output syndrome (LCOS), arrhythmias and chylothorax (the accumulation of chyle, lymph, accumulating in the chest cavity). (Cabrera *et al.*, 2024) Monitoring methods during this postoperative period will confirm the complication and specific anaesthesia will be given when suitable. This period of management will be distinctive for each patient.

## **1. Haemodynamic Monitoring after surgery for TGA**

### **Non-invasive approaches**

A. ***Transesophageal echocardiography (TEE)*** can be used to monitor ventricular performance, check for coexisting diseases and evaluate the surgical results, checking for obstruction or leaks (including the presence of VSD.) (Sarris *et al.*, 2017) In a study evaluating the use of TEE for TGA, they concluded that its longitudinal plane is suited to observe the morphologic features of the connection between the right ventricle and the aorta, whilst the horizontal plane is used to evaluate the situs (position) of the atrial and atrioventricular function. (Caso *et al.*, 1998) The patient's ability to go through TEE needs to be checked, especially for patients with haemodynamic compromise from left atrial compression. Although rare, risks of TEE could include breathing problems, heart rhythm issues and bleeding. (Johns Hopkins Medicine, 2021)

B. Blood circulation and adequacy of mixing can be evaluated using ***pulse oximetry*** and NIRS. For patients with cyanosis, the pulse oximetry can overestimate the SaO<sub>2</sub> combined with a

decrease in partial pressure of arterial oxygen (PaO<sub>2</sub>.) As mentioned previously, NIRS can be used to monitor cerebral and tissue perfusion, so both techniques will be able to evaluate oxygen trends postoperatively as well as preoperatively. However, more studies still need to be done on its accuracy. (Sarris *et al.*, 2017)

### Invasive approaches

**Central venous catheterisation** can be used to monitor the central venous pressure (CVP), which is the pressure in the superior vena cava, and this highlights the pressure in the right atrium when filled with blood. There are usually minimal complications with this method, but it could cause bleeding, infection, and puncturing of the lungs, which may lead to a pneumothorax (collection of air-fluid in the pleural cavity.) (Cascino and Shea, 2023) Left atrial pressure (LAP) is also useful and accurate for monitoring the filling of the left ventricle after ASO for TGA and can be used during weaning from cardiopulmonary bypass (CPB.) Compared to measuring the CVP, LAP is more accurate and could be used until the haemodynamic status of patients is stable. (Akhtar *et al.*, 2013) On the first day after ASO, the systolic arterial pressure should be around 50-60 mmHg and should continue to increase over the next 72 hours. (Gupta *et al.*, 2014) The risks of catheterisation, in general, are bleeding and cardiac tamponade after removing the catheter. However, these complications are relatively rare (<1%), but blood transfusions are prepared beforehand just in case. (Gottlieb and Stayer, 2016)

## **2. Anaesthesia administration for stabilisation after surgery**

### Intravenous Fluid

Intravenous (IV) fluid is used to increase cardiac output, but it should be limited to 50% maintenance. This is because lots of fluid can increase the LAP, and increasing the LAP by 8-10 mmHg should be avoided, leading to haemodynamic deterioration. Moderate hypotension of around 35 mmHg with normal mixing of venous oxygen saturation is preferred. (Gupta *et al.*, 2014) Another important factor is managing the preload (initial stretching of cardiac muscle cells before contraction) since lots of preload could lead to ventricular dysfunction and hypotension. Even small amounts of fluid (<5 mL in a single bolus injection) could lead to these risks, which is why monitoring is extremely important during postoperative care. (Cabrera *et al.*, 2024)

### Inotropic vasodilator therapy

Vasoactive drugs are commonly used to support patients with cardiac issues by causing either vasodilation or vasoconstriction, which enhances tissue perfusion and oxygenation. Patients will have already been given milrinone (0.25-0.75 mcg/kg/min), which helps strengthen heart muscles and widen blood vessels to increase cardiac output and reduce pressure in the left ventricle. Epinephrine (0.02-0.05) and low doses of dobutamine and adrenaline (0.01–0.05 mg/kg/min) could also be given until the patient becomes stable. (Gupta *et al.*, 2014) To optimise target organ perfusion and improve urine output, vasopressin and low doses of phenylephrine and norepinephrine can be used. (Sarris *et al.*, 2017) In cases where the patient is not responding

to these drugs due to poor left ventricular performance, other methods like core cooling, paralysis and peritoneal dialysis can be used. If these also fail, then mechanical support devices need to be used (refer to the section on mechanical support). (Kang *et al.*, 2004) Like the IV fluid, the principle of minimal effective dosing should be strictly followed since high inotropic support, although it might ensure stable haemodynamics, could lead to progressive lactic acidosis (built-up lactic acid in the bloodstream, which can lead to organ dysfunction) and progressive alteration of tissue perfusion markers. (Sarris *et al.*, 2017)

### **3. Mechanical support and ventilation**

#### **Mechanical Support**

As mentioned briefly, in cases where stable haemodynamics are difficult to achieve even with inotropic support, patients require mechanical circulator support like extracorporeal membrane oxygenation (ECMO) or ventricular assist device (VAD.) Using these devices for a short period allows for rapid recovery to be achieved. Research reveals that around 20% of patients who underwent ASO require the ECMO after six weeks of life. (Sarris *et al.*, 2017) However, in research to test the effectiveness of patients with CHD undergoing ECMO, two of which were TGA patients, only one of the two survived after weaning off ECMO. (Boonthim, Namchaisiri and Benjacholamas, 2019) So, in these cases, the long-term effects of weaning off ECMO need to be considered to ensure results are acceptable and would not further diminish the chances of recovery or survival.

#### **Mechanical ventilation**

In mechanical ventilation, the focus is maintaining homeostasis with optimal conditions being no hypoxemia, pH 7.35-7.45, PaCO<sub>2</sub> 35-45 mmHg and SaO<sub>2</sub> > 95%. However, these could be adjusted depending on the specific conditions of the patient. Patients should be ventilated for only 12-24 hours after surgery, except for patients with unstable haemodynamics or rhythm abnormalities, who should continue to be ventilated. (Sarris *et al.*, 2017) Weaning difficulty leads to prolonged ventilation, which could result in ventilator dependency, malnutrition, and ventilator-associated lung injury like pneumonia. This significantly impacts other factors like longer hospital stays and higher medical costs as well. (Wu *et al.*, 2022) Hence, this highlights the importance of individualised care for each patient to ensure that any risks are minimised and treatment options are weighed out in a way that would be the most beneficial to the patient.

### **4. Postoperative complications**

ASO poses quite a few complications which are managed depending on the side effects and risk factors. Below are some common complications after the ASO (Table 1). Left ventricular dysfunction and pulmonary hypertension are also conditions which can arise after ASO but they can also be associated with TGA so it has been explained in detail in the 'Patient type considerations' section of this review. Other complications not mentioned below are, for



example, central nervous system complications and renal failure which are treated independently from cardiovascular diseases.

Table 1: Complications of ASO and the postoperative management method for each

<b><u>Complication</u></b>	<b><u>Explanation</u></b>	<b><u>Postoperative management</u></b>
<i>Low Cardiac Output Syndrome</i>	Low cardiac output syndrome (LCOS) is when the cardiac output is insufficient for oxygen perfusion. This is a complication related to CPB, myocardial preservation, post-inflammatory effects of the bypass and other factors. It is also more likely to occur because of surgery delays. (Datt <i>et al.</i> , 2022)	<ul style="list-style-type: none"> <li>- monitoring by checking against criteria for LCOS including but not limited to systolic blood pressure (decreased by &gt; 20%), SvO2 (&lt; 60%), central and peripheral temperature (&gt; 5°C), decreased cerebral NIRS (Datt <i>et al.</i>, 2022)</li> <li>- vasoactive, inotropic and lusitropic drugs to improve tissue perfusion and cardiac output (Sarris <i>et al.</i>, 2017)</li> <li>- suspicion of coronary insufficiency should be addressed by coronary artery evaluation by echocardiography or return to cardiac catheterisation (Cabrera <i>et al.</i>, 2024)</li> <li>- signs of left ventricular dysfunction should also be evaluated (Sarris <i>et al.</i>, 2017)</li> </ul>
<i>Coronary artery issues</i>	The ASO is able to manage coronary artery patterns but some complex patterns result in increased mortality, especially in intramural coronary arteries. One concern is coronary arterial stenosis (buildup of plaques in the coronary artery which narrows it) since it could be fatal. This is often caused by high tension in the coronary artery, operative injury and anastomotic stenosis. (Wu <i>et al.</i> , 2016)	<ul style="list-style-type: none"> <li>- monitoring of vital signs and haemodynamics: LCOS, arrhythmia, early arterial insufficiency, ventricular dysfunction are signs of coronary artery issues</li> <li>- reoperation/surgeries performed immediately (Wu <i>et al.</i>, 2016)</li> </ul>
<i>Arrhythmia</i>	Arrhythmia is the irregular rhythm of the heartbeat which can decrease cardiac output. This can be caused during the ASO by irritation of the myocardium. Arrhythmias result in increased mortality rates and the most common type of arrhythmia after ASO is junctional ectopic tachycardia (JET), which causes a fast heart rate by the atrioventricular junction. (Sarris <i>et al.</i> , 2017)	<ul style="list-style-type: none"> <li>- ECG to follow ischemic changes</li> <li>- haemodynamic management: immediate resuscitation if the patient is hemodynamically unstable, optimise ventilation, repositioning of intrathoracic lines, correction of electrolyte and acid-base imbalances</li> <li>- managing JET: decreasing vasoactive drugs, controlling hypothermia, optimised use of electrolytes, sedatives and pacemaker</li> <li>- antiarrhythmic drugs (e.g. esmolol, procainamide, amiodarone) titrated as continuous infusions (Sarris <i>et al.</i>, 2017)</li> </ul>

<i>Haemorrhage and tamponade</i>	Haemorrhage and cardiac tamponade are caused by prolonged CPB, suture lines or anatomic dissection during ASO. Persistent bleeding can be life-threatening so they need to be managed immediately. (Gupta <i>et al.</i> , 2014)	- prevention after CPB: frozen plasma and platelets to prevent coagulopathy (impaired blood clotting) - mediastinal (part of the chest between the lungs) exploration: re-exploration when chest drainage exceeds 5 mL/kg/hr in the first hour and 4 mL/kg/hr in the second (Gupta <i>et al.</i> , 2014)
<i>Capillary leak syndrome</i>	Capillary leak syndrome is when the plasma leaks out of blood capillaries. This is caused by the CPB and worsens with higher bypass and cross-clamp time, deep hypothermic circulatory arrest or if the patient has sepsis. (Gupta <i>et al.</i> , 2014)	- replacement of lost fluid with fresh frozen plasma or 5% albumin using peritoneal dialysis catheter - vasoactive drugs like vasopressin are used to maintain systemic arterial pressure (Gupta <i>et al.</i> , 2014)

**5. Summary of postoperative care for TGA**

For postoperative care, the monitoring techniques are essentially identical to preoperative care. However, the aim for this postoperative period is patient stabilisation so different anaesthesia should be administered with dosage limits for each patient. In certain circumstances where stabilisation still cannot be achieved, mechanical support including mechanical ventilation may be required. An important consideration for all these management methods is the period of treatment and weaning off the medication or mechanical support which could affect the patient’s survival. Finally, complications as a result of surgery are inevitable, so the method of management during this period would vary, and healthcare professionals would need to prioritise certain conditions that are more critical than others.

**Patient type considerations: TGA with other CHDs**

TGA can be associated with other CHDs which results in additional preoperative and postoperative care methods or further complications which need to be considered. The most common ones are presented in the table below.

**Associated CHD (with the percentage of cases in TGA patients) and their pre and post-operative care procedures**

TGA/IVS (70%)

Patients who have TGA with intact ventricular septum (TGA/IVS) require surgery since the left ventricular muscle mass, which used to be equal to the right ventricle before birth, will become deconditioned as the pulmonary ventilation rate decreases. As a result, the left ventricle will not be able to manage the systemic workload. Patients usually undergo ASO since its mortality rate is quite low (2-5%) and the best timing for surgery is during the first three weeks of life. (Sarris *et al.*, 2017)

A. **Preoperative care:** PGE1 maintains PDA, which increases the left atrial pressure, allowing adequate mixing. Without adequate opening of the interatrial region even with PGE1, BAS must be a priority. Early ASO is preferred, so BAS is performed for specific cases. (Gupta *et al.*, 2014) However, some patients have extreme cyanosis, acidosis or severe hypoxia so they fail to respond to these therapies. This is very dangerous so ECMO could be used before the patient is ready for an ASO. (Datt *et al.*, 2022)

Methods to increase the mixed venous oxygen saturation include decreasing oxygen consumption (using mechanical ventilation, sedation, and muscle relaxants) and improving oxygen delivery. However, it is worth noting that increasing oxygen saturation can close the PDA which reduces mixing and mechanical ventilation could lead to alveolar overdistension (expansion of the alveoli which could lead to lung damage.) (Gupta *et al.*, 2014)

B. **Postoperative care:** ASO for TGA/IVS is different in the way that a ventricle pumping at low pulmonary resistance is put in a high-resistance circuit. Thus, it can take weeks for the left ventricle to adapt to the ASO. To manage this, monitoring the electrolyte balance, temperature, vascular resistance, and other factors will ensure high flow but low pressure. (Gupta *et al.*, 2014) Sedation by morphine is required for all patients and other inotropes can be added if other side effects are present. When the patient has become stable, weaning off the ventilator and muscle relaxants should be done. (Gupta *et al.*, 2014)

#### TGA/VSD (25%)

Although less common, TGA/VSD allows adequate mixing of the two circuits (when there are no obstructive lesions.) However, VSD leads to excessive ventricular workload which can cause damage to blood vessels and lead to heart failure. Some signs include tachypnoea (respiratory rate greater than normal), tachycardia (faster heart rate than normal), abnormal heart rhythms, weight gain in the first 3-6 weeks and can develop into hepatomegaly (enlarged liver.) (Gupta *et al.*, 2014) Without operation for TGA/VSD, the life expectancy is on average 22 months proving that operations would significantly affect the patient's survival. (Datt *et al.*, 2022)

A. **Preoperative care:** The echocardiogram will be able to identify the VSD and any other anatomical or functional information (e.g. the number and size of the VSD). (Datt *et al.*, 2022) The reliable result of the echocardiogram reduces the need for cardiac catheterisation and angiography which are invasive forms of monitoring. (Serraf *et al.*, 2019) In a research conducted on 100 patients with TGA/VSD, 50% of patients underwent BAS and PGE1 infusion (0.025 Jlg/kg/min) and maintained haemodynamic stability in 22 patients. (Serraf *et al.*, 2019)

The VSD should be closed prior to the ASO, and this can be done by using a patch, or if it is small, it can be left open. The patient will also have high pressure in the left ventricle, so the ASO can be done in a few weeks to correct this. (Gupta *et al.*, 2014) Abnormalities of the pulmonary artery and the aorta can often be seen alongside VSD, which is why correction is

required in the postoperative period. The patients underwent corrective procedures during the first months of life or during early infancy. (Serraf *et al.*, 2019)

**B. Postoperative care:** In the Serraf *et al.* 2019 research mentioned earlier, some outcomes include 2 patients having compulsory pacemakers for atrioventricular block and sick sinus syndrome. (Serraf *et al.*, 2019) Moreover, VSD is associated with seizures. (Wernovsky *et al.*, 1995) However, the VSD closure did not affect rhythm disturbances, and statistical tests deny any significance for operative mortality. (Serraf *et al.*, 2019) In research that compared TGA/IVS and TGA/VSD, results indicate that TGA/VSD requires longer total support and myocardial ischemia. Since patients with TGA/VSD have TGA/VSD requires longer total support and myocardial ischemia. Since patients with TGA/VSD have slightly better cardiac indexes than those with TGA/IVS on the first night, inotropic support is less needed. (Wernovsky *et al.*, 1995)

#### TGA/LVOTO (20-30%)

Left ventricle outflow tract obstruction (LVOTO) is where there is limited flow of blood from the left ventricle, which could result in the failure of the left ventricle if it is not treated. (Vilcant and Hai, 2024) It is common in TGA and is usually present alongside VSD. The selection criteria for ASO remain unspecified. Usually, if LVOTO is dynamic (no LVOT abnormalities that need surgery), an ASO can be performed before fixing the LVOTO, but if LVOTO is fixed (LVOT abnormalities present) and severe, ASO may not be the best option. (Wu *et al.*, 2010)

**A. Preoperative care:** A 2-d and Doppler echocardiogram are adequate to provide an image of the LVOTO. However, some anomalies are still difficult to visualise even with the echocardiogram so there are still ongoing debates on the method of diagnosis. (Wu *et al.*, 2010) Not much is mentioned about the care specifically for LVOTO so presumably they follow the guidelines for usual TGA management. However, the timing of surgery and precautions would need to be considered, for example, whether the LVOTO should be fixed before or after the surgery or whether the VSD needs to be fixed before LVOTO.

**B. Postoperative care:** In research comparing groups with fixed and dynamic TGA/LVOTO, both groups have similar chances of neo-aortic stenosis, pulmonary regurgitation, and pulmonary arterial stenosis. Some of these patients underwent BAS and angioplasty. Another study reveals that out of 55 patients with TGA/LVOTO, 11% of deaths were caused by LVOTO, and its long-term outcomes after ASO show that only 5% had LVOTO in the latest follow-up. (Kalfa *et al.*, 2013) Thus, ASO seems to be the most preferred surgical method for TGA/LVOTO.

#### TGA/PPHN (1-3%)

Persistent pulmonary hypertension (PPHN) is where the normal pulmonary vascular adaptation after birth fails. In basic terms, the neonate does not adapt to the new circulatory system outside of the uterus where oxygen is provided by the placenta. As a result, pulmonary vascular resistance increases and the deoxygenated blood is shunted in the systemic circulation, leading to severe hypoxemia. TGA/PPHN is a condition with high risk and can be fatal. It has been speculated that the cause of PPHN is due to restricted PFO or closure of the ductus arteriosus. Echocardiography will be able to identify the presence of TGA and the risk of PPHN. (Roofthoof *et al.*, 2007)

**A. Preoperative care:** Pulmonary vascular resistance is caused by the shunting of the ductus arteriosus, bypassing the pulmonary circulation. Thus, PGE1 could further decrease the amount of blood bypassing the lungs, resulting in hypoxia and pulmonary artery constriction. If PGE1 must be used, cautious monitoring using the echocardiogram, acid/base monitoring and cerebral oxygenation monitoring is vital. (Séguéla *et al.*, 2017) However, this suggestion is not representative because in an experiment, 14 TGA/PPHN patients were treated with PGE1, and 12 patients underwent BAS. (Roofthoof *et al.*, 2007) Thus, more research may need to evaluate the concern.

Common management methods include inotropic and vasodilator support, neuromuscular blockade, alkalisation and mechanical ventilation, which help to reduce pulmonary vascular resistance and should start immediately after birth. The main pulmonary vasodilator used is inhaled nitric oxide (iNO), which reduces respiratory distress and improves the prognosis of PPHN. (Séguéla *et al.*, 2017) This is proven by a case study where three neonates with TGA/PPHN were successfully treated with iNO. One of the neonates received iNO before BAS which improved drastically so that BAS could be delayed. (El-Segaier *et al.*, 2005) Nevertheless, another research shows that around 30% of neonates do not respond to iNO since PPHN is dependent on many factors. This means the patient becomes reliant on the ECMO. (Séguéla *et al.*, 2017)

**B. Postoperative care:** Some postoperative complications include atelectasis, chylothorax, junctional ectopic tachycardia, etc. (Roofthoof *et al.*, 2007) Another study indicates that patients with PPHN before ASO resulted in 54.5% of patients with postoperative PPHN complications. (Wongwaitaweewong *et al.*, 2022) Compared to TGA patients with normal pulmonary pressure, patients with TGA/PPHN require longer ventilatory support (around 8.2 days compared to 3.9 days for TGA without PPHN), ECMO and longer hospital stays (10.4 compared to 5.5 days). (Roofthoof *et al.*, 2007) These clearly result in high risks, which have been mentioned in the postoperative section, which is another reason why PPHN is complicated to deal with even after a successful surgery.

## **Conclusions and Outlook**

TGA remains a CHD which is difficult to manage due to the associated conditions and its many complications before or after surgery. Thus, by optimising both preoperative and postoperative care, the best outcomes for the patient could be achieved. This includes evaluating the effectiveness of procedures or medication currently being utilised and the risks or complications as well. Further data would need to be collected in areas such as haemodynamic management since some controversial aspects of these methods have yet to be resolved. Additionally, the long-term outcomes of patients are another factor in determining how safe the methods are and whether complications could be present long after their discharge. Current studies on TGA revolve around new technologies that make diagnosis, monitoring, and management more accurate, whilst other reviews discuss the quality of life of patients.

## Works Cited

- Adeyinka, A., Rouster, A.S. and Valentine, M. (2024) 'Enteric Feedings', in StatPearls. Treasure Island (FL): StatPearls Publishing. Available at: <http://www.ncbi.nlm.nih.gov/books/NBK532876/> (Accessed: 2 August 2024).
- Akhtar, M.I. et al. (2013) 'Transatrial access for left atrial pressure (LAP) monitoring line placement in arterial switch operation (ASO) in neonates', 63(11).
- Amouzesi, A. et al. (2021) 'Comparison of pulse oximetry and clinical examination for proper tissue perfusion in congenital heart patients'. Available at: <https://jsurgery.bums.ac.ir/article-1-272-en.pdf>.
- Anders, M. (2013) 'Transposition of Great Arteries', Don't Forget The Bubbles [Preprint]. Available at: <https://doi.org/10.31440/DFTB.3504>.
- Ashraf, A. (2020) Prostaglandin-induced cortical hyperostosis in infants with cyanotic congenital heart disease | Radiology Reference Article | Radiopaedia.org, Radiopaedia. Available at: <https://doi.org/10.53347/rID-78514>.
- Beca, J. et al. (2009) 'Pre-Operative Brain Injury in Newborn Infants With Transposition of the Great Arteries Occurs at Rates Similar to Other Complex Congenital Heart Disease and Is Not Related to Balloon Atrial Septostomy', *Journal of the American College of Cardiology*, 53(19), pp. 1807–1811. Available at: <https://doi.org/10.1016/j.jacc.2009.01.061>.
- Becker, K.C. et al. (2014) 'Necrotizing Enterocolitis in Infants with Ductal-Dependent Congenital Heart Disease', *American Journal of Perinatology*, 32, pp. 633–638. Available at: <https://doi.org/10.1055/s-0034-1390349>.
- Beerman, L.B. (2023) Transposition of the Great Arteries (TGA) - Pediatrics, MSD Manual Professional Edition. Available at: <https://www.msmanuals.com/professional/pediatrics/congenital-cardiovascular-anomalies/transposition-of-the-great-arteries-tga> (Accessed: 15 July 2024).
- Boonthim, K., Namchaisiri, J. and Benjacholas, V. (2019) 'ECMO support after arterial switch operation: A case report series from a single institution', *Chulalongkorn Medical Journal*, 63(3). Available at: <https://he05.tci-thaijo.org/index.php/CMJ/article/view/301> (Accessed: 18 August 2024).
- Cabrera, A.G. et al. (2024) 'Transposition of the Great Arteries'.
- Cascino, T. and Shea, M.J. (2023) Central Venous Catheterization - Heart and Blood Vessel Disorders, MSD Manual Consumer Version. Available at: <https://www.msmanuals.com/home/heart-and-blood-vessel-disorders/diagnosis-of-heart-and-blood-vessel-disorders/central-venous-catheterization> (Accessed: 8 August 2024).
- Caso, P. et al. (1998) 'Diagnostic value of transesophageal echocardiography in the assessment of congenitally corrected transposition of the great arteries in adult patients', *American Heart Journal*, 135(1), pp. 43–50. Available at: [https://doi.org/10.1016/s0002-8703\(98\)70341-2](https://doi.org/10.1016/s0002-8703(98)70341-2).
- Cheung, A. et al. (2022) 'Detection of hypoxia by near-infrared spectroscopy and pulse oximetry: a comparative study', *Journal of Biomedical Optics*, 27(7), p. 077001. Available at: <https://doi.org/10.1117/1.JBO.27.7.077001>.

- Cirstoveanu, C. et al. (2023) ‘Impact of Bedside Balloon Atrial Septostomy in Neonates with Transposition of the Great Arteries in a Neonatal Intensive Care Unit in Romania’, *Life*, 13(4), p. 997. Available at: <https://doi.org/10.3390/life13040997>.
- Cohen, M.S. and Mertens, L.L. (2019) ‘EDUCATIONAL SERIES IN CONGENITAL HEART DISEASE: Echocardiographic assessment of transposition of the great arteries and congenitally corrected transposition of the great arteries’, *Echo Research and Practice*, 6(4), pp. R107–R119. Available at: <https://doi.org/10.1530/ERP-19-0047>.
- Datt, V. et al. (2022) ‘Perioperative anesthetic management of transposition of great arteries: a review’, *The Egyptian Journal of Cardiothoracic Anesthesia*, 16(2), p. 23. Available at: [https://doi.org/10.4103/ejca.ejca\\_17\\_21](https://doi.org/10.4103/ejca.ejca_17_21).
- Doshi, H., Venugopal, P. and MacArthur, K. (2012) ‘Does a balloon atrial septostomy performed before arterial switch surgery increase adverse neurological outcomes?’, *Interactive Cardiovascular and Thoracic Surgery*, 15(1), pp. 141–143. Available at: <https://doi.org/10.1093/icvts/ivr145>.
- Ebishima, H., Kurosaki, K. and Shiraishi, I. (2015) The Effects For Growth of Perioperative Enteral Feeding in Neonates With Tga | *Intensive Care Medicine Experimental* | Full Text. Available at: <https://icm-experimental.springeropen.com/articles/10.1186/2197-425X-3-S1-A495> (Accessed: 2 August 2024).
- El-Segaier, M., Hellström-Westas, L. and Wettrell, G. (2005) ‘Nitric oxide in neonatal transposition of the great arteries’, *Acta Paediatrica* (Oslo, Norway: 1992), 94(7), pp. 912–916. Available at: <https://doi.org/10.1111/j.1651-2227.2005.tb02010.x>.
- Ghumman, B., Kumar, A. and Kumar, S. (2019) ‘Inadequacy of Near-Infrared Spectroscopy Cerebral Oximetry Monitoring for Detecting Neurological Complication’, *Annals of Cardiac Anaesthesia*, 22(3), pp. 321–324. Available at: [https://doi.org/10.4103/aca.ACA\\_147\\_18](https://doi.org/10.4103/aca.ACA_147_18).
- Gottlieb, E.A. and Stayer, S.A. (2016) ‘A Single Catheter to Monitor Both Central Venous Pressure and Left Atrial Pressure\*’, *Pediatric Critical Care Medicine*, 17(3), p. 261. Available at: <https://doi.org/10.1097/PCC.0000000000000650>.
- Gupta, A.K. et al. (2014) ‘Perioperative Management for Transposition of Great Arteries’, *Journal of Pediatric Critical Care*, 1(3), p. 161. Available at: <https://doi.org/10.21304/2014.0103.00031>.
- Haeffele, C. and Lui, G.K. (2015) ‘Dextro-Transposition of the Great Arteries: Long-term Sequelae of Atrial and Arterial Switch’, *Cardiology Clinics*, 33(4), pp. 543–558. Available at: <https://doi.org/10.1016/j.ccl.2015.07.012>.
- InformedHealth.org (2023) ‘In brief: What happens during cardiac catheterization?’, in *InformedHealth.org* [Internet]. Institute for Quality and Efficiency in Health Care (IQWiG). Available at: <https://www.ncbi.nlm.nih.gov/books/NBK355300/> (Accessed: 1 August 2024).
- Johns Hopkins Medicine (2021) *Transesophageal Echocardiogram*. Available at: <https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/transesophageal-e-chocardiogram> (Accessed: 8 August 2024).



- Kalfa, D.M. et al. (2013) 'Arterial Switch for Transposition With Left Outflow Tract Obstruction: Outcomes and Risk Analysis', *The Annals of Thoracic Surgery*, 95(6), pp. 2097–2103. Available at: <https://doi.org/10.1016/j.athoracsur.2013.01.077>.
- Kang, N. et al. (2004) 'Extending the Boundaries of the Primary Arterial Switch Operation in Patients With Transposition of the Great Arteries and Intact Ventricular Septum', *Circulation*, 110(11\_suppl\_1), p. II–123. Available at: <https://doi.org/10.1161/01.CIR.0000138221.68312.33>.
- Lewis, A.B. et al. (1981) 'Side effects of therapy with prostaglandin E1 in infants with critical congenital heart disease', *Circulation*, 64(5), pp. 893–898. Available at: <https://doi.org/10.1161/01.cir.64.5.893>.
- Martins, P. and Castela, E. (2008) 'Transposition of the great arteries', *Orphanet Journal of Rare Diseases*, 3, p. 27. Available at: <https://doi.org/10.1186/1750-1172-3-27>.
- McQuillen, P.S. et al. (2006) 'Balloon atrial septostomy is associated with preoperative stroke in neonates with transposition of the great arteries', *Circulation*, 113(2), pp. 280–285. Available at: <https://doi.org/10.1161/CIRCULATIONAHA.105.566752>.
- Mukherjee, D. et al. (2010) 'Analysis of 8681 neonates with transposition of the great arteries: outcomes with and without Rashkind balloon atrial septostomy', *Cardiology in the Young*, 20(4), pp. 373–380. Available at: <https://doi.org/10.1017/S1047951110000296>.
- Pérez Moreno, J.C. et al. (2018) 'Cerebral oximetry monitoring in the management of severe hypoxaemia associated with transposition of the great arteries with balloon atrial septostomy', *Revista Española de Anestesiología y Reanimación (English Edition)*, 65(5), pp. 294–297. Available at: <https://doi.org/10.1016/j.redare.2018.04.004>.
- Roofthoof, M.T.R. et al. (2007) 'Persistent Pulmonary Hypertension of the Newborn With Transposition of the Great Arteries', *The Annals of Thoracic Surgery*, 83(4), pp. 1446–1450. Available at: <https://doi.org/10.1016/j.athoracsur.2006.11.001>.
- Sarris, G.E., (Chairperson) (Greece) et al. (2017) 'Clinical guidelines for the management of patients with transposition of the great arteries with intact ventricular septum', *European Journal of Cardio-Thoracic Surgery*, 51(1), pp. e1–e32. Available at: <https://doi.org/10.1093/ejcts/ezw360>.
- Séguéla, P.-E. et al. (2017) 'Transposition of the great arteries: Rationale for tailored preoperative management', *Archives of Cardiovascular Diseases*, 110(2), pp. 124–134. Available at: <https://doi.org/10.1016/j.acvd.2016.11.002>.
- Serraf, A. et al. (2019) Anatomic correction of transposition of the great arteries with ventricular septal defect: Experience with 118 cases - ScienceDirect. Available at: <https://www.sciencedirect.com/science/article/pii/S0022522319365936> (Accessed: 19 August 2024).
- Soongswang, J. et al. (2000) 'Limitation of Transthoracic Echocardiography in the Diagnosis of Congenital Heart Diseases', 83.

- Szymanski, M.W. et al. (2023) ‘Transposition of the Great Arteries’, in StatPearls. Treasure Island (FL): StatPearls Publishing. Available at: <http://www.ncbi.nlm.nih.gov/books/NBK538434/> (Accessed: 20 June 2024).
- Tabbutt, S., Helfaer, M.A. and Nichols, D.G. (2006) ‘Chapter 7 - Pharmacology of Cardiovascular Drugs’, in D.G. Nichols et al. (eds) *Critical Heart Disease in Infants and Children (Second Edition)*. Philadelphia: Mosby, pp. 173–203. Available at: <https://doi.org/10.1016/B978-032301281-2.50009-6>.
- Teixeira, O.H.P. et al. (1983) Long-term prostaglandin E1 therapy in congenital heart defects - ScienceDirect. Available at: <https://www.sciencedirect.com/science/article/pii/S0735109784802624> (Accessed: 15 July 2024).
- Therapeutic Goods Administration (2022) Limitations of pulse oximeters and the effect of skin pigmentation | Therapeutic Goods Administration (TGA). Therapeutic Goods Administration (TGA). Available at: <https://www.tga.gov.au/news/safety-updates/limitations-pulse-oximeters-and-effect-skin-pigmentation> (Accessed: 1 August 2024).
- Vilcant, V. and Hai, O. (2024) ‘Left Ventricular Outflow Tract Obstruction’, in StatPearls. Treasure Island (FL): StatPearls Publishing. Available at: <http://www.ncbi.nlm.nih.gov/books/NBK470446/> (Accessed: 21 August 2024).
- Villafañe, J. et al. (2014) ‘D-TRANSPOSITION OF THE GREAT ARTERIES: Hot Topics in the Current Era of the Arterial Switch Operation’, *Journal of the American College of Cardiology*, 64(5), pp. 498–511. Available at: <https://doi.org/10.1016/j.jacc.2014.06.1150>.
- Wernovsky, G. et al. (1995) ‘Postoperative Course and Hemodynamic Profile After the Arterial Switch Operation in Neonates and Infants’, *Circulation*, 92(8), pp. 2226–2235. Available at: <https://doi.org/10.1161/01.CIR.92.8.2226>.
- Willis, L. et al. (2008) ‘Enteral Feeding in Prostaglandin-Dependent Neonates: Is It a Safe Practice?’, *The Journal of Pediatrics*, 153(6), pp. 867–869. Available at: <https://doi.org/10.1016/j.jpeds.2008.04.074>.
- Wongwaitawee Wong, K. et al. (2022) ‘Predictors of persistent pulmonary hypertension of the newborn in D-transposition of the great arteries: experiences at Songklanagarind Hospital, Songkhla’. Available at: <https://doi.org/10.21203/rs.3.rs-1764466/v1>.
- Wu, K.-L. et al. (2010) ‘Left Ventricular Outflow Tract Obstruction in Complete Transposition of the Great Arteries - Echocardiography Criteria for Surgical Strategies -’, *Circulation Journal*, 74(6), pp. 1214–1218. Available at: <https://doi.org/10.1253/circj.CJ-09-0081>.
- Wu, Q.-Y. et al. (2016) ‘Surgical Treatment of Complete Transposition of the Great Arteries in Newborn’, *Chinese Medical Journal*, 129(19), pp. 2381–2383. Available at: <https://doi.org/10.4103/0366-6999.190663>.
- Wu, Xiaoming et al. (2022) ‘Individualized Analysis and Treatment of Difficult Weaning From Ventilation Following Open Cardiac Surgery in Young Children With Congenital Heart Disease’, *Frontiers in Cardiovascular Medicine*, 9, p. 768904. Available at: <https://doi.org/10.3389/fcvm.2022.768904>.

# Reading Help for the Visually Impaired: the Case of Talking Ads By Stefan Maxim

## Abstract

Most existing grocery ads for visually impaired people are processed and recited manually, which can be very time-consuming. Furthermore, there is a lack of appropriate tools available for optimizing and personalizing the audio. This paper proposes an automated method for generating audio ads for visually impaired people by transforming the traditional paper or online ads into an enhanced audio version that selectively speeds up audio portions based on the pitch of the words, all while organizing them based on categories. The generated audio files take advantage of the fact that visually impaired people are more capable of hearing high-pitched words and can process audio information at much higher speeds (2X to 3x) than most sighted people. By re-training a pre-trained model for image detection alongside a classical code base for text-to-audio and audio augmentation, we were able to vertically integrate the process of reading ads in a more compact and intuitive way. This has resulted in a method that is not only more accessible for those with visual impairments but still remains portable and light enough to be utilized on mobile devices.

## Introduction

For most people, AI and technology have greatly increased the rate of consumption of various visual media types like articles, newspapers, and journals. However, the same cannot be said about visually impaired people, as they are forced to find a different approach to interpreting media. Currently, approximately 304.1 million Rupert R A Bourne [2020] people have moderate to severe blindness while around 49.1 million are completely blind, a stark 42.8 % increase from the 34.4 million in 1990. While medical research is helping reduce the number of preventable cases, challenges like access to information and mobility remain a problem. Specifically, there is a gap of 55 words per minute (wpm) between the average reading rate (238wpm) and speaking rate (183wpm) per Rupert R A Bourne [2020]. This results in a 23 % increase in time spent by a person who consumes audio-only information. In response, this paper proposes a method to automate the process by which grocery adverts ( example ad in Figure 1.a ) are converted to audio form while decreasing the time to consume.

Figure 1.a illustrates the top-level diagram of the conversion application. It begins by processing the image and then uses ML detection to denote the picture, text, and price. It then classically sorts the ads, indexing at text level to create a searchable database organized by pre-established criteria such as the type of object or price. A selective speed-up of the low-pitch words is performed as they are split into phonemes. This is followed by the text-to-audio conversion using variable speed based on the user feedback and neurological research of the visually impaired. Figure 1.b shows the simplified diagram of the YOLOv8n (nano) single-pass real-time object detection region-based convolutional neural network that uses anchor-free detection, decoupled head and modified loss function for processing speed-up.

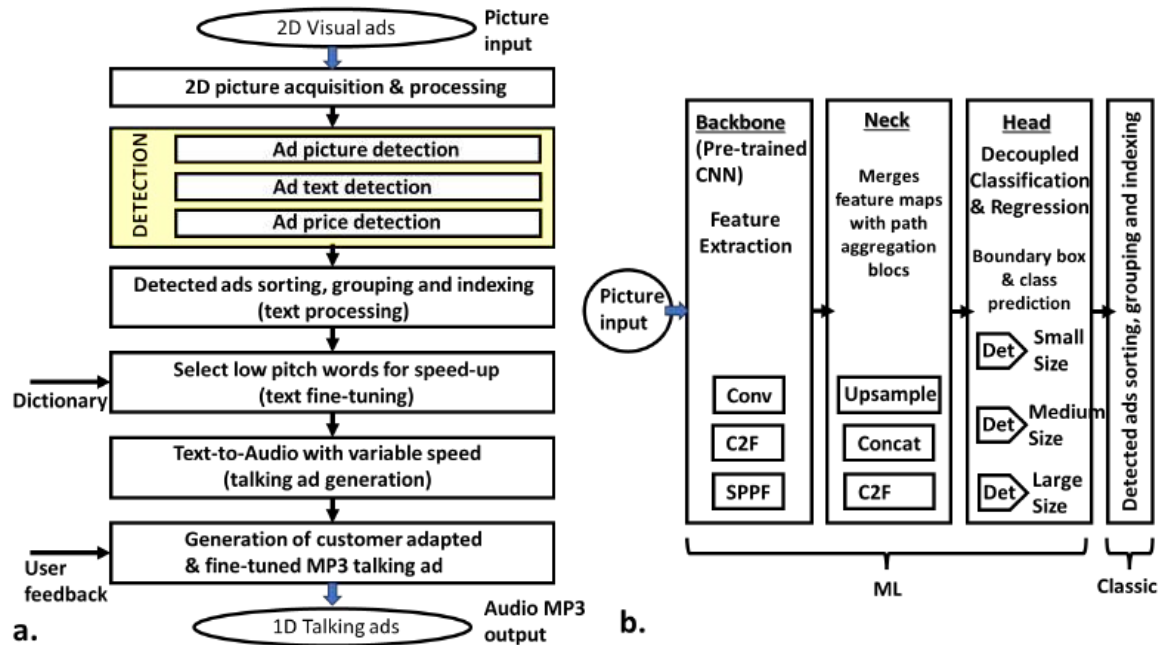


Figure 1: a. Bloc diagram of the proposed talking ads generation and fine-tuning using both ML and classic processing; b. YOLOv8n real-time single-pass object-detection R-CNN simplified diagram

## Related work

The most common method of converting the ads to audio form is via a human manually reciting them. Alternatively, one can use existing LLMs like ChatGPT, but it is very high-power and unwieldy for those with blindness. Furthermore, the LLMs output will still need to be recited either manually or automatically at the established 183 wpm, making the total time much slower than the 238 wpm Brysbaert [2019] of reading. Furthermore, when prompted to read an ad page, ChatGPT produced an ordered list of all the words from a page. However, when asked to make its results more useful for those with blindness, the result just becomes more verbose while keeping its sequentiality. This is the same way that ads are often manually read, which is often detrimental to the listeners as people usually skim until seeing parts they are interested in, something that this archaic method makes impossible.

Several papers Kawamura and Rekimoto [2023] propose varying audio speed to increase listening efficiency, however, these solutions are focused on sighted individuals. Many studies like Fields [2010] highlight the fact that blind people can understand speech at speeds 2-3 times faster than sighted individuals. Other studies like Wan et al. [2010], Occelli et al. [2016], and Gougoux et al. [2004] show that congenitally blind people perform better than sighted people on a range of verbal tasks like pitch discrimination (spectral and temporal ability to distinguish pitch changes Bertonati et al. [2021]), auditory memory (short term memory recall of words), and verbal fluency (both semantic and letter fluency). Instead of only focusing on these traits as neurological insights, I decided to use them to tailor the audio produced in my visual-to-audio program to be more efficient and allow greater listening speeds for people with visual impairments.

## Methodology

Our solution has three parts: automatic verbal ad generation, variable speech adaptation for low-pitch words, and improvements in text-to-audio translation.

### Building the pipeline

For detecting ads, we used the pre-trained model yolov8n because of performance, size (32m parameters), and ease of use Ultralytics. By looking through hundreds of ads, we noticed that most have the same format: a rectangular area with an image, a description, and a price. For our data, we used a combination of “real” images created by manually labeling ads as well as synthetic images (item pictures are sourced from Fruits360 dataset Molnár and generated text). We then built three training datasets: 200 synthetic, 200 real, and a combination of 200 real and 900 synthetic images, and one validation dataset with 50 real images. We are interested in 3 key metrics: the F1 score per class, the mean average precision (both for 0.5 and 0.95 IOU), and the amount of data needed to achieve a good accuracy. Our hypothesis is that our dataset is sufficient to achieve a  $>0.9$  mAP50 accuracy. After the text and the price are detected, we are using Easy OCR Jaided AI to convert them to text and gTTF Google to convert the text into speech.

### Accelerating low pitch words

Since visually impaired individuals are able to discern high-pitch sounds, we hypothesize that selectively increasing the speed of low-pitch words should maintain the intelligibility of a text. For each ad text, we decompose it into words and then further down into phonemes. For the words with low-pitch phonemes, we increased the speed of pronunciation. We created a control sample with all words at a constant speed (1) and several variable speed ads with speed values of 1.3X, 1.5X, 1.7X, 2.0X, and 2.7X respectively. We tested the resulting audio files on 33 sighted individuals. Each participant was given ads at both constant and variable speed and was asked to evaluate the audio on a 1-10 scale where 1 meant unintelligible and 10 meant clear.

### 2D to 1D transformation

We were then tasked with efficiently transforming the image (2D) into audio (1D). Our goal was to answer “what information can be eliminated from the 2D space when moving it to a 1D space?”. The 33 sighted individuals who participated in our study were asked to describe how they read the weekly ads, how long it takes, and what they remember from it.

## Results

We will start with the results from Section 3.1. Our experiments (source code, data and more results: Maxim, S) returned 6 models, each with a different combination of dataset size/composition and epochs, and obtained the mAP50 scores and mAP50-95 depicted in table 1.

mAP scores for each model			
Model Name	Notes	mAP50	mAP50-95
Model 1	200 synthetic images for 10 epochs	0.82705	0.46586
Model 2	200 real images for 10 epochs	0.97442	0.75793
Model 3	1100 hybrid images for 10 epochs	0.98654	0.76533
Model 4	200 synthetic images for 50 epochs	0.91534	0.54357
Model 5	200 real images for 50 epochs	0.94266	0.7137
Model 6	1100 hybrid images for 50 epochs	0.98649	0.73275

Table 1: Models with mAP50 and mAP50-90 metrics

Model 1, trained on synthetic data obtained a mAP50-95 score of approximately 0.46, which was less than that of the purely manual data Model 2, which obtained a value of 0.76. In Model 4, by increasing the epochs from Model 1, the corresponding mAP score increased to 0.54, an increase of around 0.077. Interestingly, in Model 5, the same increase in epochs from its predecessor Model 2 actually lowered the mAP score to 0.71, a decrease of around 0.044. This pattern held true with the composite dataset models (namely models 3 and 6), albeit to a less severe degree as the mAP score only decreased by 0.032, from around 0.77 to 0.73. Lastly, we depicted the F1 curve for the Best Model (Model 3) as well as that of Model 2, the second best, side by side in Figure 2. Qualitatively, we can see that the mAP, as well as the overall F1 score, seems to be tighter as well as higher for our best model. Analyzing the Section 3.2 data, it is observed that people can easily understand ads at speeds up to 1.5x with an average understanding score of 7.636 out of 10, yet this number drops to 4.182 for any higher speeds.

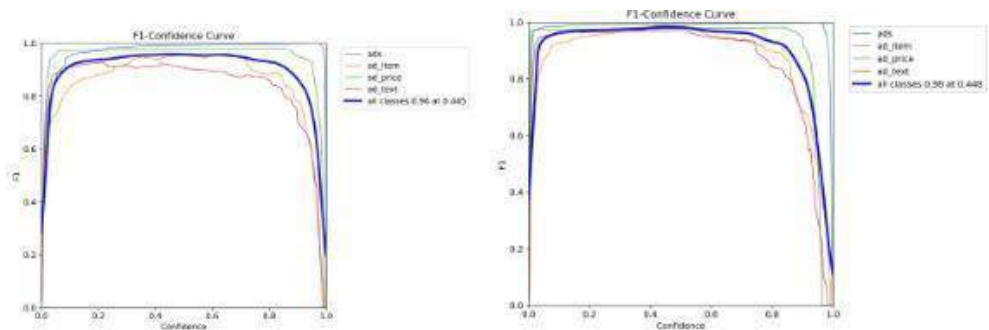


Figure 2: F1 curve for best and model trained with only real data

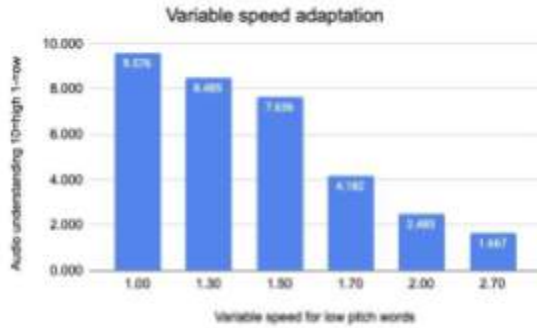


Figure 3: Understanding of an audio ad that has words at variable speed



Figure 4: Sample synthetic data

Our results in Section 3.3 found that people spend an average of 20 seconds per ad due to skimming the visuals and focusing on the category of interesting items. They then remembered 2-3 items per page.

## Discussion

The model we trained that took an image as input and detects the ad image, text, and price demonstrates the feasibility of our automated ad detection. It showed that the most effective way of training the model while minimizing labor costs would be a combination of real and synthetic data, thereby confirming the efficacy of our synthesized data (highlighted in Man and Chahl [2022] as well). We were also able to validate our hypothesis that speeding up low-pitch words in a sentence does not lead to information loss. The results show that speeds up to 1.5x do not significantly change the understanding of the content, which for visually impaired individuals who re-play the ad at even higher speeds, means the efficiency in ad listening is even greater. Lastly, we highlighted the need to process the text based on grouping rather than the current sequential processing. Our proposed method of grouping the items is in line with how ads are actually visualized.

## Conclusion and future work

We have studied the detection accuracy as a function of the dataset size, using both real grocery ads and synthetically generated ones. To ensure detection accuracy above the 95 threshold level, the training data needs to contain at least 1k synthetic data samples. When compared with the manually generated audio ads for visually impaired people, the proposed adaptive speed and selective emphasis of low-pitch words have resulted in at least 1.5x increase in speed for certain words. Using the automated ads grouping and index listing can result in 5 to 10x speed-up of the ad listening time, by jumping directly to the sections of interest. Due to the nature of grocery ads that are rectangular-based, these results can be generalized internationally. Additionally, an extension of this work can focus on improving further the 2D(image) to 1D(audio) translation at the full page level.

## Works Cited

- G. Bertonati, M. B. Amadeo, C. Campus, and M. Gori. Auditory speed processing in sighted and blind individuals. *PLOS ONE*, 16(9):e0257676, 2021. doi: 10.1371/journal.pone.0257676. URL <https://doi.org/10.1371/journal.pone.0257676>.
- Marc Brysbaert. How many words do we read per minute? a review and meta-analysis of reading rate. *Journal of Memory and Language*, 109:104047, 2019. ISSN 0749-596X.
- R. Douglas Fields. Why can some blind people process speech far faster than sighted persons? *Scientific American*, Dec 13 2010.
- Google. GTTS Documentation. <https://gtts.readthedocs.io/en/latest/>.
- Frédérique Gougoux, Franco Lepore, Maryse Lassonde, Patrice Voss, Robert J. Zatorre, and Pascal Belin. Neuropsychology: pitch discrimination in the early blind. *Nature*, 430(6997):309, Jul 15 2004. doi: 10.1038/430309a.
- Jaied AI. EasyOCR. <https://github.com/JaiedAI/EasyOCR>.
- Kazuki Kawamura and Jun Rekimoto. Aix speed: Playback speed optimization using listening comprehension of speech recognition models. In *Proceedings of the Augmented Humans International Conference 2023*, page 200–208, 2023.
- K. Man and J. Chahl. A review of synthetic image data and its use in computer vision. *J Imaging*, 8(11):310, Nov 21 2022. doi: 10.3390/jimaging8110310.
- Maxim, S. Code and Documentation. <https://github.com/StefanMaxim/AdsProject>.
- B. Molnár. Fruits 360 dataset <https://www.kaggle.com/datasets/moltean/fruits>.
- Valeria Occelli, Simon Lacey, Careese Stephens, and Krish Sathian. Superior verbal abilities in congenital blindness. *Society for Imaging Science and Technology*, 2016. doi: 10.2352/ISSN.2470-1173.2016.16HVEI-094. DOI: 10.2352/ISSN.2470-1173.2016.16HVEI-094.
- Seth Flaxman Paul Briant Michele Bottone Theo Vos Kovin Naidoo Tasanee Braithwaite Maria Cicinelli Jost Jonas Hans Limburg Serge Resnikoff Alex Silvester Vinay Nangia Hugh R Taylor Rupert R A Bourne, Jaimie Adelson. Global prevalence of blindness and distance and near vision impairment in 2020: progress towards the vision 2020 targets and what the future holds. In *Invest. Ophthalmol. Vis. Sci.*, page 61(7):2317, 2020.
- Ultralytics. Ultralytics Documentation. <https://docs.ultralytics.com>.
- Catherine Y. Wan, Amanda G. Wood, David C. Reutens, and Sarah J. Wilson. Early but not lateblindness leads to enhanced auditory perception. *Neuropsychologia*, 48(1):344–348, 2010. ISSN 0028-3932.



# ***Salmonella's* Type 6 Secretion System: Pathogenesis, Host Cell Manipulation, and Response**

## **By Vana Scorza**

### **Abstract**

Both *Salmonella enterica* and *typhi* are bacterial pathogens and cause food poisoning, typhoid fever, and gastroenteritis, which can be lethal. These two strains use the Type 6 secretion system (T6SS), a needle-like apparatus, to inject effectors, which are proteins produced by bacteria that manipulate host cell functions. Effectors can be injected into other bacterial or host cells, thereby inducing infection. Previous research has shown that the T6SS in *Salmonella* is crucial for bacterial pathogenicity, as it injects effector proteins that can manipulate host cells to promote infection and also allows *Salmonella* to outcompete other bacteria by inhibiting their growth. The central question this paper addresses is the Type 6 secretion system's role in pathogenicity as well as effector function and subsequent host cell response. This paper aims to clarify how T6SS contributes to *Salmonella's* ability to cause disease, which is key for developing targeted treatments. The findings reveal the structure and function of T6SS, its role in host cell manipulation, and its genetic regulation, emphasizing its importance in both infection and interbacterial competition.

### **Introduction**

*Salmonella* infections, which include typhoid fever and food poisoning, represent a substantial global health issue. These infections vary due to the different strains and serotypes of *Salmonella*. Typhoid fever, caused by *Salmonella typhi*, is mainly spread through contaminated food and water, leading to serious systemic illness. (CDC) In contrast, food poisoning is often linked to non-typhoidal *Salmonella* serotypes like *Salmonella enteritidis* and *Salmonella typhimurium*, and is typically associated with eating tainted food, resulting in gastroenteritis. The global impact of *Salmonella* infections is severe, with millions of cases and hundreds of thousands of deaths each year, emphasizing the urgent need to explore the mechanisms behind their pathogenicity. *S. typhi*, responsible for Typhoid fever, is transmitted mainly through contaminated food and water, presenting symptoms such as high fever, weakness, stomach pain, headache, diarrhea or constipation, cough, and loss of appetite. Typhoid fever affects approximately 9.2 million people worldwide each year, with an estimated 5,700 illnesses and 620 hospitalizations occurring in the United States, primarily among individuals who have traveled internationally (CDC). On the other hand, food poisoning, often caused by non-typhoidal *Salmonella* serotypes like *Salmonella enteritidis* and *Salmonella typhimurium*, results in about 1.35 million infections, 26,500 hospitalizations, and 420 deaths annually in the United States alone. This type of infection is typically spread through the consumption of contaminated food, particularly poultry, eggs, and dairy products, leading to symptoms like diarrhea, fever, abdominal cramps, and vomiting. Non-typhoidal *Salmonella* affects approximately 93.8 million people and causes 160,000 fatalities globally each year, while typhoid fever affects an estimated 9 million people and results in 110,000 deaths annually

(Lamichhane et al.). Together, these *Salmonella* infections lead to around 270,000 deaths each year, highlighting the significant global impact of this bacterium on public health.

A strain refers to a genetic variant or subtype of a microorganism, which is crucial for understanding the specific characteristics and behaviors of each variant (Robertson). Typhoidal *Salmonella*, such as *Salmonella typhi*, is adapted to humans and causes systemic infections, whereas non-typhoidal *Salmonella*, including various serotypes of *Salmonella enterica*, is responsible for gastroenteritis and can infect a wide range of hosts (Saleh et al.). Morphologically, *Salmonella* are rod-shaped, flagellated, and Gram-negative bacteria that are facultative anaerobes, capable of growing in both aerobic and anaerobic environments (Magana-Arachchi and Wanigatunge). They infect organs like the intestines and bloodstream, depending on the strain and infection type.

*Salmonella* species share common genes involved in virulence, survival, and adaptation, including genes related to secretion systems. Secretion systems are specialized structures that bacteria use to transport molecules, such as toxins or enzymes, from inside their cells to the outside or into other cells; they often resemble complex, needle-like appendages attached to the bacterial cell wall. *Salmonella* have specific genes encoding Type III Secretion Systems (T3SS) and Type VI Secretion Systems (T6SS), which facilitate the injection of bacterial effector proteins into host cells and virulent competitors (Gal-Mor et al.). An effector is a protein produced by bacteria that can manipulate host cell functions to benefit the bacteria, aiding in infection and survival. T6SS, a needle-like apparatus, evolved for interbacterial competition and predation (Miyata et al.). Its effectors manipulate host cell functions or kill competing bacteria, aiding in bacterial survival and pathogenicity (Hernandez et al.). These effectors can disrupt immune responses, alter cell signaling pathways, and induce cell death, thereby enhancing bacterial infection and colonization. This paper will investigate the different secretion systems in *Salmonella*, focusing on T6SS, its effectors, and their roles in pathogenicity and interbacterial competition. Each subsection will cover the structure and mechanism of T6SS, the specific functions of its effectors, and the genetic regulation of these systems, aiming to provide a comprehensive understanding of how *Salmonella* uses T6SS to interact with its environment and hosts.

### **T6SS in *Salmonella* and its effector and genetics**

The Type VI Secretion System (T6SS) is a complex molecular machine used by *Salmonella* and other Gram-negative bacteria to deliver effector proteins into adjacent cells or the extracellular space surrounding the cells (Monjarás Feria and Valvano). Effectors are proteins that can alter the host cell's processes to benefit the bacteria, such as by causing cell damage or disrupting immune responses (Hernandez et al.). This benefits pathogenic bacteria because it can lead to the release of nutrients and molecules that the bacteria can utilize for growth and replication. For example, Hernandez *et al.* demonstrated that certain pathogenic bacteria exploit host cell damage to enhance their virulence and survival. This shows how bacteria can manipulate host cellular mechanisms to their advantage, promoting infection and disease

progression. In this system, effectors act as molecular tools that help bacteria establish infections and compete with other microorganisms (Monjarás Feria and Valvano).

The T6SS operates through a precise mechanism involving a contractile needle structure that punctures neighboring cells or the extracellular environment (Hernandez et al.). This sophisticated process is mediated by various proteins that regulate the assembly, functioning, and disassembly of the secretion apparatus. Despite significant research, the exact mechanisms of effector delivery and its regulation remain uncertain (Jurėnas and Journet). Hypotheses suggest that effectors can be directed either into target cells or released into the extracellular space, depending on the effector's nature and the bacterial strategy (Song et al.). The size of the effectors compared to the size of the secretion system poses significant challenges for precise delivery and understanding the exact regulatory mechanisms involved (Mak and Thurston). Research is ongoing to understand how environmental signals and specific factors influence these delivery processes and the efficiency of the T6SS (Amaya et al.; Monjarás Feria and Valvano). This highlights the importance of the T6SS in bacterial pathogenicity and the challenges in understanding its precise mechanisms, showing the complexity and significance of studying bacterial infection strategies for developing therapeutic interventions.

*Salmonella* species produce various effectors that perform different functions (Blondel et al.). The diversity of these effectors varies among *Salmonella* species, reflecting their adaptation to distinct host environments (Amaya et al.). Over 40 effector proteins are currently known, with many more predicted by bioinformatic analyses, though they remain to be experimentally confirmed (Pillay et al.). Studies show that *Salmonella* effectors interact with host cell mechanisms to enhance bacterial survival and replication (Schadich et al.). For example, the effector protein SopE helps *Salmonella enterica* serovar Typhimurium to invade host cells by inducing actin cytoskeleton rearrangements, facilitating bacterial uptake and replication within host cells (Röder and Hensel). This mechanism helps bacteria evade the host's immune response. Additionally, there's been an increased effort to investigate the genetics and functionality of *Salmonella* effectors, given their importance in disease and inducing damage to the host. For example, research demonstrated that specific effectors can manipulate host cell processes to support bacterial virulence (Blondel et al.). A recent study demonstrated that specific *Salmonella* effectors manipulate host cell processes by modulating the host's immune response, disrupting cellular signaling pathways, and altering host cell cytoskeleton dynamics. These manipulations support bacterial virulence by creating a more favorable environment for *Salmonella* survival and replication within the host. Understanding how effectors modify host cell processes is crucial for developing targeted therapies and preventive measures against *Salmonella* infections.

Regarding T6SS genetics, the system's assembly and effector delivery are controlled by genes encoding various components of the secretion apparatus (Schadich et al.). The *tssA* gene encodes a protein that forms part of the inner membrane complex and is crucial for anchoring the T6SS to the bacterial membrane (Russell et al.). The *tssB* and *tssC* genes encode proteins that assemble into a complex forming the baseplate and sheath of the T6SS. The *hcp* gene encodes the hemolysin coregulated protein, which is part of the T6SS tube that translocates effectors. The

*vgrG* gene produces a protein that aids in the formation of the needle complex, which is essential for delivering effectors to target cells. These genes are regulated in response to environmental factors and competitive interactions with other bacteria, ensuring that the T6SS is active when needed for infection or competition, thereby optimizing the bacterium's survival and pathogenicity. Understanding these genetic mechanisms provides insights into *Salmonella*'s adaptation to host environments and potential strategies for controlling bacterial pathogenicity. However, further research is needed to fully unravel the complexities of T6SS and its genetic regulation. In summary, *Salmonella*'s T6SS is essential for delivering effectors that support infection and competition. Clarifying T6SS function and genetics is crucial for developing effective treatments. These insights enhance our understanding of bacterial pathogenesis and adaptation strategies.

### **How the T6SS effector interacts with human cells**

This section will explore how effectors from the two major *Salmonella* serotypes—*Salmonella typhi* and *Salmonella enterica*—interact with the human body to induce disease. *Salmonella typhi* primarily infects humans through the ingestion of contaminated food or water, a common occurrence in areas with poor sanitation and limited access to clean drinking water (WHO). This bacterium targets cells in the gastrointestinal (GI) tract, specifically the epithelial cells lining the intestines and immune cells such as macrophages. It invades the intestinal epithelium and subsequently spreads to the lymphatic system and bloodstream (Nguyen et al.; Ramanujan). The infection process involves ingestion of the contaminated substance, followed by invasion of the intestinal epithelial cells. The bacteria then survive and replicate within macrophages to avoid immune detection, highlighting its inherent nature as an intracellular pathogen (Galan). Following this, the bacteria spread through the lymphatic system to other organs, such as the liver and spleen. This progression highlights the sophisticated tactics *Salmonella typhi* uses to evade the host's immune defenses and spread within the body, ultimately resulting in systemic infection and typhoid fever.

*S. typhi* employs various T6SS effectors to manipulate host cell processes, modulating the host's immune response and enhancing bacterial survival (Lien and Lai). While specific effectors and their exact mechanisms are still under investigation, they are known to interfere with immune signaling pathways and promote bacterial replication within host cells. For instance, some effectors can inhibit the production of pro-inflammatory cytokines, thereby dampening the immune response and facilitating bacterial persistence (Zhou et al.). Additionally, these effectors can manipulate host cell death pathways, preventing apoptosis of infected cells and allowing the bacteria to replicate within a protected intracellular environment (Lien and Lai). These interactions illustrate the complex strategies employed by *S. typhi* to evade the host immune response and sustain infection.

*Salmonella enterica* is primarily transmitted through the consumption of contaminated food, particularly undercooked meat, poultry, and eggs, or through contact with infected animals or their environment (Giannella; Ly and Casanova). This serotype targets cells in the GI tract,

similar to *S. typhi*, but also affects a broader range of host cells, including epithelial cells, and various immune cells within the intestines (Li; Ly and Casanova). The infection process involves ingestion of contaminated food, invasion of intestinal epithelial cells, and survival and replication within epithelial and immune cells. Next, the cells spread locally within the intestines, causing inflammation and diarrhea. *S. enterica* can be both an intracellular and extracellular pathogen, depending on the infection stage (Maciel et al.). This bacterium also uses T6SS to deliver effectors that disrupt various cellular processes, such as immune signaling and cytoskeletal dynamics, facilitating bacterial invasion and survival. This bacterium also uses the T6SS to deliver effectors that disrupt various cellular processes, such as immune signaling and cytoskeletal dynamics, facilitating bacterial invasion and survival. For instance, the T6SS effector proteins can modulate host cell cytoskeletal rearrangements and immune responses, which are crucial for bacterial evasion of host defenses and successful infection. The specific effectors and their mechanisms vary, contributing to the diverse clinical manifestations of *S. enterica* infections (Blondel et al.). These multifaceted strategies enable *S. enterica* to effectively colonize and persist within a wide range of hosts, leading to varied clinical outcomes.

There are similarities and differences between *S. typhi* and *S. enterica* infections. Both serotypes are foodborne pathogens; however, *S. typhi* is more associated with waterborne transmission in areas with poor sanitation, while *S. enterica* is commonly linked to contaminated food sources (Giannella; Ly and Casanova; Saleh et al.). Both target cells in the GI tract, but *S. typhi* is more likely to invade immune cells and spread systemically, while *S. enterica* typically causes localized intestinal infections (Li; Ly and Casanova; Nguyen et al.; Ramanujan). The effectors delivered by T6SS in both serotypes manipulate host cell processes, but the specific effectors and their targets can differ significantly between the serotypes, affecting their pathogenic mechanisms. For instance, *S. typhi*'s effectors often disrupt host cell signaling pathways and immune responses, aiding in systemic infection and prolonged survival. In contrast, *S. enterica*'s effectors generally focus on manipulating intestinal cell function and disrupting tight junctions to enhance bacterial invasion and persistence in the gut (Blondel et al.). These variations result in different clinical manifestations: typhoid fever, associated with prolonged fever and systemic symptoms, is caused by *S. typhi*, while gastroenteritis, which leads to acute gastrointestinal issues such as diarrhea and vomiting, is caused by *S. enterica*. (Gal-Mor et al.; Johnson et al.). The distinct pathogenic mechanisms and clinical presentations of *S. typhi* and *S. enterica*, including their use of the type VI secretion system (T6SS), reveal how differences in effector delivery influence their interactions with host cells and overall pathogenicity. Overall, deepening our understanding of T6SS in these serotypes is crucial for elucidating how their effectors contribute to pathogenicity and host-cell interactions. This knowledge is essential for developing targeted therapeutic strategies and improving disease management for infections caused by these *Salmonella* serotypes.

## How human cells respond to *Salmonella* effectors and pathogenesis

*Salmonella typhi* primarily infects human intestinal epithelial cells and can invade M cells, which are specialized epithelial cells in the Peyer's patches of the small intestine (Ménard et al.). Once inside, *S. typhi* can survive and replicate within macrophages, which are immune cells that typically ingest and destroy pathogens, making these cells ineffective in combating the infection (Weinstein et al.). The T6SS of *S. typhi* facilitates this process by injecting effector proteins into host cells, and promoting bacterial survival and replication within macrophages (Jepson and Clark). Additionally, *S. typhi* can interact with dendritic cells, another type of immune cell responsible for antigen presentation and initiation of the immune response, further complicating the host's ability to mount an effective defense (Weinstein et al.). These interactions highlight the complex mechanisms employed by *S. typhi* to establish and maintain infection within the human host (Galán). Understanding these interactions is crucial for developing strategies to combat the infection.

The body's response to *S. typhi* effectors, particularly those delivered by the T6SS, involve several immune mechanisms (Hersch et al.). One key response is the activation of macrophages and the production of pro-inflammatory cytokines, which are signaling molecules that mediate and regulate immunity, inflammation, and hematopoiesis (Bumann and Valdivia). These cytokines, such as TNF-alpha, IL-1, and IL-6, play crucial roles in orchestrating the immune response and recruiting other immune cells to the site of infection (Hersch et al.). However, *Salmonella typhi* has evolved sophisticated mechanisms to subvert these immune responses, allowing it to survive and replicate within the hostile environment of the host. One of the key strategies employed by *S. typhi* is the secretion of effector proteins through the Type VI Secretion System (T6SS). These effector proteins can manipulate host cell processes to the bacteria's advantage. For instance, they can inhibit the fusion of phagosomes with lysosomes within macrophages, a critical step in the host's attempt to degrade and eliminate pathogens. By preventing this fusion, *S. typhi* effectively avoids being destroyed within the macrophages, enabling it to persist within the host and contribute to the chronic nature of typhoid fever. This manipulation of the host immune response highlights the complex interplay between *S. typhi* and the host, underscoring the challenges in developing effective treatments. This allows *S. typhi* to persist and replicate within host cells, leading to systemic infection and dissemination of the pathogen throughout the body (Bumann and Valdivia). The body's immune response to *S. typhi* involves macrophage activation and cytokine production, but *S. typhi* can evade these defenses through T6SS-mediated inhibition of phagosome-lysosome fusion, allowing it to persist and spread within the host.

*S. enterica* serovars, including *S. typhimurium*, share similarities with *S. typhi* in their ability to infect various host cells, particularly intestinal epithelial cells and macrophages (Swart and Hensel; Ye et al.). Like *S. typhi*, *S. enterica* utilizes the T6SS to deliver effector proteins that manipulate host cell functions, allowing the bacteria to evade immune destruction and establish infection (Mistry et al.). However, unlike *S. typhi*, which predominantly causes systemic infections, *S. enterica* often targets a broader range of immune cells, including dendritic cells and

neutrophils (Ye et al.). This interaction with additional immune cells can lead to altered immune responses, further aiding *S. enterica*'s evasion of the host immune system (Swart and Hensel). This highlights the broader immune disruption capabilities of *S. enterica* compared to *S. typhi*, though both pathogens ultimately promote their survival by subverting host defenses.

The host immune response to *S. enterica* involves the recognition of bacterial components, such as lipopolysaccharides (LPS) and flagella, by pattern recognition receptors (PRRs) on immune cells (Haraga et al.). LPS, a major component of the outer membrane of *S. enterica*, is detected by Toll-like receptor 4 (TLR4), leading to the activation of inflammatory pathways, while flagella are recognized by Toll-like receptor 5 (TLR5), triggering immune responses aimed at neutralizing the bacterial threat. This recognition triggers the activation of signaling pathways that lead to the production of pro-inflammatory cytokines and the recruitment of immune cells to the site of infection (Hurley et al.). However, *S. enterica* has developed strategies to evade these responses, such as the secretion of effector proteins through its T6SS, which can manipulate host cell signaling and immune responses (Ye et al.). For instance, these effectors can inhibit the production of reactive oxygen species (ROS) and nitric oxide (NO), which are crucial for the killing of intracellular pathogens (Haraga et al.). Despite these evasion tactics, the host immune system can mount a robust response involving both innate and adaptive immunity to control and clear the infection (Hurley et al.). This is achieved through the activation of other immune pathways, such as the production of interferon-gamma by T cells, which enhances the bactericidal activity of macrophages. Additionally, the adaptive immune system generates specific antibodies that neutralize bacterial toxins and promote the clearance of the pathogen. These immune responses, although sometimes delayed or suppressed by the pathogen, can eventually lead to the elimination of *S. enterica* from the host. The host immune response to *S. enterica* involves recognizing bacterial components like LPS and flagella through PRRs, which activates inflammatory pathways, but the pathogen's T6SS-mediated effector proteins can suppress critical immune functions like ROS and NO production, though the host can still counteract these evasion strategies through robust innate and adaptive immune responses.

The body's defense against Salmonella infection involves a coordinated response from both the innate and adaptive immune systems (Mistry et al.). Innate immune cells, such as macrophages, dendritic cells, and neutrophils, play a critical role in the initial detection and containment of the bacteria through phagocytosis and the production of antimicrobial molecules (Kurtz et al.). Although *S. typhi* has evolved to subvert these immune cells by inhibiting phagosome-lysosome fusion within macrophages, the immune system can still mount an effective response through the activation of alternative immune pathways. For instance, the secretion of pro-inflammatory cytokines and activation of adaptive immune cells, such as T lymphocytes, can help to control and eventually eliminate the infection, despite the pathogen's efforts to evade immune destruction. This demonstrates the immune system's ability to adapt and overcome bacterial evasion tactics, ensuring effective pathogen clearance. Adaptive immunity is characterized by the activation of T-cells and B-cells, which provide long-term protection

through the production of specific antibodies and the generation of memory cells that can quickly respond to subsequent infections (Mistry et al.).

The host's ability to control and eliminate *Salmonella* relies on the effective collaboration between these immune components, despite the pathogen's sophisticated evasion strategies. For instance, CD4+ helper T-cells enhance the immune response by secreting cytokines that boost the activity of macrophages, thereby improving their ability to kill intracellular bacteria like *S. typhi*. In addition, CD8+ cytotoxic T-cells directly target and destroy infected cells, which helps limit bacterial replication. B-cells also play a crucial role by producing specific antibodies against *Salmonella* antigens. These antibodies neutralize the pathogen and mark it for destruction through processes like opsonization. Furthermore, memory T-cells and B-cells are generated during the initial infection and persist in the body long-term. Upon re-exposure to *Salmonella*, these memory cells enable a rapid and robust immune response, significantly reducing the severity of reinfection and aiding in the swift clearance of the pathogen. These examples highlight how adaptive immunity effectively combats *Salmonella* infections, even when the pathogen employs advanced mechanisms to evade immune detection. This section highlights how *Salmonella* uses the Type VI Secretion System (T6SS) to deliver effector proteins that subvert host immune responses, enabling the bacteria to invade and survive within various host cells. Despite these evasion strategies, the host immune system employs both innate and adaptive responses, such as pro-inflammatory cytokine production and antibody generation, to control and clear the infection.

## **Conclusion**

The Type VI Secretion System (T6SS) in *Salmonella* is a crucial virulence factor that significantly contributes to the bacterium's ability to cause disease (Sana et al.). T6SS enables *Salmonella* to inject effector proteins directly into host cells, manipulating their functions to the bacterium's advantage. This system is particularly important for *Salmonella's* survival within the host, as it allows the pathogen to evade immune detection and establish a persistent infection. In species like *S. typhi* and *S. enterica*, T6SS disrupts host immune responses, leading to severe and often systemic infections. The complexity and effectiveness of *Salmonella's* evasion strategies through T6SS pose considerable challenges for the host's immune defenses, highlighting the urgent need for novel therapeutic interventions. A deeper understanding of T6SS and its role in *Salmonella's* pathogenicity is essential for developing targeted treatments that could significantly reduce the global burden of *Salmonella* infections.

Leveraging knowledge of the T6SS provides several potential avenues for developing new treatments against *Salmonella* infections. One promising approach is the development of T6SS inhibitors that could prevent *Salmonella* from delivering effector proteins into host cells, thereby reducing the bacterium's ability to cause disease (Mulder et al.). This strategy is particularly important as it could help curb the rise of antibiotic resistance by offering an alternative to traditional antibiotics. Another potential treatment involves creating vaccines that prime the immune system to recognize and neutralize T6SS effectors, thereby enhancing the



host's ability to combat infection. Research also suggests that boosting specific aspects of the immune response, such as enhancing the production of reactive oxygen species (ROS), could counteract T6SS-mediated immune evasion and improve outcomes for infected individuals. These approaches are closely connected to the key points highlighted earlier, as they specifically address the mechanisms through which T6SS allows *Salmonella* to interfere with host functions and evade immune responses. By aiming to inhibit or neutralize the T6SS, these strategies tackle the challenges posed by *Salmonella's* complex pathogenic tactics. Ultimately, developing such targeted treatments could greatly decrease the global impact of *Salmonella* infections and offer a vital tool in combating antibiotic-resistant strains.

Despite these advances, there remain significant gaps in our understanding of the T6SS that researchers continue to explore. The full range of T6SS effectors and their specific targets within host cells are still being explored, which is crucial for developing comprehensive therapeutic strategies. Recent research highlights the T6SS's role in bacterial adaptation to stress conditions, such as metal ion uptake and resistance to oxidative stress, as well as its involvement in temperature and pH changes (Yu et al.). This system not only aids in interbacterial competition and environmental adaptation, but also in mediating stress responses. Furthermore, understanding how T6SS influences microbial community dynamics in the human gut, including its role in signal transduction among bacteria, is essential. This knowledge could impact both the progression of infections and the effectiveness of treatments. Ongoing research into these aspects of T6SS will be crucial for fully elucidating its role in *Salmonella* pathogenicity and developing innovative strategies to combat infections.

## Works Cited

- Amaya, Fernando A., et al. "Identification of Type VI Secretion Systems Effector Proteins That Contribute to Interbacterial Competition in Salmonella Dublin." *Frontiers in Microbiology*, vol. 13, 2022, p. 811932. *PubMed*, <https://doi.org/10.3389/fmicb.2022.811932>.
- Blondel, Carlos J., et al. "Identification and Distribution of New Candidate T6SS Effectors Encoded in Salmonella Pathogenicity Island 6." *Frontiers in Microbiology*, vol. 14, 2023, p. 1252344. *PubMed*, <https://doi.org/10.3389/fmicb.2023.1252344>.
- Bumann, Dirk, and Raphael H. Valdivia. "Identification of Host-Induced Pathogen Genes by Differential Fluorescence Induction Reporter Systems." *Nature Protocols*, vol. 2, no. 4, Apr. 2007, pp. 770–77. *www.nature.com*, <https://doi.org/10.1038/nprot.2007.78>.
- CDC. "Salmonella Homepage | CDC." *Center for Disease Control and Prevention*, 18 June 2024, <https://www.cdc.gov/salmonella/index.html>.
- Galan, Jorge. "Molecular Mechanisms of Salmonella Typhi Pathogenicity." *USDA National Agricultural Library*, 2011, <https://www.nal.usda.gov/research-tools/food-safety-research-projects/molecular-mechanisms-salmonella-typhi-pathogenicity>.
- Galán, Jorge E. "Interactions of Salmonella with Host Cells: Encounters of the Closest Kind." *Proceedings of the National Academy of Sciences*, vol. 95, no. 24, Nov. 1998, pp. 14006–08. *pnas.org (Atypon)*, <https://doi.org/10.1073/pnas.95.24.14006>.
- Gal-Mor, Ohad, et al. "Same Species, Different Diseases: How and Why Typhoidal and Non-Typhoidal Salmonella Enterica Serovars Differ." *Frontiers in Microbiology*, vol. 5, Aug. 2014, p. 391. *PubMed Central*, <https://doi.org/10.3389/fmicb.2014.00391>.
- Giannella, Ralph A. "Salmonella." *Medical Microbiology*, edited by Samuel Baron, 4th ed., University of Texas Medical Branch at Galveston, 1996. *PubMed*, <http://www.ncbi.nlm.nih.gov/books/NBK8435/>.
- Haraga, Andrea, et al. "Salmonellae Interplay with Host Cells." *Nature Reviews Microbiology*, vol. 6, no. 1, Jan. 2008, pp. 53–66. *www.nature.com*, <https://doi.org/10.1038/nrmicro1788>.
- Hernandez, Ruth E., et al. "Type VI Secretion System Effector Proteins: Effective Weapons for Bacterial Competitiveness." *Cellular Microbiology*, vol. 22, no. 9, 2020, p. e13241. *Wiley Online Library*, <https://doi.org/10.1111/cmi.13241>.
- Hersch, Steven J., et al. "Defending against the Type Six Secretion System: Beyond Immunity Genes." *Cell Reports*, vol. 33, no. 2, Oct. 2020, p. 108259. *DOI.org (Crossref)*, <https://doi.org/10.1016/j.celrep.2020.108259>.
- Hurley, Daniel, et al. "Salmonella–Host Interactions – Modulation of the Host Innate Immune System." *Frontiers in Immunology*, vol. 5, Oct. 2014, p. 481. *PubMed Central*, <https://doi.org/10.3389/fimmu.2014.00481>.
- Jepson, Mark A., and M. Ann Clark. "The Role of M Cells in Salmonella Infection." *Microbes and Infection*, vol. 3, no. 14, Nov. 2001, pp. 1183–90. *ScienceDirect*,

- [https://doi.org/10.1016/S1286-4579\(01\)01478-2](https://doi.org/10.1016/S1286-4579(01)01478-2).
- Johnson, Rebecca, et al. "Comparison of Salmonella Enterica Serovars Typhi and Typhimurium Reveals Typhoidal Serovar-Specific Responses to Bile." *Infection and Immunity*, vol. 86, no. 3, Feb. 2018, p. 10.1128/iai.00490-17. *journals.asm.org (Atypon)*, <https://doi.org/10.1128/iai.00490-17>.
- Jurėnas, Dukas, and Laure Journet. "Activity, Delivery, and Diversity of Type VI Secretion Effectors." *Molecular Microbiology*, vol. 115, no. 3, 2021, pp. 383–94. *Wiley Online Library*, <https://doi.org/10.1111/mmi.14648>.
- Kurtz, Jonathan R., et al. "Salmonella Infection: Interplay between the Bacteria and Host Immune System." *Immunology Letters*, vol. 190, Oct. 2017, pp. 42–50. *PubMed Central*, <https://doi.org/10.1016/j.imlet.2017.07.006>.
- Lamichhane, Bibek, et al. "Salmonellosis: An Overview of Epidemiology, Pathogenesis, and Innovative Approaches to Mitigate the Antimicrobial Resistant Infections." *Antibiotics*, vol. 13, no. 1, 1, Jan. 2024, p. 76. *www.mdpi.com*, <https://doi.org/10.3390/antibiotics13010076>.
- Li, Qiao. "Mechanisms for the Invasion and Dissemination of Salmonella." *The Canadian Journal of Infectious Diseases & Medical Microbiology = Journal Canadien Des Maladies Infectieuses et de La Microbiologie Médicale*, vol. 2022, June 2022, p. 2655801. *PubMed Central*, <https://doi.org/10.1155/2022/2655801>.
- Lien, Yun-Wei, and Erh-Min Lai. "Type VI Secretion Effectors: Methodologies and Biology." *Frontiers in Cellular and Infection Microbiology*, vol. 7, 2017, p. 254. *PubMed*, <https://doi.org/10.3389/fcimb.2017.00254>.
- Ly, Kim Thien, and James E. Casanova. "Mechanisms of Salmonella Entry into Host Cells." *Cellular Microbiology*, vol. 9, no. 9, Sept. 2007, pp. 2103–11. *PubMed*, <https://doi.org/10.1111/j.1462-5822.2007.00992.x>.
- Maciel, Bianca Mendes, et al. "Salmonella Enterica: Latency." *Current Topics in Salmonella and Salmonellosis*, IntechOpen, 2017. *www.intechopen.com*, <https://doi.org/10.5772/67173>.
- Magana-Arachchi, D. N., and R. P. Wanigatunge. "Chapter 2 - Ubiquitous Waterborne Pathogens." *Waterborne Pathogens*, edited by Majeti Narasimha Vara Prasad and Anna Grobelak, Butterworth-Heinemann, 2020, pp. 15–42. *ScienceDirect*, <https://doi.org/10.1016/B978-0-12-818783-8.00002-5>.
- Mak, Hazel, and Teresa L. M. Thurston. "Interesting Biochemistries in the Structure and Function of Bacterial Effectors." *Frontiers in Cellular and Infection Microbiology*, vol. 11, Feb. 2021. *Frontiers*, <https://doi.org/10.3389/fcimb.2021.608860>.
- Ménard, Sandrine, et al. "Cross-Talk Between the Intestinal Epithelium and Salmonella Typhimurium." *Frontiers in Microbiology*, vol. 13, June 2022. *Frontiers*, <https://doi.org/10.3389/fmicb.2022.906238>.
- Mistry, Jayna J., et al. "ROS-Mediated PI3K Activation Drives Mitochondrial Transfer from Stromal Cells to Hematopoietic Stem Cells in Response to Infection." *Proceedings of the National Academy of Sciences*, vol. 116, no. 49, Dec. 2019, pp. 24610–19. *pnas.org*

- (*Atypon*), <https://doi.org/10.1073/pnas.1913278116>.
- Miyata, Sarah T., et al. “Type VI Secretion System Regulation as a Consequence of Evolutionary Pressure.” *Journal of Medical Microbiology*, vol. 62, no. 5, 2013, pp. 663–76. *Microbiology Society Journals*, <https://doi.org/10.1099/jmm.0.053983-0>.
- Monjarás Feria, Julia, and Miguel A. Valvano. “An Overview of Anti-Eukaryotic T6SS Effectors.” *Frontiers in Cellular and Infection Microbiology*, vol. 10, Oct. 2020. *Frontiers*, <https://doi.org/10.3389/fcimb.2020.584751>.
- Mulder, David T., et al. “Type VI Secretion System-Associated Gene Clusters Contribute to Pathogenesis of *Salmonella* Enterica Serovar Typhimurium.” *Infection and Immunity*, vol. 80, no. 6, May 2012, pp. 1996–2007. *journals.asm.org (Atypon)*, <https://doi.org/10.1128/iai.06205-11>.
- Nguyen, Tri, et al. “The Role of 9-O-Acetylated Glycan Receptor Moieties in the Typhoid Toxin Binding and Intoxication.” *PLOS Pathogens*, vol. 16, no. 2, Feb. 2020, p. e1008336. *PLoS Journals*, <https://doi.org/10.1371/journal.ppat.1008336>.
- Pillay, Timesh D., et al. “Speaking the Host Language: How *Salmonella* Effector Proteins Manipulate the Host.” *Microbiology*, vol. 169, no. 6, June 2023, p. 001342. *PubMed Central*, <https://doi.org/10.1099/mic.0.001342>.
- Ramanujan, Krishna. “Study Finds Key Mechanism for How Typhoid Bacteria Infects.” *The Cornell Chronicle*, 21 Feb. 2020, <https://news.cornell.edu/stories/2020/02/study-finds-key-mechanism-how-typhoid-bacteria-infects>.
- Robertson, Ruairi. “What Is a Strain? Strain-Level Identification in Microbiome Analyses.” *Microbiome Insights*, 22 Nov. 2023, <https://blog.microbiomeinsights.com/what-is-a-strain-strain-level-identification-in-microbiome-analyses>.
- Röder, Jennifer, and Michael Hensel. “Presence of SopE and Mode of Infection Result in Increased *Salmonella*-Containing Vacuole Damage and Cytosolic Release during Host Cell Infection By.” *Cellular Microbiology*, vol. 22, no. 5, 2020, p. e13155. *Wiley Online Library*, <https://doi.org/10.1111/cmi.13155>.
- Russell, Alistair B., et al. “Type VI Secretion System Effectors: Poisons with a Purpose.” *Nature Reviews. Microbiology*, vol. 12, no. 2, Feb. 2014, pp. 137–48. *PubMed*, <https://doi.org/10.1038/nrmicro3185>.
- Saleh, Sara, et al. “*Salmonella* Typhi, Paratyphi A, Enteritidis and Typhimurium Core Proteomes Reveal Differentially Expressed Proteins Linked to the Cell Surface and Pathogenicity.” *PLoS Neglected Tropical Diseases*, vol. 13, no. 5, May 2019, p. e0007416. *PubMed Central*, <https://doi.org/10.1371/journal.pntd.0007416>.
- Sana, Thibault G., et al. “*Salmonella* Typhimurium Utilizes a T6SS-Mediated Antibacterial Weapon to Establish in the Host Gut.” *Proceedings of the National Academy of Sciences*, vol. 113, no. 34, Aug. 2016, pp. E5044–51. *pnas.org (Atypon)*, <https://doi.org/10.1073/pnas.1608858113>.

- Schadich, Ermin, et al. "Role of Salmonella Typhi Vi Antigen and Secretory Systems on Immune Response." *Current Pharmaceutical Design*, vol. 22, no. 41, 2016, pp. 6251–60. *PubMed*, <https://doi.org/10.2174/1381612822666160829142308>.
- Song, Li, et al. "Contact-Independent Killing Mediated by a T6SS Effector with Intrinsic Cell-Entry Properties." *Nature Communications*, vol. 12, no. 1, Jan. 2021, p. 423. [www.nature.com](http://www.nature.com), <https://doi.org/10.1038/s41467-020-20726-8>.
- Swart, A. Leoni, and Michael Hensel. "Interactions of Salmonella Enterica with Dendritic Cells." *Virulence*, vol. 3, no. 7, Nov. 2012, pp. 660–67. *PubMed Central*, <https://doi.org/10.4161/viru.22761>.
- Weinstein, Debra L., et al. "Differential Early Interactions between Salmonella Enterica Serovar Typhi and Two Other Pathogenic Salmonella Serovars with Intestinal Epithelial Cells." *Infection and Immunity*, vol. 66, no. 5, May 1998, pp. 2310–18.
- WHO. "Typhoid." *World Health Organization*, 30 Mar. 2023, <https://www.who.int/news-room/fact-sheets/detail/typhoid>.
- Ye, Chenglin, et al. "Salmonella Enterica Serovar Typhimurium Interacts with CD209 Receptors To Promote Host Dissemination and Infection." *Infection and Immunity*, vol. 87, no. 8, July 2019, pp. e00100-19. *PubMed Central*, <https://doi.org/10.1128/IAI.00100-19>.
- Yu, Kai-Wei, et al. "T6SS Mediated Stress Responses for Bacterial Environmental Survival and Host Adaptation." *International Journal of Molecular Sciences*, vol. 22, no. 2, 2, Jan. 2021, p. 478. [www.mdpi.com](http://www.mdpi.com), <https://doi.org/10.3390/ijms22020478>.
- Zhou, Guodong, et al. "Manipulation of Host Immune Defenses by Effector Proteins Delivered from Multiple Secretion Systems of Salmonella and Its Application in Vaccine Research." *Frontiers in Immunology*, vol. 14, Apr. 2023. *Frontiers*, <https://doi.org/10.3389/fimmu.2023.1152017>.

## **A Review on Colorectal Cancer in Young Vs Old Patients, Including Related Symptoms, Diet, and Gut Microbiomes By Anirudh R. Guddeti, David H. Parkinson**

### **Summary**

Colorectal cancer (CRC) is a major global health concern, with significant differences in incidence, clinical characteristics, and outcomes between early-onset (EOCRC) and late-onset CRC (LOCRC). This review explores the distinctions between EOCRC and LOCRC, focusing on clinical features, dietary patterns, and gut microbiome composition. EOCRC is often more aggressive and diagnosed at a later stage compared to LOCRC, likely due to the lack of routine screening in younger populations. The increasing incidence of EOCRC in high-income countries highlights the need for revised screening guidelines. Dietary factors, particularly the consumption of processed meats and a Western diet, are associated with increased CRC risk, while diets rich in fiber and calcium are protective. The gut microbiome plays a critical role in CRC development, with pathogenic bacteria such as *Fusobacterium nucleatum* and *Escherichia coli* being more prevalent in CRC patients, while beneficial species are diminished. This review underscores the importance of understanding the microbiome's role in CRC and the necessity for tailored prevention and treatment strategies. Further research is needed to explore the underlying mechanisms and develop microbiome-based diagnostics and therapies.

### **Introduction**

Colorectal Cancer is the third most common cancer in the world and second highest cause of cancer related mortality (1, 2). In 2020, colorectal cancer represented 9.4% of cancer-related deaths worldwide (3). The global estimated incidence in colorectal cancer was 1.9 million cases in 2020 and is predicted to rise to 3.2 million in 2040 (4). The term colorectal cancer refers to cancer in the colon or rectum, caused by uncontrollable multiplication of the inner layer of glandular epithelial cells in the colon or rectum (4). Colorectal cancer (CRC) is associated with aging (5). Colorectal cancer can be classified based on age of diagnosis, with age < 50 considered early onset colorectal cancer (EOCRC), and age > 50 considered late onset colorectal cancer (LOCRC) (5). EOCRC is more likely to be metastatic, and more likely to be diagnosed at a more advanced stage of cancer (4). The incidence of colorectal cancer also varies by location, with higher incidence in higher income countries (5). Since 1994, the incidence and mortality of EOCRC specifically has increased by 2% each year, especially in high income countries (5). The surge in CRC incidence in young adults is particularly alarming as the overall CRC frequency has been decreasing in high income countries (5). In the U.S, the incidence of EOCRC is predicted to double by 2030 (5). This has pushed the U.S. Preventative Task Force (USPTF) to reduce the screening age to 45 in 2021 (2). While some people are genetically predisposed to have an increased risk of CRC, the increase in cases has largely been attributed to an unhealthy diet and lifestyle (4,5). LOCRC is more commonly diagnosed at an early stage due to patients undergoing screening regularly, whereas patients under the age of 50 are not screened unless there is a family history of CRC (6).

### **Diet and CRC**

CRC is influenced by many factors, both genetic and environmental (7). There is evidence that suggests increased consumption of red meat and alcohol are attributed to increased CRC risk, while dietary fiber, calcium, and yogurt are attributed to a lower risk (7). BMI and physical activity also play a crucial role in CRC incidence and mortality (7, 8). Healthy lifestyle advice is given to most CRC patients post diagnosis; however, studies have shown that this guidance is more of a preventative measure and may not be useful post diagnosis (7). CRC survivors are given advice to help prevent recurrence, but the same advice can be used to prevent getting cancer in the first place (7). A western diet, generally characterized by high consumption of red or processed meat, refined grains, fats, and salts is associated with CRC incidence (7, 8). Studies have shown that a western diet caused a threefold increased risk of development, recurrence, or mortality from CRC (7).

Inflammation also plays a significant role in CRC development. It has been shown that a diet consisting of fruits, vegetables, wholegrains, fish, nuts, and low-fat dairy products, while limiting red meat and full-fat dairy can decrease inflammation in the gut (7, 8). Calcium, which is important for human health, is known to inhibit colonic KRAS mutations which cause CRC, but it is recommended by doctors that calcium intake should come from diet rather than supplements (7). Key sources of calcium in the diet include dairy products and leafy greens such as kale (7,8). Nuts are also full of nutrients, and antioxidants as well as anti-carcinogenic and anti-inflammatory properties (7,8). Tree nuts have been associated with both disease free and recurrence free survival, but peanuts had no effect (7). Coffee was shown to significantly affect CRC specific mortality, with patients who drank 4 or more cups per day having a 52% decreased risk of CRC death.

Dietary fiber, found in wholegrains, fruits, vegetables, and legumes, has been associated with a decreased risk of CRC (7, 8), and a healthier gut microbiome. Ultra processed foods, which are rich in sodium, sugar, and saturated fats, increase risk of CRC (8) and can harm helpful gut bacteria. For example, a high fat diet can lead to increased pathogenic *Bacteroides* species (9), whereas a plant rich diet leads to increased *Prevotella* species (9), which aid in the breakdown of fibers and lead to a healthy gut.

### **The gut microbiome and CRC**

The human gut hosts tens of thousands of species and approximately 100 trillion microorganisms, which make up the gut microbiome (9). The organisms that compose the gut microbiome may number more than the total cells in the human body (11). The gut microbiome plays various important roles in the human body, including metabolism, and acting as a barrier against pathogens (10). Because of this, microbiome dysbiosis can cause many problems, and large studies have shown that gut microbiome dysbiosis is linked to colorectal carcinogenesis (10).

Alterations or dysbiosis in the gut microbiome can be caused by dietary patterns or shifts in environmental factors (9). The main tumor promoting activities of the microbiota are bacterial

toxins, metabolic products, and immune modulation or reaction (10). Several studies have also shown that the microbiome is associated with CRC carcinogenesis via inflammatory pathways. Many species have already been identified to contribute to inflammation and carcinogenesis (11). *Escherichia coli*, more commonly known as *E. coli*, has been found in many CRC patient gut microbiomes. There is evidence that these bacteria are colonizing CRC lesions and nearby tissue (10). *E. coli* is also known to affect the immune system and help tumor cells survive by inhibiting macrophages (10). Emerging evidence also suggests that the microbiome may impact the effectiveness of cancer treatment (10). One study demonstrated that a fecal transplant from patients with CRC to healthy mice led to increased tumorigenesis under the influence of carcinogens (11), which suggests that there is a relationship between microbial species and development of CRC.

### **Clinical and pathological features of early and late onset CRC**

Early onset and late onset colorectal cancer tend to have different clinical characteristics (12). Early-onset CRCs more frequently display adverse histopathological features, all of which are suggestive of an unfavorable tumor biology and are associated with worse oncological outcomes (12). Younger patients with CRC are more likely to have synchronous and metachronous lesions and typically display more advanced disease stage at presentation compared with their older counterparts (12). Higher lymph node involvement, worse outcomes, and advanced stage cancers are also associated with EOCRC (12).

Symptoms of colorectal cancer can include abdominal pain, rectal bleeding, anemia, and diarrhea. Having 3 or more of these symptoms is an extreme red flag for EOCRC (13). However, there are approximately 34 million patients with abdominal pain, 28 million with diarrhea, and almost 8 million with anemia, that are under the age of 50 (13). Thus, it would take on average about 60,000 colonoscopies to diagnose just 2 patients with CRC, which is neither possible nor realistic (13).

Many previous studies have provided a comprehensive examination of the relationship between colorectal cancer (CRC) and various factors including microbiome composition, age, dietary habits, and genetic mutations. Here, the aim is to emphasize key findings and highlight common findings among these studies.

## **RESULTS**

*Of the 12 studies (14-25), 5 were case studies (14,16,19,21,23) and 5 were cohort studies (15,20,22,24,25), 1 was a meta-analysis (17), and 1 was a case control study (18).*

### **Microbiome in CRC**

Multiple studies linked microbial species with CRC. *Fusobacterium nucleatum* was implicated in CRC development in 75% of the studies, including detection inside CRC tumors (14, 15, 16, 17, 21, 22, 23, 24). *Clostridium symbiosum*, *Peptostreptococcus stomatis*, *Parvimonas micra*, and *Hungatella hathewayi* were enriched in both early-onset CRC (EOCRC)



and late-onset CRC (LOCRC) patients, serving as marker taxa for CRC (15). In contrast, beneficial species like *Eubacterium rectale*, *Roseburia intestinalis*, and *Bifidobacterium adolescentis* were found to be less abundant in CRC patients (15, 25). The odds of CRC increased significantly with the presence of certain bacterial species, with one study reporting a 593% increase with the presence of *Fusobacterium*, 102% with the presence of *E. coli*, and 92% with the presence of colibactin (17). Additionally, the presence of *E. coli* seemed to be less associated with CRC in younger patients compared to older patients (17).

### **Age and Microbiome Variation**

Age was shown to have a significant effect on the gut microbiome in CRC patients. Several studies reported weak correlations between age and microbial alpha diversity (15, 17, 24). Specific species, such as *Prevotella stercorea*, *Bifidobacterium dentium*, and *Prevotella bivia*, were found to have age-related variations (15). However, CRC appeared to override age-related changes in the microbiome, suggesting that the disease exerts a stronger influence on microbial composition than age alone (15). *Fusobacterium nucleatum* and *E. coli* are prevalent in CRC patients, with the presence of these bacteria linked to an increased CRC risk (17, 23). *E. coli* was notably associated with older patients, whereas *Fusobacterium* was more commonly found across all ages in CRC cases (17, 23).

### **Alpha and Beta Diversity in CRC**

Alpha diversity refers to the number of microbial taxa within a specific sample, while beta diversity refers to differences in species composition between multiple samples. Increased microbial diversity is a common finding in CRC patients compared to healthy controls, with a shift in bacterial species (18, 24, 25). Studies found differences in microbial alpha diversity between CRC patients and healthy controls, with CRC patients showing increased levels of pathogenic bacteria and decreased levels of beneficial bacteria (18, 22, 25). Elevated alpha diversity was also observed in patients with KRAS-mutant CRC compared to those with wild-type CRC (21). CRC progression is often associated with a higher prevalence of pathogenic bacteria such as *Fusobacterium* and *Enterococcus* and a reduction in beneficial species like *Roseburia intestinalis* and *Bifidobacterium adolescentis* (25). There is also a significant difference in beta diversity between CRC patients and healthy controls across all disease stages, highlighting variations in microbial communities in relation to disease progression (24, 25).

### **Diet and CRC Microbiota**

Dietary factors consistently emerge as influential in CRC risk and microbiota composition. The consumption of processed meats and high-calorie diets is associated with increased CRC risk and alterations in gut microbiota (16, 20). The American diet's shift towards increased caloric sweeteners and ultra-processed foods is linked to changes in gut microbiota composition, promoting CRC through mechanisms like inflammation and carcinogenic metabolite production (16). Pro-inflammatory bacteria such as *Fusobacterium* and decreased

populations of beneficial bacteria like *Faecalibacterium prausnitzii* were observed in CRC patients who consumed more processed foods and less dietary fiber (20).

## **DISCUSSION**

This study shows the complex relationship between the gut microbiome and colorectal cancer. Understanding specific bacterial species' roles in CRC can lead to new biomarkers for early detection and potential therapeutic targets. The findings underscore the significance of the microbiome in cancer development, suggesting that alterations in gut bacteria may directly contribute to CRC pathogenesis. Along with this, knowing the differences in CRC between young and older patients is crucial for several reasons. Firstly, the rising incidence of EOCRC, especially in high-income countries, poses a significant public health challenge. The traditionally accepted view has been that CRC primarily affects older individuals, which has influenced screening guidelines and public health strategies. However, the increasing occurrence of CRC in younger populations creates a need for a reevaluation of these guidelines and a broader public health approach. Moreover, recognizing the unique clinical and pathological characteristics of EOCRC, such as its more aggressive nature and advanced stage at diagnosis, can lead to better treatment methods and potentially improved outcomes for younger patients.

In the past, CRC was commonly believed to affect mostly older patients, leading to screening programs targeting people over the age of 50. EOCRC was a rare occurrence. The assumption was that the risk factors, such as age-related genetic mutations and longer exposure to environmental factors, made older adults more susceptible. However, recent data challenges this notion, showing an increase in CRC cases among younger adults who may not have these traditional risk factors. This shift in understanding highlights the importance of considering a wider range of potential risk factors, including dietary patterns and gut microbiota composition. In addition, the gut microbiome's role in CRC was not fully understood, with limited recognition of specific bacteria like *Fusobacterium nucleatum*. This study confirms and expands upon the idea that the microbiome can influence cancer risk and progression, showing that microbial diversity and specific bacterial profiles are closely associated with CRC, regardless of age. It challenges previous assumptions that age alone significantly alters the gut microbiome, instead suggesting that CRC can dominate these changes.

EOCRC and LOCRC exhibit distinct clinical and pathological features. EOCRC is often diagnosed at a more advanced stage and is more likely to be metastatic compared to LOCRC. This can be attributed to the lack of routine screening in younger populations, which delays diagnosis until the disease is more advanced. The U.S. Preventative Task Force's decision to lower the screening age to 45 reflects the rising incidence of EOCRC and the need for earlier detection. Additionally, EOCRC tends to present more frequently in the distal colon or rectum, often with constricting lesions rather than the more common ascending tumors seen in older patients.

Diet and lifestyle are significant factors influencing CRC risk across all age groups. The increasing prevalence of a Western diet, characterized by high consumption of red and processed

meats, refined grains, and high-fat foods, has been linked to higher CRC incidence. This diet is associated with increased inflammation and carcinogenesis, potentially due to its impact on the gut microbiome and the production of harmful metabolites. Conversely, diets rich in fiber, fruits, vegetables, and calcium are associated with a reduced risk of CRC. The beneficial effects of such diets include reduced inflammation, improved gut microbiome diversity, and protective effects against genetic mutations like KRAS mutations. The study also highlights that the unhealthy dietary patterns observed in younger populations, which may include higher consumption of fast food and processed foods, could be contributing to the rising incidence of EOCRC.

The gut microbiome plays a crucial role in CRC development, with significant differences observed between the microbiomes of CRC patients and healthy controls. The study found that pathogenic bacteria like *Fusobacterium nucleatum* and *Escherichia coli* are more prevalent in CRC patients, while beneficial species such as *Eubacterium rectale* and *Bifidobacterium adolescentis* are less abundant. These microbial shifts are associated with inflammation, immune modulation, and the production of carcinogenic metabolites. Age-related differences in the gut microbiome composition are evident, with certain species showing varying prevalence across age groups. However, CRC appears to have a more significant influence on the gut microbiome than age alone. This suggests that the presence of CRC can override typical age-related microbial changes, emphasizing the disease's impact on gut health.

### **Limitations**

The review was limited to studies published during and after 2023, potentially overlooking valuable insights from earlier research. The exclusion criteria, such as excluding animal studies and those not directly discussing CRC in relation to diet or the microbiome, might have limited the scope of findings. The studies reviewed often focused on associations rather than hard results, which limits the ability to draw definitive conclusions about the mechanisms involved. The analysis was limited by the heterogeneity of the included studies, varying methodologies, and the reliance on existing data, which may not capture all relevant microbial factors. This may lead to inconsistencies in findings and make it challenging to compare results directly. Some studies may have had limited sample sizes or lacked diversity in demographics, potentially affecting the generalizability of the findings.

### **For the Future**

Future studies should include interventional studies that test specific dietary or microbiome interventions could help determine their effectiveness in CRC prevention or treatment. Larger, more comprehensive studies can also confirm the specific bacterial species' roles in CRC and their mechanisms. In the future, researchers could control what each experimental group eats, to be able to look at dietary pattern effects on CRC without any other dietary factors in play. These studies should also include diverse populations in terms of age, ethnicity, and lifestyle which will help generalize findings across different demographic groups. Along with this, development of microbiome-based diagnostic tools and personalized treatment

strategies could improve CRC detection and management. Further research should focus on integrating microbiome analysis into routine clinical practice.

### **Materials and Methods**

For this review a PubMed search was conducted using the following keywords: "Colorectal Cancer", "CRC", "Gut", "Microbiota", "microbiome", "gut bacteria", "Age", "early", "onset", "late." The following exclusion criteria were applied: Studies before 2023, Studies not discussing colorectal cancer, Studies not discussing colorectal cancer in relation to diet or gut microbiome, and animal studies (Figure 1).

## Works Cited

1. Liu, Hongwei et al. "Comparative characteristics of early-onset vs. late-onset advanced colorectal cancer: a nationwide study in China." *BMC cancer* vol. 24,1 503. 20 Apr. 2024, doi:10.1186/s12885-024-12278-7
2. Jeri-Yabar, Antoine et al. "Survival Analysis of Metastatic Early-Onset Colorectal Cancer Compared to Metastatic Average-Onset Colorectal Cancer: A SEER Database Analysis." *Cancers* vol. 16,11 2004. 25 May. 2024, doi:10.3390/cancers16112004
3. Aeschbacher, Pauline et al. "Obesity and overweight are associated with worse survival in early-onset colorectal cancer." *Surgery* vol. 176,2 (2024): 295-302. doi:10.1016/j.surg.2024.03.037
4. Kwao-Zigah, Genevieve et al. "Microbiome Dysbiosis, Dietary Intake and Lifestyle-Associated Factors Involve in Epigenetic Modulations in Colorectal Cancer: A Narrative Review." *Cancer control : journal of the Moffitt Cancer Center* vol. 31 (2024): 10732748241263650. doi:10.1177/10732748241263650
5. Su, Jiao et al. "The global burden and trends analysis of early-onset colorectal cancer attributable to dietary risk factors in 204 countries and territories, 1990-2019: a secondary analysis for the global burden of disease study 2019." *Frontiers in nutrition* vol. 11 1384352. 31 May. 2024, doi:10.3389/fnut.2024.1384352
6. Lukman, Kiki et al. "The analysis between clinicopathological aspect of early-onset vs. late-onset colorectal cancer and mortality rate: a cross-sectional study." *Annals of medicine and surgery (2012)* vol. 85,6 2496-2501. 3 May. 2023, doi:10.1097/MS9.0000000000000757
7. Fretwell, Anna et al. "Still too little evidence: the role of diet in colorectal cancer survivorship - a systematic review." *Critical reviews in food science and nutrition*, 1-13. 11 Jun. 2024, doi:10.1080/10408398.2024.2360068
8. Oliveira, Manoela Lima et al. "A Perspective Review on Diet Quality, Excess Adiposity, and Chronic Psychosocial Stress and Implications for Early-Onset Colorectal Cancer." *The Journal of nutrition* vol. 154,4 (2024): 1069-1079. doi:10.1016/j.tjn.2024.03.002
9. Chawrylak, Katarzyna et al. "Gut Microbiota-Adversary or Ally? Its Role and Significance in Colorectal Cancer Pathogenesis, Progression, and Treatment." *Cancers* vol. 16,12 2236. 15 Jun. 2024, doi:10.3390/cancers16122236
10. Wang, Yan, and Hui Li. "Gut microbiota modulation: a tool for the management of colorectal cancer." *Journal of translational medicine* vol. 20,1 178. 21 Apr. 2022, doi:10.1186/s12967-022-03378-8
11. Dai, Rui et al. "The Impact of the Gut Microbiome, Environment, and Diet in Early-Onset Colorectal Cancer Development." *Cancers* vol. 16,3 676. 5 Feb. 2024, doi:10.3390/cancers16030676
12. REACCT Collaborative et al. "Characteristics of Early-Onset vs Late-Onset Colorectal Cancer: A Review." *JAMA surgery* vol. 156,9 (2021): 865-874. doi:10.1001/jamasurg.2021.2380

13. Yen, Timothy, and Swati G Patel. "Symptoms and early-onset colorectal cancer: red flags are common flags!" *Journal of the National Cancer Institute* vol. 115,8 (2023): 883-885. doi:10.1093/jnci/djad093
14. Tito, Raúl Y et al. "Microbiome confounders and quantitative profiling challenge predicted microbial targets in colorectal cancer development." *Nature medicine* vol. 30,5 (2024): 1339-1348. doi:10.1038/s41591-024-02963-2
15. Qin, Youwen et al. "Consistent signatures in the human gut microbiome of old- and young-onset colorectal cancer." *Nature communications* vol. 15,1 3396. 22 Apr. 2024, doi:10.1038/s41467-024-47523-x
16. Atchade, Axelle Mayode et al. "Unraveling the complexities of early-onset colorectal cancer: a perspective on dietary and microbial influences." *Frontiers in public health* vol. 12 1370108. 4 Apr. 2024, doi:10.3389/fpubh.2024.1370108
17. Kharofa, Jordan et al. "Metagenomic analysis of the fecal microbiome in colorectal cancer patients compared to healthy controls as a function of age." *Cancer medicine* vol. 12,3 (2023): 2945-2957. doi:10.1002/cam4.5197
18. Carson, Tiffany L et al. "A case-control study of the association between the gut microbiota and colorectal cancer: exploring the roles of diet, stress, and race." *Gut pathogens* vol. 16,1 13. 11 Mar. 2024, doi:10.1186/s13099-024-00608-w
19. Barot, Shimoli V et al. "Distinct intratumoral microbiome of young-onset and average-onset colorectal cancer." *EBioMedicine* vol. 100 (2024): 104980. doi:10.1016/j.ebiom.2024.104980
20. Xiao, Yi et al. "Association between sulfur microbial diet and the risk of colorectal cancer precursors in older adults." *Frontiers in nutrition* vol. 10 1167372. 14 Aug. 2023, doi:10.3389/fnut.2023.1167372
21. Huang, Zigui et al. "Identification of KRAS mutation-associated gut microbiota in colorectal cancer and construction of predictive machine learning model." *Microbiology spectrum* vol. 12,5 (2024): e0272023. doi:10.1128/spectrum.02720-23
22. Alhhazmi, Areej A et al. "Identification of Gut Microbiota Profile Associated with Colorectal Cancer in Saudi Population." *Cancers* vol. 15,20 5019. 17 Oct. 2023, doi:10.3390/cancers15205019
23. Joo, Jihoon E et al. "Intratumoral presence of the genotoxic gut bacteria pks+ *E. coli*, Enterotoxigenic *Bacteroides fragilis*, and *Fusobacterium nucleatum* and their association with clinicopathological and molecular features of colorectal cancer." *British journal of cancer* vol. 130,5 (2024): 728-740. doi:10.1038/s41416-023-02554-x
24. Guo, Xiaodong et al. "Gut microbiota and serum metabolite signatures along the colorectal adenoma-carcinoma sequence: Implications for early detection and intervention." *Clinica chimica acta; international journal of clinical chemistry* vol. 560 (2024): 119732. doi:10.1016/j.cca.2024.119732
25. Adnan, Darbaz et al. "Early-onset Colon Cancer Shows a Distinct Intestinal Microbiome and a Host-Microbe Interaction." *Cancer prevention research (Philadelphia, Pa.)* vol. 17,1 (2024): 29-38. doi:10.1158/1940-6207.CAPR-23-0091

## Figures and Figure Captions

Place your figures in this section with their caption located below the figure.

**Figure 1.** Consort diagram showing how studies were included or excluded in this study. Studies before 2023 were removed along with animal studies, and studies not related to CRC and diet or the microbiome

Study	Type of study	Sample size	Variables included	Date	Findings
14	Case study	589 patients	Microbiome, CRC development	April 2024	High fecal calprotectin reflects higher risk of CRC development, and is caused by increased <i>F. nucleatum</i> and a low dietary fiber.
15	Cohort study	460 patients	Age, microbiome in CRC	April 2024	No significant changes of marker taxa in CRC gut microbiome related to age, but there are changes compared to healthy controls. Findings suggest that CRC has more effect on the gut microbiome than age does.
16	Case study	900 patients	Age, diet, microbiome in CRC	April 2024	A westernized diet, including ultra-processed foods like bread, chips, candy bars etc., have a very negative impact on human health and gut microbiome, and can worsen CRC tumorigenesis. There is an abundance of pro-inflammatory bacteria, and a decrease in short chain fatty acid producing bacteria.
17	Meta-analysis	692 patients 602 controls	Age, microbiome in CRC	February 2023	The odds of CRC were increased with the presence of colibactin, <i>F. nucleatum</i> , and <i>E. coli</i> was 2. However, <i>E. coli</i> was increased in younger healthy controls and decreased in younger CRC patients.
18	Case Control Study	11 patients 22 controls	Microbiome, CRC development, diet	March 2024	Alpha diversity was positively associated with CRC. <i>Fusobacteria</i> , <i>Gemellaceae</i> , and <i>Peptostreptococcus</i> were strongly, positively associated with CRC.
19	Case study	276 patients	Age, microbiome in CRC	February 2024	EOCRC patients were almost 2 times more likely to have stage IV cancer than LOCRC (28% vs 15%). EOCRC had higher alpha diversity, and an enrichment of <i>Akkermansia</i> and <i>Bacteroides</i> , while LOCRC tumors showed greater relative abundances of <i>Fusobacterium</i> , and <i>Escherichia</i> . <i>Fusobacterium</i> abundance was related to patient survival in EOCRC.
20	Cohort study	17627 patients	Age, diet, microbiome in CRC	August 2023	A diet high in sulfur-rich foods increases the risk of CRC. In contrast, consuming more fruits, yellow vegetables, whole grains, legumes, leafy

					vegetables, and cruciferous vegetables is linked to lower risk, as reflected in a lower score. A higher score comes from the intake of foods like low-calorie beverages, French fries, red and processed meats.
21	Case study	94 patients	Microbiome, KRAS mutation, CRC development	April 2024	26 species were found with significant differences between KRAS mutant and wild type patients. 14 of these were more abundant in mutant patients and 12 were more abundant in wild type patients (21). <i>Anaerostipes</i> were found to notably decrease in wild-type patients and <i>Clostridium cadaveris</i> was found to greatly increase in KRAS mutant patients.
22	Cohort study	25 patients 25 controls	Microbiome in CRC	October 2023	In CRC <i>Streptococcus</i> , <i>Lactobacillus</i> , <i>Klebsiella</i> , <i>Intestinibacter</i> , <i>Ralstonia</i> , <i>Alistipes</i> , <i>Pseudomonas</i> , <i>Peptostreptococcus</i> , <i>Faecalibaculum</i> and many others were found to be abundant (22).
23	Case study	3363 patients	Age, microbiome in CRC	June 2024	<i>E. coli</i> , <i>B. fragilis</i> , and <i>F. nucleatum</i> are very prevalent in CRC and inside tumors. Inside of tumors the prevalence of <i>E. coli</i> , <i>B. fragilis</i> , and <i>F. nucleatum</i> are 20.7%, 6.1%, and 8.9% respectively. CRC with intratumoral <i>F. nucleatum</i> are associated with more aggressive, and abnormal cancer cells. CRC with intratumoral <i>E. coli</i> was associated with LOCRC
24	Cohort study	45 patients 20 controls	Microbiome in CRC	May 2024	Significant difference in beta diversity. Enterococcus, Pseudomonas, Fusobacterium, Alistipes and Blautia were dominant in CRC patient microbiomes, while Subdoligranulum, Bacteroides, Eubacterium were the primary bacteria in healthy controls.
25	Cohort study	701 patients 693 controls	Age, microbiome in CRC	Jan 2024	There was a difference in beta diversity between CRC and healthy controls across all disease stages. EOCRC saw an increase in known CRC promoting bacteria such as Akkermansia muciniphila, Bacteroides fragilis, Methanobrevibacter smithii. LOCRC patients had an increase in CRC promoting bacteria and a decrease in protective species.

Table 1 shows individual findings of each study. This was included so that any details not in common with other papers, and therefore not in the results section, were expressed somewhere.



# **Understanding Financial Literacy in Immigrants and Non-English Speakers in the United States By Jack Andreuccetti**

## **Abstract**

The financial landscape is becoming increasingly complex, with the rise of diverse financial products, technologies, and new investment options, making financial knowledge more crucial than ever. However, immigrants and non-English speakers in the United States face significant obstacles in acquiring the financial literacy essential to their economic security. This paper examines the relationship between financial knowledge, immigration, time spent living in the U.S., and English language proficiency. Drawing on data from two national data sets, the research reveals that non-native English speakers demonstrate lower financial literacy in the United States, proportionate to their English proficiency. New immigrants also exhibit poor financial sophistication, which improves the longer they reside in the U.S. While financial education courses positively affect financial literacy, their impact varies across demographic groups. This paper concludes by underscoring the critical need for financial literacy initiatives tailored to immigrants and non-English speakers, to increase familiarity and remove barriers to the U.S. financial system. Expanding access to financial education and banking services is essential for integrating these populations, ultimately enhancing their economic security.

**Keywords** financial literacy, immigrants, non-English speakers, financial capabilities, U.S. financial system, English proficiency, financial education, economic security, financial inclusion

## **I. Introduction**

Finances are becoming increasingly complex due to the growth of financial products, fintech innovations, and new investment options like digital currencies, making financial knowledge more important than ever. As a result, financial literacy, the skills and knowledge needed to make informed money-related decisions, has gained prominence in economic research. Despite this increase in its importance, financial literacy remains low in the United States and throughout the world (Lusardi and Mitchell, 2014). Research has shown that this is especially true for ethnic/racial minorities and immigrants (Lusardi and Tufano, 2015), with these groups being disadvantaged compared to the white majority and native-born.

While many authors have considered ethnicity/race and separately the impact of financial education on literacy, there has been little research done on how proficiency in the native language or time in the country since immigration impacts financial knowledge, or how that plays into the benefits from financial education.

This paper aims to explore whether there is a connection between native language proficiency or time spent in the country and financial knowledge. In order to do this, the analysis will focus on several key areas: financial literacy levels among different racial and ethnic groups in the United States, including immigrants, the relationship between financial sophistication and English proficiency, and the effects of financial education.

The analysis finds that non-native English speakers are less likely to be financially sophisticated in the United States, in proportion to how well they speak English. Similarly, the time an immigrant has spent in the country is positively correlated with financial sophistication, ownership of a bank account or home, and participation in the stock market.

This paper begins with a literature review on financial literacy and education. This is followed by a description of the data sources, the NFCS and SIPP, and an outline of the methodology for examining financial literacy across different demographic groups. Next, findings are presented on how financial literacy correlates with English proficiency, immigration status, and financial education. We conclude with policy recommendations to improve financial education for immigrants and non-English speakers.

## **II. Literature Review**

This section explores the existing body of research on financial literacy, with an emphasis on the impact of economic education, demographic disparities, the behavioral implications of financial knowledge, and the effectiveness of financial literacy interventions.

### **Economic Education and Financial Literacy**

Earlier studies establish a significant positive correlation between economic education and financial literacy. This strong positive correlation largely stems from the fact that economic education enables individuals to understand and navigate the world of finances, resulting in better financial decisions and outcomes. An example of this is the research conducted by Lusardi and Mitchel (2014), whose work displays how individuals who experienced formal economic education had higher levels of financial literacy, and thus were more likely to engage in conscientious actions that improved their long-term financial well-being. These included participation in stock markets, cautious borrowing, and saving.

Prior research performed by Mandell (2008) establishes the importance of early financial education. His findings highlight how students who took high school courses on personal finances or management of money tended to score significantly higher when their proficiency in financial literacy was assessed. Thus, these results implicate that early education is crucial in improving the financial literacy competence among the general population. However, other studies have found that the impact of financial education programs fluctuates depending on several variables. Research by Cole, Paulson, and Shastry (2016) found that albeit financial education is generally effective, factors such as quality of instruction, ethnicity, and the socioeconomic background of students influence the success of such education. In addition, an ongoing debate exists regarding the efficacy of financial education initiatives. The pessimistic position of Willis (2011) on this subject stems from her research, which explains how numerous financial programs fail to make a lasting improvement in financial literacy. She argues that a fault exists between education and situations or applications in the real world. Furthermore, a “forgetting curve” may limit the benefits of financial education courses, as individuals lose previously gained financial knowledge as more time passes without reiteration (Fernandes,

Lynch, & Netemeyer, 2014). Thus, these evaluations suggest more research should be done on the relationship between financial education and the actual application of economic principles.

### **Financial Literacy Across Demographics**

Past studies on financial literacy consistently reveal large disparities between demographic groups. This research establishes how race, ethnicity, gender, income, and age all play an immense role in assessing one's financial literacy. Such research also establishes that the disparities exist for a number of reasons, including broader systematic inequalities, family structure, and societal norms, as well as unequal access to opportunities such as education. The relationship between gender and financial literacy levels contains one of the most distinct and significant findings in the literature. There exists a large gender gap in financial literacy, with women having lower levels of financial literacy, on average, than men. This phenomenon exists across numerous studies and situations. An example is the work of Bucher-Koenen et al. (2017), which attributes this imbalance to differences in socialization, risk management, and the gendered division of work in households. Women are, on average, less involved in household financial decisions, hence having lower understandings of financial literacy. This gendered difference in financial knowledge also has implications for women's financial security in the long run, as they are more likely to have financial hardships due to lower lifetime earnings and longer life expectancies. Lusardi and Mitchell (2011) discovered similar occurrences, but also found that women are aware of their shortfalls in regards to financial literacy, whereas men are less conscientious of their own levels of financial knowledge. In other words, while women are on average less financially literate, they are generally more aware of their position in this regard. There also exist significant racial and ethnic disparities in financial literacy. Lusardi and Tufano (2015) noticed that Black and Hispanic individuals are much more likely to display low levels of financial literacy. They suggest this racial imbalance stems from these groups being more likely to use high-cost borrowing methods and being less likely to own savings or retirement accounts. Ultimately, the roots of these disparities are varied, including historical discrimination and inequalities in access to education (FINRA, 2022). Additionally, Lusardi et al. (2023) found that those who identified as Black and Hispanic were much more financially vulnerable than other ethnicities such as White and Asian. The researchers attested the imbalances to a higher likelihood of single parenthood, increased debt, and poor money management practices, as well as income shocks and a lack of savings and wealth. Moreover, the combination of race and gender disparities creates compounded disadvantages for Black and Hispanic women, who Clark et al. (2021) found face higher levels of economic insecurity and are less likely to have access to resources that aid their financial literacy.

The effect of these disparities goes beyond the scope of individual finances and extends to broader financial inequality. The lower levels of financial literacy among previously discussed demographic groups contribute to widening the wealth gap, since individuals in these marginalized groups have proven less likely to perform wealth-building activities such as investing in financial markets or real estate. Furthermore, financial illiteracy worsens the effects

of economic shocks, like losing a job or experiencing a medical emergency. This occurs because individuals with low financial literacy are less likely to have savings or funds specifically for emergencies (Lusardi, Schneider, & Tufano, 2011). Thus, these findings accentuate the importance of financial literacy in broader efforts to reduce financial inequality.

### **Impact of Financial Literacy on Financial Behavior**

Literature and prior studies done on the subject show that financial literacy plays a critical role in determining financial behavior, as those with greater financial literacy have a higher likelihood of engaging in beneficial economic activities. Research done by Lusardi and Mitchell (2011) exhibits the strong relationship between financial literacy and retirement planning. These authors found that answering one additional financial question correctly was associated with a 3-4 percentage point higher chance of planning for retirement. Also, individuals with higher financial literacy are much more likely to calculate their retirement needs and use retirement savings plans, such as a 401(k). Planning for retirement is key to accumulating wealth – individuals who have a retirement plan accumulate significantly more wealth than those who do not. The correlation between low financial literacy and lack of retirement planning poses a concerning issue, as future financial security is becoming increasingly dependent on decisions made by individuals and less on pensions.

Furthermore, The National Financial Capability Study (2022) supports the link between financial literacy and positive financial behavior. In the study, individuals who expressed higher levels of financial literacy were found more likely to pay their credit card balances in full each month, avoid late fees, and have better credit scores. The above practices lead to less financial stress and greater stability. However, some research suggests that, despite its necessity, financial literacy is not sufficient to guarantee positive financial habits. Psychological biases, such as overconfidence or recency bias, can also lead individuals to make poor financial choices, regardless of their financial literacy (Kahneman & Tversky, 1979). Additionally, outside factors, such as economic conditions or societal standards, affect financial decisions irrespective of measures of financial literacy (Thaler & Sunstein, 2008). Hence, these findings suggest efforts to improve financial literacy should be bolstered by an environment that limits cognitive biases and promotes positive financial decisions.

### **Financial Literacy Interventions**

Researchers are unclear on the benefits of financial education programs, since studies have produced mixed results. While some approaches led to positive results, others had quite limited effects, causing uncertainty about the best form of financial education. Lusardi et al. (2020) claim the efficacy of financial literacy education depends on factors including content, delivery method, and timing. More specifically, their research found that education fit to the specific needs and situations of individuals tended to be more effective than universal programs. Brown et al. (2016) provide an example of the influence of specialized financial education. Programs that focused on practical skills, such as budgeting and saving, and

that delivered these concepts through interactive methods, were more likely to cause positive behavioral changes. In addition, timing of educational efforts matters, as initiatives that occurred during the transition to adulthood or before major financial decisions, had a significantly larger influence than those delivered at other times. Furthermore, the results of Corsini and Giannelli (2021) displayed a strong connection between specific financial education and a significant increase in financial literacy.

On the other hand, Fernandes, Lynch, and Netemeyer (2014) found the average effect size of educational efforts was small in their analysis of financial literacy programs. This result suggests many programs have a limited impact on financial behavior and decisions. The researchers claim financial education requires other forms of support to be most effective. The other forms of support suggested include behavioral nudges, access to financial products, and regulations to protect individuals. The Consumer Financial Protection Bureau (CFPB) backs this view, emphasizing the importance of combining financial literacy programs with greater efforts of financial inclusion and consumer protection.

While many studies have examined the link between ethnicity or race and financial literacy, as well as the effects of financial education on financial knowledge, there is limited research on how English language proficiency or the duration of time since immigration affects financial literacy. Furthermore, the impact of these factors on the effectiveness of financial education has received little attention. This paper aims to explore these topics in depth.

### **III. Data**

This paper utilizes two primary data sources to explore the financial knowledge of diverse populations in the United States: (1) the 2021 National Financial Capability Study (“NFCS”) administered by the Financial Industry Regulatory Authority<sup>1</sup>, and (2) the 2022 Survey of Income and Program Participation (“SIPP”), from the United States Census Bureau<sup>2</sup>.

#### **National Financial Capability Study (NFCS), 2021**

The NFCS is a large-scale national survey designed to gather extensive information on financial literacy and related demographic data across the United States. The survey has 27,118 unique participants and includes questions specifically targeting respondents’ experiences with financial education courses, their financial behaviors, and their financial literacy. Financial literacy is tested by a set of seven multiple-choice questions. This survey also asks the self-perceived financial knowledge of respondents, which can be compared against their actual results from the financial literacy test. The NFCS is one of the few surveys that have data for finance-specific education and tests for financial literacy. This, coupled with robust demographics, make it an ideal data set to explore the topics addressed in this paper. The NFCS aimed to create a set of questions that could reliably measure a person’s financial literacy by testing their ability to apply basic financial concepts to real-life situations. See Appendix 1 for the seven specific financial literacy test questions asked, which includes the “Big

Three” questions that were developed to assess financial literacy (Lusardi and Mitchell 2008). They are the first three questions in Appendix 1 and were constructed to evaluate:

- Numeracy: The ability to perform a simple calculation on interest compounding;
- Understanding of Inflation: The ability to understand inflation in a direct financial decision;
- Understanding of Risk Diversification: Knowledge of risk diversification, in this case in relation to individual stocks and mutual funds.

The remaining questions were derived to cover the topics of interest rates, mortgage payment amortization, and bond prices. All questions were intended to be simple, avoiding complex calculations, and formatted as multiple-choice options to make it easier for the survey respondents. Respondents were also given the option to select “Do not know”, “Refuse to answer”, or “Prefer not to say” to try and eliminate inaccuracies in the results from guessing. Before being included in the survey, the NCFSS questions were crafted to ensure they accurately measured the financial literacy topics outlined above. This involved pilot testing with diverse demographic groups to ensure the questions were understandable and could effectively differentiate between varying levels of financial literacy.

### **Survey of Income and Program Participation (SIPP), 2022**

The 2022 SIPP is a comprehensive survey conducted by the U.S. Census Bureau, which collected detailed data on income sources, program participation, and labor force activities among 41,070 individuals and their corresponding households across the United States. This survey captures a broad spectrum of financial information, including employment, income, and wealth; and includes unique language and immigration status data such as English language proficiency, language spoken at home, and the amount of time living in the United States. From a financial perspective, the data does not directly capture financial literacy, but instead has detailed data on types of financial products owned across a comprehensive set of potential product types. The survey is backward looking to the previous calendar year; data collected in 2022 reflect experiences from 2021, etc. To ensure its validity, the survey undergoes pretesting, subject matter expert reviews, and post-collection data consistency checks.

## **IV. Methodology**

This section outlines the empirical strategy, data sources, and financial literacy indicators used across two separate national data sets.

**Empirical strategy.** To analyze differences in financial literacy across proficiency in the English language, immigration status, time in the U.S., and whether a financial education course has previously been taken.

To research these variables, three financial literacy measures and six proxy indicators were analyzed across the two data sets, using multivariate regression analyses – both ordinary least squares (OLS) and Probit (binary). In the models a financial literacy score ( $Y_i$ ) of individual  $i$  is related to a set of controls, where  $X_k$  is a vector of  $k$  demographic variables including the respondent's race/ethnicity, gender, age, education, and household or personal income. For the NCFS data, additional variables of a financial education course and the number of financial education course locations are included. The SIPP data has immigration status, and time living in the United States, and proficiency speaking English as supplemental variables. The additional variables are represented as  $X_z$ , a vector of  $z$  variables below.

$$Y_i = \beta_0 + \beta_k X_k + \beta_z X_z + \varepsilon$$

The NFCS dataset contains three variables that we use to evaluate different measures of financial literacy. The first variable is the result of a financial literacy test given to each participant as part of the survey. The second is a self-assessment asking the participant to rate their own financial knowledge, while the third asks if a financial literacy course was ever taken. A fourth proxy variable is created that measures financial product ownership as an alternative to the other three financial literacy variables.

From the SIPP data set this paper creates four different proxy variables to assess financial sophistication, a construct that measures various tiers of financial product ownership as a substitute for a direct financial literacy variable.

### **Measures of Financial Literacy**

A Total Financial Literacy score was created from the answers to the financial literacy questions described above and found in Appendix 1. The respondent's score was the sum of each correctly answered question out of the seven total questions, with the Total Financial Literacy variable having a maximum possible score of seven and a minimum of zero. Additionally, this paper explores respondents' view of their own financial knowledge, as self-assessed in the survey. The survey asked how the respondent would assess their overall financial knowledge, on a scale from one to seven: one meaning very low and seven meaning very high. See Appendix 2 for the definition of each NFCS variable used in the analysis.

### **Measures of Financial Sophistication, a proxy for Financial Literacy**

Unlike the NCFS survey, the SIPP data did not have a direct financial knowledge test, so proxy variables were created to estimate financial literacy.

The SIPP Total Financial Sophistication variable was constructed as the sum of financial and investment products that each respondent owned. To create the variable, participants were given a single point for ownership of a financial and investment product in each of ten separate categories. The points were then summed to create a variable that ranged from zero to ten, with zero being no ownership of any financial or investment products, and ten being ownership of at

least one product in each of the ten separate categories. The ten categories are: 1) Annuities; 2) Bank Accounts; 3) Bonds and Fixed Income; 4) Educational Savings Accounts; 5) General Partnerships, Limited Partnerships, and Investments in Businesses; 6) Life Insurance; 7) Other Investments – art, collectibles, etc.; 8) Real Estate, including ownership of a primary home; 9) Retirement Accounts; and 10) Stocks and Equity.

For example, if the respondents owned a checking account and also their primary home, they would get a point each for the Bank Accounts and Real Estate categories, giving them a Total Financial Sophistication score of two.

In addition to the total score, four tiers of financial sophistication were created as subsets of SIPP Total Financial Sophistication, to segment participants into groups. This was to help determine if there were characteristics unique to each grouping. The Tiers are:

- 1) Tier 1 – Any respondent who had a Bank Account product;
- 2) Tier 2 – Respondents with a Bank Account and also owned Real Estate;
- 3) Tier 3 – A Bank Account, Real Estate, and investments in Stocks, Bonds or other Fixed Income;
- 4) Tier 4 – The respondents with the above and additional ownership of other financial products or investments in any other category.

The concern with this approach, as an indirect proxy for financial knowledge, is that the ownership of financial and investment products may be more correlated to wealth than financial sophistication. To address this, the SIPP Total Financial Sophistication is compared to a similar variable constructed using the NCFS data in Section V. While not decisive, the comparison does support the approach of financial product ownership as a proxy for financial literacy.

See Appendix 3 for the definition of each SIPP variable used in the analysis. Note the income variable is for each participant's household and is a self-reported annual number. So, each member of the household, regardless of age and employment status, should have the same annual household income number. Furthermore, data from respondents of all ages was used. A secondary analysis looking solely at those of working age, just high school age, or just retirees, etc. could be considered.

## **V. Analysis of Empirical Findings**

The primary aim of the analysis is to explore the determinants of financial literacy, especially as it relates to race/ethnicity, immigration, time in the United States, and proficiency in the English language.

### **Multivariate Regression Results – SIPP Data**

#### *Total Financial Sophistication Score*

Table 1 provides multivariate regression results of the factors shaping Total Financial Sophistication, the proxy for financial literacy. These results indicate that proficiency in speaking



English is positively correlated with financial sophistication, as does being born in the United States. For immigrants, there is a clear trend of improved financial sophistication the longer an immigrant has been in the United States, with the shortest time periods having the most negative correlation coefficients. As time in the country increases to over 20 years, the coefficient becomes negligible and is not statistically significant. A possible hypothesis is that a newcomer's unfamiliarity with the U.S. financial system creates a barrier to financial system access. After 20 years, immigrants may reach a level of familiarity with the system that looks like financial knowledge comparable to that of native-born individuals. Similarly, it's easy to see English as an obstacle to the financial system that improves with competence in the language. Nevertheless, the results show that English proficiency and the length of time immigrants have spent in the country have a positive relationship with financial literacy.

*[Insert Table 1 about here]*

#### *Financial Sophistication Tiers*

Table 1 also shows results for the Financial Sophistication Tiers 1-4. For Tier 1 (ownership of a bank checking or savings product), many fewer factors are significant when determining who has a bank account. The only significant factors are age, gender, education, time in the United States, and being either Black or White. It is reasonable that income no longer plays an important role when deciding to own a bank account, as the utility of an account does not change much with wealth. However, being in the United States for less than five years is an important factor that makes bank account ownership less probable. Investigating the challenges immigrants face with U.S. banking products could facilitate quicker integration into the financial system. This warrants further analysis, as access to a bank account serves as the gateway to the U.S. financial system, and an increase in bank accounts among new immigrants could positively impact their financial well-being.

More surprisingly, speaking English well and being born in the United States are no longer significant for ownership of a bank account. Additionally, identifying as White is positively correlated and Black is negatively linked to owning bank accounts, which is consistent with previous research (Lusardi et al, 2023). Age remains an important factor, but this might only be relevant for pre-adult ages (0-18 years old) where bank account ownership is uncommon. Tier 2 adds ownership of real estate, such as a primary residence, and Tier 3 layers on further ownership of stock or fixed income investments. These tiers, as well as the most advanced Tier 4, have similar characteristics to the overall Total Financial Sophistication score, with the notable exception of Tiers 2 and 3 Financial Sophistication having a negative relationship to participants who identify as Black.

For Financial Sophistication, an Ordinary Least Squares model (OLS) was used, while for the Financial Sophistication Tiers 1-4 probit binary models were used. All estimations use robust standard errors.

#### **Multivariate Regression Results – NCFS Data**

Table 2 provides multivariate regression results of the factors shaping the four dependent variables: Total Financial Literacy, Financial Knowledge Self-Assessment, Participation in a Financial Education Class, and a Financial Sophistication score constructed from the NCFS data.

*[Insert Table 2 about here]*

Predictably, having taken a financial education course in the past is positively correlated to the Total Financial Literacy score, however, the number of places a participant has received financial education (high school, college, employer, or the military) is not. Taking any financial education class is highly correlated to the number of places a respondent received financial education, and either variable can be interchangeable in the models. Still, it is surprising that having taken financial education in more places does not confer additional financial literacy. This may be the result of the type of questions being asked; one would suspect that the frequency, intensity, and level of financial education achieved will start to matter for survey questions that are more difficult. As mentioned in Section III, the NCFS survey was designed to only include straightforward questions that do not require considerable calculation. Note, however, that an increase in financial course locations does positively influence Self-Assessment scores as well as Financial Sophistication, adding further credence to the theory that more financial education is useful but only basic financial education is sufficient for the NCFS financial literacy test questions. Finally, it is important to emphasize how few individuals have participated in financial education classes. In the NCFS survey, only 21% of respondents reported having taken any form of financial education course. This number is roughly the same for Whites, Hispanics, and Asians, but increases to 25% for Blacks.

When looking at the Total Financial Literacy score, race/ethnicity follow patterns shown in previous literature, including Whites and Asians positive correlations, while Hispanic and Blacks have negative correlations to financial literacy. Yet these patterns break down when looking at the participant's self-assessment of financial knowledge, which is either not statistically significant or, in the case of Asians, is negatively correlated, the opposite of actual financial literacy test results. Self-assessment of personal abilities has historically been unreliable, and cultural factors may influence how individuals evaluate their financial skills compared to their peers. Nonetheless, there is a well-documented tendency to overestimate one's actual abilities in survey responses (Kahneman and Tversky, 1979). Consequently, the findings from the Financial Knowledge Self-Assessment model are of limited interest. One last finding worth noting is that Black respondents were more likely to report having taken a financial education course, while Asian respondents were less likely. Although not shown in Table 2, a similar pattern was observed in the number of locations where financial education was received, with Black respondents showing a positive relationship and Asian respondents a negative correlation.

*Validation of the SIPP Financial Sophistication construct, as a proxy for financial literacy*

Finally, to verify the robustness of the SIPP Financial Sophistication variable as a proxy for financial literacy, a comparable variable labeled "financial sophistication" was constructed using

data from the NCFS survey. The NCFS survey was far less detailed than the SIPP survey regarding financial and investment product ownership, featuring only four questions on product ownership compared to the 50+ detailed questions in the SIPP survey. A summary of the differences:

- SIPP Financial Sophistication variable: Derived from 10 product categories that consolidated 51 in-depth survey questions, encompassing all specialized financial and investment products, along with details on how they are owned;
- NCFS Financial Sophistication variable: Only four survey questions on financial and investment products: home, stocks, banking, and retirement account ownership.

As a result, the NCFS variable exhibits a narrower range of differentiation, with scores ranging from zero to four, compared to zero to ten for the SIPP variable. Despite the limited scope and depth of the NCFS financial product questions, a regression model utilizing the NCFS Financial Sophistication variable could still validate the financial sophistication concept as a strong proxy for financial literacy if:

- The NCFS Financial Sophistication regression model exhibits characteristics that closely resemble those of the regression model for NCFS Total Financial Literacy test scores;
- The determinants of the NCFS Financial Sophistication model are also similar to those of the SIPP Financial Sophistication model, which uses a different dataset but comparable factors (such as gender, age, income, education, race/ethnicity, etc.).

Indeed, a model of NCFS Financial Sophistication has the same statistically significant determinants as the NCFS Financial Literacy model (see Table 2). Likewise, the determinants for the NCFS Financial Sophistication model share similar attributes with those of the SIPP Financial Sophistication model. Although not conclusive, these findings support the validity of SIPP Total Financial Sophistication as a proxy for financial literacy. For the Total Financial Literacy, Financial Knowledge Self-Assessment, and Financial Sophistication (constructed from NCFS) models, Ordinary Least Squares was used. For the Financial Education Class variable, a probit binary model was used. All estimations use robust standard errors.

### **Correlation Matrices**

Table 3 presents a correlation matrix for variables using the NCFS data. Notably, there is a strong correlation (86%) between taking any financial education course and the variable that sums financial course locations (up to four: high school, college, employer, and/or the military). As a result, these two variables are nearly interchangeable in the models, and either remains statistically significant in all models when the other is excluded.

*[Insert Table 3 about here]*

Table 4 shows a correlation matrix for variables using the SIPP data. As expected, being White and being born in the United States are both highly correlated with speaking English proficiently.

*[Insert Table 4 about here]*

The analysis in this section reveals that proficiency in English and being born in the United States are positively correlated with financial sophistication. For immigrants, financial knowledge improves the longer they have lived in the U.S., becoming comparable to that of native-born individuals after around 20 years. Income is not a significant factor in bank account ownership, though newly arrived immigrants (less than five years) are less much likely to own an account. As expected, financial education has a strong influence on financial literacy.

## **VI. Conclusion**

The results clearly indicate that English-speaking ability and time spent in the United States are positively associated with financial literacy, a relationship that, while intuitive, has not been widely explored in previous research. This finding suggests several potential policy implications, including the importance of simultaneously improving English proficiency and financial literacy for maximum impact. A program designed to address both areas concurrently should be developed and its outcomes carefully evaluated.

Moreover, examining the barriers or unfamiliarity immigrants face with U.S. banking products could help speed up their integration into the financial system. In addition to low Total Financial Sophistication scores, immigrants who have been in the country for less than five years were also less likely to own a bank account. This issue merits further study, as bank accounts are a critical entry point into the U.S. financial system and expanding access for new immigrants could positively impact their overall financial well-being. Potential policy initiatives could include simplifying access to banking products for new immigrants, supported by subsidies to encourage bank account ownership and offset any additional credit risk for banks. These special gateway accounts could be incorporated into new arrival materials for immigrants, possibly as auto-enroll or opt-out programs.

Lastly, it is important to highlight how few individuals have participated in a financial education class. Only 21% of respondents in the NCFS survey reported taking any type of financial education course. Achieving universal financial education remains a significant challenge, with financial literacy particularly low among new immigrants and non-English speakers. However, even small changes in the language and timeliness in which financial education is delivered could have a meaningful impact on financial security for all United States residents.

**Table 1 – Determinants of Financial Sophistication (SIPP Data)**

**Financial Sophistication Scores (Total and Tiers 1-4)**

	Total	Tier 1	Tier 2	Tier 3	Tier 4
	OLS	Probit (binary)	Probit (binary)	Probit (binary)	Probit (binary)
Gender	0.126 *** (0.011)	0.033 * (0.018)	0.042 *** (0.014)	0.157 *** (0.015)	0.112 *** (0.016)
Age	0.013 *** (3.23E-4)	0.018 *** (4.76E-4)	0.017 *** (4.60E-4)	0.017 *** (4.84E-4)	0.012 *** (4.49E-4)
Income	1.58E-06 *** (1.11E-7)	-2.51E-08 (8.49E-8)	1.44E-06 *** (1.97E-7)	1.50E-06 *** (1.84E-7)	1.31E-06 *** (1.41E-7)
Education	0.344 *** (0.003)	0.466 *** (0.006)	0.294 *** (0.004)	0.396 *** (0.006)	0.346 *** (0.006)
White	0.237 *** (0.028)	0.122 ** (0.048)	0.198 *** (0.043)	0.275 *** (0.049)	0.306 *** (0.055)
Black	-0.043 (0.031)	-0.241 *** (0.053)	-0.330 *** (0.048)	-0.210 *** (0.054)	0.026 (0.060)
Hispanic	-0.045 (0.029)	-0.063 (0.050)	0.032 (0.046)	0.010 (0.053)	0.023 (0.060)
Asian	0.215 *** (0.036)	0.001 (0.061)	0.179 *** (0.052)	0.208 *** (0.059)	0.209 *** (0.065)
BorninUS	0.185 *** (0.032)	0.070 (0.048)	0.135 *** (0.039)	0.162 *** (0.044)	0.228 *** (0.047)
WellSpeakEnglish	0.078 *** (0.009)	0.001 (0.015)	0.016 (0.013)	0.084 *** (0.017)	0.122 *** (0.019)
TimeinUSA_5	-0.251 *** (0.049)	-0.332 *** (0.083)	-0.613 *** (0.082)	-0.757 *** (0.112)	-0.573 *** (0.128)
TimeinUSA_5_10	-0.115 ** (0.048)	0.079 (0.079)	-0.319 *** (0.067)	-0.351 *** (0.079)	-0.250 *** (0.087)
TimeinUSA_10_20	-0.099 ** (0.043)	0.152 ** (0.070)	-0.085 (0.056)	-0.133 ** (0.065)	0.009 (0.070)
TimeinUSA_20	0.019 (0.037)	0.218 *** (0.059)	0.095 ** (0.046)	0.038 (0.052)	0.047 (0.057)
Constant	-1.372 *** (0.196)	-3.036 *** (0.353)	-2.617 *** (0.371)	-4.805 *** (0.650)	-9.788 (18.494)
R <sup>2</sup> [N = 41,070]	0.483	0.486	0.284	0.306	0.239

Note: Robust standard errors in parenthesis; \*\*\* p<0.01; \*\* p<0.05; \*p<0.1. Other controls include region of residence.

Tier 1 - Financial Sophistication - Any respondent who had a Banking product

Tier 2 - Financial Sophistication - Respondents with a Banking product and also owned Real Estate

Tier 3 - Financial Sophistication - A Banking product, Real Estate, and investments in Stocks, Bonds or other Fixed Income

Tier 4 - The respondents with the above and additional ownership of other financial product or investments in any other another category

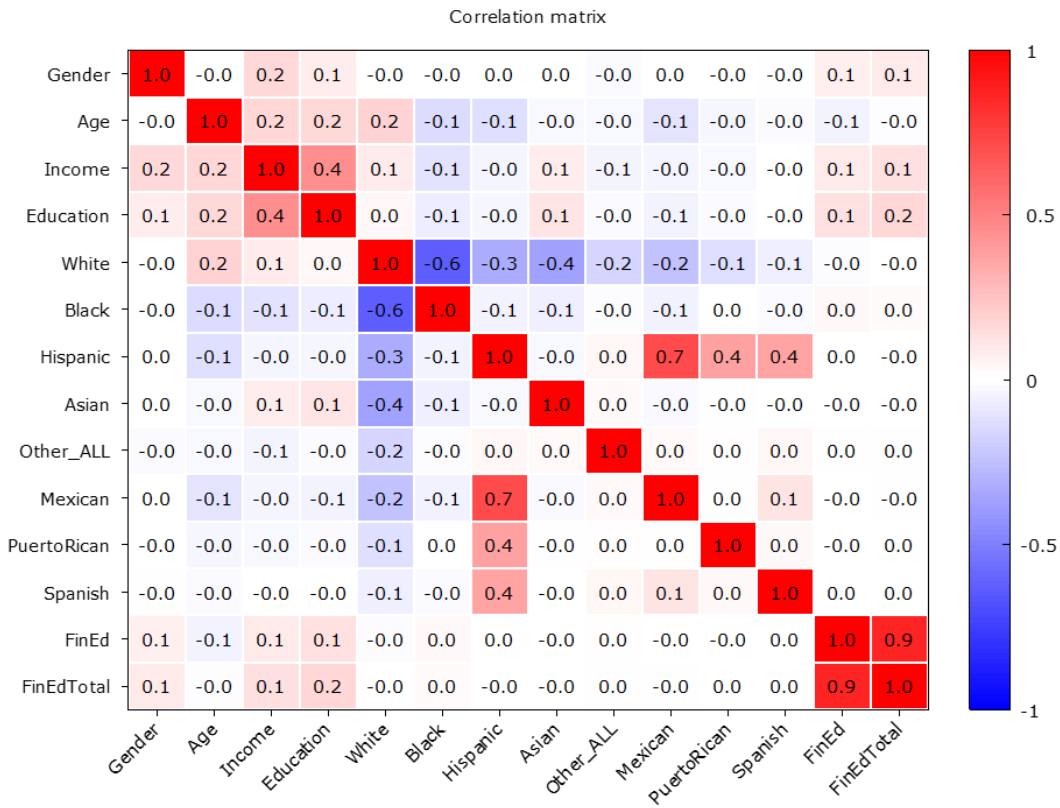
**Table 2 – Determinants of Financial Literacy, Financial Knowledge Self-Assessment, and Financial Sophistication (NCFS Data)**

	<b>Total Financial Literacy</b>	<b>Financial Knowledge Self-Assessment</b>	<b>Taken a FinEd Class</b>	<b>Financial Sophistication (NCFS)</b>
	OLS	OLS	Probit (binary)	OLS
Gender	0.687 *** (0.020)	0.299 *** (0.018)	0.154 *** (0.018)	0.125 *** (0.011)
Age	0.239 *** (0.006)	0.115 *** (0.005)	-0.068 *** (0.006)	0.136 *** (0.003)
Income	0.145 *** (0.005)	0.133 *** (0.005)	0.031 *** (0.005)	0.242 *** (0.003)
Education	0.232 *** (0.007)	0.083 *** (0.006)	0.103 *** (0.006)	0.105 *** (0.004)
White	0.223 *** (0.046)	-0.006 (0.043)	-0.003 (0.040)	0.096 *** (0.027)
Black	-0.462 *** (0.052)	0.070 (0.050)	0.158 *** (0.045)	-0.071 ** (0.031)
Hispanic	-0.116 *** (0.045)	-0.090 ** (0.041)	0.006 (0.039)	-0.059 ** (0.026)
Asian	0.358 *** (0.062)	-0.194 *** (0.053)	-0.148 *** (0.053)	0.058 * (0.035)
Taken a Financial Ed	0.407 *** (0.049)	0.214 *** (0.038)	N/A	0.010 (0.028)
How many Fin Educa	0.014 (0.027)	0.204 *** (0.019)	N/A	0.105 *** (0.015)
Constant	0.238 *** (0.056)	3.316 *** (0.054)	-1.240 *** (0.051)	0.263 *** (0.033)
$R^2$ [N = 27,118]	0.265	0.139	0.028	0.408

Note: Robust standard errors in parenthesis; \*\*\* p<0.01; \*\* p<0.05; \*p<0.1. Other controls include region of residence.

**Table 3 – NCFS Correlation Matrix**

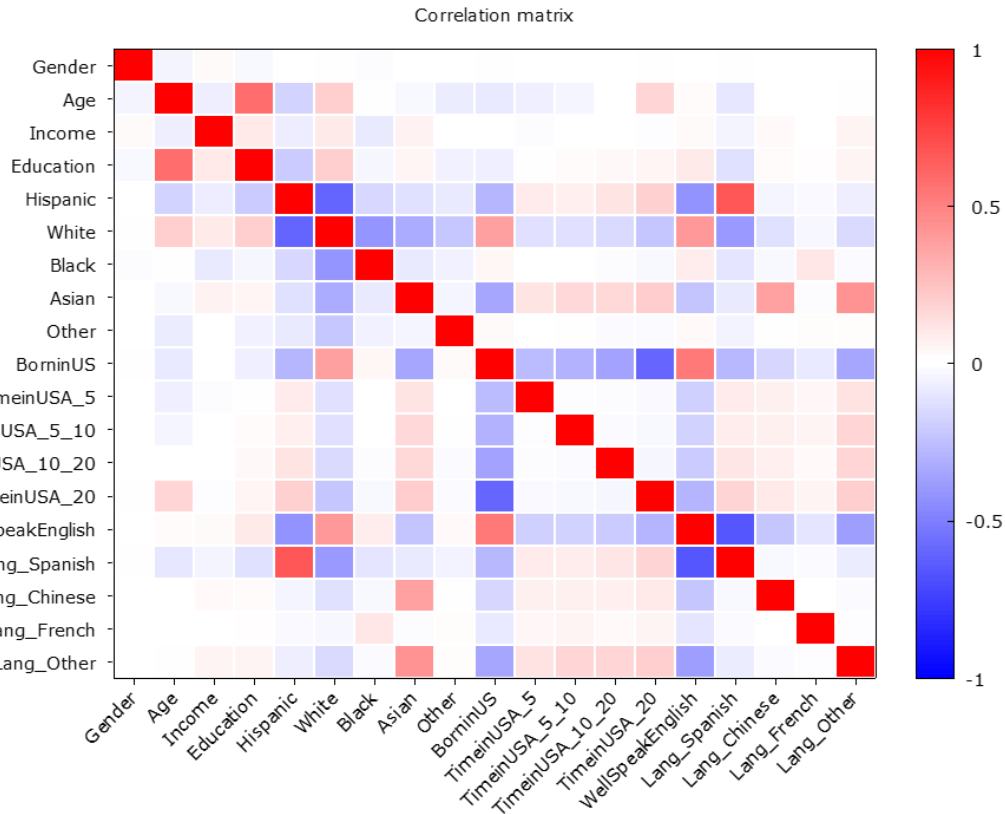
	Gender	Age	Income	Education	White	Black	Hispanic	Asian	Other	Mexico	Puerto Rico	Spain	Financial Education	#of Financial Education Locations
Gender	1.000	-0.009	0.156	0.079	-0.006	-0.005	0.003	0.002	-0.027	0.004	-0.005	-0.011	0.0709	0.0893
Age		1.000	0.164	0.159	0.179	-0.145	-0.133	-0.036	-0.037	-0.113	-0.045	-0.022	-0.0548	-0.0147
Income			1.000	0.445	0.085	-0.116	-0.047	0.077	-0.052	-0.046	-0.034	-0.006	0.0912	0.128
Education				1.000	0.039	-0.078	-0.043	0.109	-0.022	-0.054	-0.024	-0.007	0.1257	0.1643
White					1.000	-0.633	-0.331	-0.372	-0.170	-0.247	-0.130	-0.069	-0.0197	-0.008
Black						1.000	-0.054	-0.070	-0.018	-0.057	0.006	-0.023	0.0331	0.0232
Hispanic							1.000	-0.035	0.042	0.707	0.372	0.366	0.0028	-0.0062
Asian								1.000	0.032	-0.029	-0.010	-0.003	-0.0055	-0.0082
Other									1.000	0.033	0.016	0.044	0.011	0.0128
Mexico										1.000	0.010	0.112	-0.0019	-0.0109
Puerto Rico											1.000	0.033	-0.0025	0.002
Spain												1.000	0.0098	0.0074
Financial Education													1	0.8616
#of Fin Ed Locations														1



Correlation Coefficients, using the observations 1 – 27118  
 Two-tailed critical values for n = 27118: 5% 0.0119, 1% 0.0156

**Table 4 – SIPP Correlation Matrix**

	Gender	Age	Income	Education	Hispanic	White	Black	Asian	Other	BorninUS	TimeinUSA_5	TimeinUSA_5_10	TimeinUSA_10_20	TimeinUSA_20	WellSpeakEnglish	Lang_Spanish	Lang_Chinese	Lang_French	Lang_Other
Gender	1.000																		
Age	0.090	1.000																	
Income	-0.024	0.514	1.000																
Education	-0.092	0.192	0.092	1.000															
Hispanic	-0.172	0.194	-0.077	0.088	1.000														
White	0.194	0.162	0.193	-0.042	0.611	1.000													
Black	0.088	0.128	-0.086	0.093	0.162	0.421	1.000												
Asian	0.059	0.045	-0.063	-0.069	0.045	-0.332	-0.225	1.000											
Other	-0.005	-0.005	-0.005	-0.009	0.000	0.083	0.083	0.083	1.000										
Born in US	0.000	-0.091	-0.091	-0.091	-0.091	-0.091	-0.091	-0.091	-0.091	1.000									
TimeinUSA<5yrs	-0.002	-0.094	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	1.000								
TimeinUSA 5-10yr	-0.006	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	1.000							
TimeinUSA 10-20yr	0.000	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	1.000						
TimeinUSA>20yrs	0.000	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	-0.097	1.000					
EnglishProf	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000				
Lang_Spanish	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000			
Lang_Chinese	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000		
Lang_French	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	
Lang_Other	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000



Correlation Coefficients, using the observations 1 – 41070  
 Two-tailed critical values for n = 41070: 5% 0.0097, 1% 0.0127



## Works Cited

- Brown, Meta, John Grigsby, Wilbert van der Klaauw, Jaya Wen, and Basit Zafar. "Financial Education and the Debt Behavior of the Young." *Review of Financial Studies* 29, no. 9 (February 24, 2016): 2490–2522. <https://doi.org/10.1093/rfs/hhw006>.
- Bucher-Koenen, Tabea, Annamaria Lusardi, Rob J.M. Alessie, and Maarten van Rooij. "How Financially Literate Are Women? An Overview and New Insights." *SSRN Electronic Journal*, August 2017. <https://doi.org/10.2139/ssrn.2585187>.
- Clark, Robert L., Annamaria Lusardi, Olivia S. Mitchell, and Hallie Davis. "Financial Well-Being among Black and Hispanic Women ." *SSRN Electronic Journal*, February 2021. <https://doi.org/10.2139/ssrn.3792558>.
- Cole, Shawn, Anna Paulson, and Gauri Kartini Shastri. "High School Curriculum and Financial Outcomes: The Impact of Mandated Personal Finance and Mathematics Courses." *Journal of Human Resources* 51, no. 3 (November 30, 2016): 656–98. <https://doi.org/10.3368/jhr.51.3.0113-5410r1>.
- Corsini, Lorenzo, and Gianna Claudia Giannelli. "Economics Education and Financial Literacy Acquisition: Evidence from a Field Experiment." *Journal of Behavioral and Experimental Finance* 32 (August 16, 2021): 100556. <https://doi.org/10.1016/j.jbef.2021.100556>.
- Fernandes, Daniel, John G. Lynch, and Richard G. Netemeyer. "Financial Literacy, Financial Education, and Downstream Financial Behaviors." *Management Science* 60, no. 8 (August 2014): 1861–83. <https://doi.org/10.1287/mnsc.2013.1849>.
- Gibson, John, David McKenzie, and Bilal Zia. "The Impact of Financial Literacy Training for Migrants." *Policy Research Working Papers*, May 2012. <https://doi.org/10.1596/1813-9450-6073>.
- Kahneman, Daniel, and Amos Tversky. "Prospect Theory: An Analysis of Decision under Risk." *The Econometric Society* 47 (March 1979): 263–92. <https://doi.org/10.21236/ada045771>.
- Lin, Judy T., Christopher Bumcrot, Gary Mottola, Olivia Valdes, Robert Ganem, Christine Kieffer, Annamaria Lusardi, and Gerri Walsh (2022). "Financial Capability in the United States: Highlights from the FINRA Foundation National Financial Capability Study." FINRA Investor Education Foundation. [www.FINRAFoundation.org/NFCSReport2021](http://www.FINRAFoundation.org/NFCSReport2021)
- Lusardi, Annamaria, and Olivia S. Mitchell. "Financial Literacy Round the World: An Overview." *SSRN Electronic Journal*, June 2011. <https://doi.org/10.2139/ssrn.1810551>.
- Lusardi, Annamaria, and Olivia S. Mitchell. "The Economic Importance of Financial Literacy: Theory and Evidence." *Journal of Economic Literature* 52, no. 1 (March 1, 2014): 5–44. <https://doi.org/10.1257/jel.52.1.5>.
- Lusardi, Annamaria, and Peter Tufano. "Debt Literacy, Financial Experiences, and Overindebtedness." *Journal of Pension Economics and Finance* 14, no. 4 (October 2015): 332–68. <https://doi.org/10.1017/s1474747215000232>.
- Lusardi, Annamaria, Andrea Hasler, and Paul J. Yakoboski. "Building up Financial Literacy and Financial Resilience." *Mind & Society* 20, no. 2 (July 31, 2020): 181–87. <https://doi.org/10.1007/s11299-020-00246-0>.

- Lusardi, Annamaria, Andrea Hasler, Olivia S. Mitchell, and Alessia Sconti. "Understanding Financial Vulnerability among Asians, Blacks, and Hispanics in the United States." SSRN Electronic Journal, June 29, 2023. <https://doi.org/10.2139/ssrn.4425740>.
- Lusardi, Annamaria, Daniel Schneider, and Peter Tufano. "Financially Fragile Households: Evidence and Implications." National Bureau of Economic Research, May 2011. <https://doi.org/10.3386/w17072>.
- Lusardi, Annamaria. "Financial Literacy and the Need for Financial Education: Evidence and Implications." *Swiss Journal of Economics and Statistics* 155, no. 1 (January 24, 2019). <https://doi.org/10.1186/s41937-019-0027-5>.
- Mandell, Lewis. "Financial Literacy of High School Students." *Handbook of Consumer Finance Research*, 2008, 163–83. [https://doi.org/10.1007/978-0-387-75734-6\\_10](https://doi.org/10.1007/978-0-387-75734-6_10).
- Mitchell, Olivia S., and Annamaria Lusardi. "Financial Literacy and Retirement Planning in the United States." SSRN Electronic Journal, October 2011. <https://doi.org/10.2139/ssrn.1810550>.
- Thaler, Richard H., and Cass R. Sunstein. "Nudge: Improving Decisions about Health, Wealth and Happiness." Yale University Press 19 (August 22, 2008): 356–60. <https://doi.org/10.37974/alf.89>.
- Willis, Lauren E. "The Financial Education Fallacy." *American Economic Review* 101, no. 3 (May 1, 2011): 429–34. <https://doi.org/10.1257/aer.101.3.429>.

# Natural Disasters Effect on Puerto Rican Educational Output By Matthew Cattaneo

## Abstract

This paper analyzes persistent disparities between Puerto Rican educational output and United States educational output. Puerto Rican students frequently underperform on standardized tests, have low graduation rates, and fail multiple classes. These disparities are compared with many different causes including low average income, Puerto Rican refusal to embrace bilingualism, and the constant threat of natural disasters on the island. Each cause has an impact on the poor output, but the research dives deeply into the effect that natural disasters leave on the Puerto Rican education system. The hurricanes and earthquakes are analyzed historically, and their impacts are noted. The two key impacts mentioned are post-disaster trauma and the closing of schools. The paper finishes by analyzing how these effects continue the cycle of poor educational results.

## 1. Introduction

Puerto Rican students' educational outcomes are significantly lower than students in the 50 US states. Puerto Rico has lower standardized test scores and graduation rates than school systems in the United States do. On top of that, they also have higher fail rates as well. This research aims to discuss the statistics behind these disparities along with a deep dive into one of the potential causes mentioned.

The paper proceeds as follows: Section 2 explicitly states and explains the poor educational outcomes. Section 3 introduces and explains the potential causes of these outcomes which are average income of teachers/families, Commonwealth refusal to embrace bilingualism, and the threat of natural disasters. The paper then goes into more detail on how natural disasters affect the students. Section 4 explains the different types of natural disasters that hit the island and does a historical analysis of their effects on Puerto Rico. Section 5 details how this cause affects academic output among Puerto Rican students.

## 2. Poor Educational Outcomes

Puerto Rico's education system has been plagued by a variety of poor educational outcomes. These problems emphasize the deeply rooted underlying issues within Puerto Rico's education system. In *Economic Storm: The Crisis of Education in Puerto Rico*, Stefan Trines states how Puerto Rico is behind the US writing, "PR [Puerto Rico] keeps trailing the U.S. mainland in a variety of education indicators" (Trines, 2018). He then writes, "The net enrollment rate in secondary education in PR was only 66.6 percent in 2015 compared to 80.5 percent in the U.S." (Trines, 2018). The rest of this section now goes into education indicators that Puerto Rico struggles with, which are low standardized test scores, low graduation rates, and high fail rates.

### 2.1 Low Standardized Test Scores

Puerto Rican students take the same standardized tests as American students and are graded on the same scale, yet Puerto Rican students consistently perform poorer than American students. In *Puerto Rico*, the author writes, “Only 6% of students in Puerto Rico scored at or above grade level in eighth grade compared to 68% of students in the Continental United States” (League of United Latin American Citizens, 2023). This shocking statistic shows just how big the gap in educational output is between Puerto Rico and the United States. Puerto Rico public eighth grade students scored 23% below the national average in standardized tests in 2013. In *Economic Storm: The Crisis of Education in Puerto Rico*, Trines analyzes that statistic stating, “The tests demonstrated that 89 percent of Puerto Rican fourth-graders and 95 percent of eighth-graders had mathematics abilities below the ‘basic level’” (Trines, 2018).

## **2.2 Low Graduation Rates**

Students of Puerto Rico often do not graduate from high school. In fact, many Puerto Rican students drop out of schooling altogether. Since 2012, the total enrollment of students in the Puerto Rican public school system has been steadily declining. The total enrollment in 2012 was 434,609 students whereas the total enrollment in 2022 was 250,668 students, almost 200,000 less. In *Puerto Rico*, the author states, “19.10% of all students do not finish 9<sup>th</sup> grade” (League of United Latin American Citizens, 2023). This means that a little over 80% of students finish their first year of high school with the rest dropping out and not graduating. In *¿Cual Fue la Tasa de Graduación?*, the data shows that “75.70%” of eligible students graduated in 2021, “73.76%” graduated in 2022, and “74.99%” graduated in 2023 (Perfil Escolar, 2023). The lowest graduation rate for the 2021-2022 school year was 75.71% in Arizona, and the highest was 93% in Virginia. Students consistently do not graduate high school or pursue higher education, putting them further behind US school systems.

## **2.3 Fail Rates**

Puerto Rico’s students experience high fail rates. In *Are the Challenges of Puerto Rico's Schools a Taste of What Other Districts Will Face*, Kavitha Cardoza, the author, writes “In 2021, school officials announced that 13,000 students failed all of their classes” (Cardoza, 2023). There are about 276,000 students in the Puerto Rican school system, meaning approximately 4.7% of students failed every single class they had in 2021. As students fail, they are forced to make up the classes they failed during the summer. Students are not able to be taught properly due to being out of school consistently or having to be online and in turn they fail their classes and fall farther behind students the same age who go to continental US schools.

## **3. Different Potential Causes**

Puerto Rico faces many different obstacles that serve as potential causes of lower academic results. The causes that this paper goes into are the average income of teachers/families, the Puerto Rican Commonwealth’s refusal to embrace bilingualism, and the threat of natural disasters towards schooling. Other potential causes that this paper does not go

into are mismanagement of money meant for the educational system and a lack of proper teacher training. Each one of the causes this paper goes into contributes to poor results, but the paper goes further into the wide range of effects that natural disasters cause.

### **3.1 Average Income of Teachers/Families**

Many people in Puerto Rico have a low average income, especially teachers. In *Richest Puerto Rican Zip Codes (2024)*, Kristen Carney writes, “Average household income: \$38,227” (Carney, 2024). The cost of living for the average four-person household is \$44,886. With little money coming in, students struggle to have adequate supplies to prepare themselves properly for school. Without proper supplies, the quality of students learning decreases leading to worse overall learning. Often, teachers must buy school supplies for their schools despite having a low income as well. In *The Disappearing Schools of Puerto Rico*, Jonathan Katz writes, “As in many schools across the country, the teachers at Ramón Torres Rivera Elementary and Middle School often have to buy their own classroom materials” (Katz, 2019). In 2023, the base salary for teachers was \$1,750 a month, a number that has not changed since 2008. The 2023 single person living costs were at an average of \$2,683 per month. The supplies the teachers buy for students, while helpful, are minimal due to the low base salary that the teachers make.

### **3.2 Commonwealth Refusal to Embrace Bilingualism**

The Puerto Rican education system is very similar to the US including their grading scale and what standardized tests they take. There is one key difference that significantly damages reading scores and standardized test scores. This issue is the Puerto Rican Commonwealth’s refusal to embrace bilingualism. Puerto Ricans refuse to embrace bilingualism because they believe it would damage the Puerto Rican culture. Many times, the Puerto Rican government tried to mandate an English reading standardized test, but many Puerto Ricans do not know English. In *Economic Storm: The Crisis of Education in Puerto Rico*, when talking about the issue of bilingualism Trines writes, “Recent Puerto Rican governments have therefore tried to push the use of English in education” (Trines, 2018). This results in extremely low scores for the Puerto Rican students. Teachers also resist the change to English because they believe that it would take up too much time and derive from other subjects of learning. This refusal by families and teachers disconnects Puerto Rican education from the US’s education despite having a similar education system. This in turn causes roadblocks in the Puerto Rican education system resulting in less learning and overall worse educational output by students.

### **3.3 Threat of Natural Disasters and Why Natural Disasters are the Focus**

The island of Puerto Rico is under a constant threat of natural disasters due to its location in the humid Caribbean climate and its proximity to a tectonic plate boundary. As disasters hit the island, power outages are common leading to many issues for Puerto Ricans. Disasters also result in billions of dollars in damage that leads to the island struggling to rebuild due to the costs. These disasters destroy homes and schools along with everything else that gets in their

way, leading to students having less resources readily available to assist them in their learning. Disasters also lead to a shortage of food throughout the island, gravely affecting everyone who lives there. This cause is the center of the research because of the vast implications that disasters have on students and their academic results. There are many different problems in the Puerto Rican school environment brought about by natural disasters that make it hard for students to have a proper learning environment. Every time Puerto Rico tries to resolve the issues brought about by the most recent disaster, a new one strikes, renewing the cycle of struggle for Puerto Rican students.

#### **4. Different Types of Natural Disasters that Hit the Island**

Puerto Rico has two main natural disasters that have gravely affected the students and island as a whole: hurricanes and earthquakes. Puerto Rico has been hit by over ten hurricanes since 2010 including Hurricane Irma, Hurricane Maria, and Hurricane Teddy. In *Education in Puerto Rico*, the author writes, “The disruption of Hurricane Maria is an issue. The closing of hundreds of schools in the wake of falling enrollments, increases in psychological needs among students suffering traumatic experiences, and very high absentee rates all affect the effectiveness of the schools” (PR51st, 2022). These hurricanes cause a plethora of issues in Puerto Rican education. When not faced with hurricanes, Puerto Rico still has the threat of earthquakes, however. The most recent severe earthquake was in 2020 leaving 2/3rds of Puerto Ricans without power.

##### **4.1 Hurricanes**

The island of Puerto Rico has been ravaged by hurricanes since its discovery in 1508, namely Hurricane Maria and Hurricane Fiona. Hurricane Maria was a Category 4 hurricane in 2017 and cut off all the power to the island Causing over \$90 million in damages. In *The Facts: Hurricane Maria’s Effect on Puerto Rico*, the author explains, “Electricity was cut off to 100 percent of the island, and access to clean water and food became limited for most” (Mercy Corps, 2020). The hurricane also led to schools closing for extended time. In *Average Puerto Rican Student Missed 78 Days of School after Maria, Study Finds*, Andrew Ujifusa wrote, “The calculation that the average student missed 78 days of school refers to children ages 5-17. In addition, the study found that the average child in preschool lost 92 days at their educational institution” (Ujifusa, 2018). That becomes a problem because it is considered chronic absenteeism (missing over 10% of the school year) for all students. In *Why is School Attendance Important? The effects of Chronic Absenteeism*, the author explains, “When children are absent from school, they miss out on consistent instruction that is needed to develop basic skills. Children in early grades are particularly susceptible to falling behind in fundamental reading skills, which can have a snowball effect that impacts future learning” (School of Education American University Washington, 2021). Students were not able to be in class and it led to falling behind in their learning. Three years after the hurricane people were still suffering from issues such as power outages for a day. Two years later, Hurricane Fiona hit the island in 2022.

In *Fiona and Ian are Retired as Hurricane Names*, when talking about Hurricane Fiona Johnny Diaz writes, “It caused more than \$3 billion in damage as it tore a destructive path across the Caribbean...” (Diaz 2023). Like Hurricane Maria, Hurricane Fiona caused significant power outages, but not as bad as Hurricane Maria’s were. Regardless, both hurricanes had their respective names retired due to the destruction they caused on the societies they hit.

## **4.2 Earthquakes**

Earthquakes, while not as frequent in Puerto Rico as hurricanes, cause significant damage and problems to Puerto Rico. They cause significant damage with costs that Puerto Rico does not have the money to offer. Earthquakes happen frequently because Puerto Rico lies near the boundary of the North American and Caribbean tectonic plates. The most recent significant earthquakes in Puerto Rico occurred in 2020. These caused damages to a Puerto Rican island that was still struggling from the aftermath of Hurricane Maria that hit three years prior. These earthquakes damaged homes and buildings. In *Puerto Rico Earthquakes*, the author states, “In Yauco, 3,261 homes were damaged including 62 that collapsed and 245 that are compromised and at risk of collapsing in the continuing quake activity” (Center for Disaster Philanthropy, 2020). The earthquakes caused serious damage that Puerto Rico struggled to recover from.

## **5. How Natural disasters Affect Education**

Natural disasters affect Puerto Rican children’s schooling in numerous ways. They cause post-disaster trauma within the students that causes trouble concentrating, leading to a poor learning environment. They have also resulted in the shutting down of numerous schools making students scramble to find a school they could reasonably get to. Without a proper learning environment caused by these effects, students in turn have lower educational output including their standardized test scores, graduation rates, and fail rates.

### **5.1 Post-Disaster Trauma**

As a result of the constant bombardment of natural disasters on the Puerto Rican commonwealth, many students have developed Post-Traumatic Stress Disorder (PTSD). This trauma has a significant effect on the students’ academic performance, whether it affects their standardized test scores or their ability to learn and focus on classes. In *How Trauma Impacts School Performance*, the author states, “Experiencing trauma, especially at a young age, disrupts young people’s ability to relate to others and manage emotions” (Mental Health America, n.d.). Students cannot convey their emotions or how they feel when they have experienced trauma. Trauma also makes it harder for students to comprehend new information and lowers reading comprehension. The Puerto Rican students have experienced this significant trauma at a young age and their academic results are struggling because of it. Trauma also causes an increase in stress. This increase in stress leads to more nerves before standardized testing and the students educational output suffers because of it. As stress levels in an individual increase the likelihood of the person underperforming on standardized tests increases. In *Testing, Stress, and*

*Performance: How Students Respond Physiologically to High-Stakes Testing*, the authors of the study write, “We find that high-stakes testing is related to cortisol responses, and those responses are related to test performance. Those who responded most strongly, with either increases or decreases in cortisol, scored 0.40 standard deviations lower than expected on the high-stakes exam” (Heissel, Adam, Doleac, Figlio, Meer, 2021). The constant natural disasters cause unrelenting trauma which is a significant reason for Puerto Rican students’ low academic scores.

## **5.2 Different Trauma Causes**

Students have experienced trauma through many different things they witnessed during the storms. In a study titled *Mental Health of Puerto Ricans Living On and Off the Island After Hurricane Maria*, the authors Gerardo Fernandez present the results that state, “The study revealed that 83.9% [of Puerto Rican students] experienced seeing housing destruction, 57% had friends or family members leave the island, 45.7% reported damages to their home, 32.3% experienced shortage of food and water...” (Fernandez & Baum, 2022). There were other traumatic events mentioned in the study showing how a high percentage of students have experienced an event that could induce trauma in them.

## **5.3 Closing of Schools**

The unrelenting disasters have led to the closing of many schools on the island. Some schools were closed temporarily due to power outages and other infrastructure issues. Temporary closings led to an average of 78 missed school days for students of Puerto Rico. Missing school leads to many adverse effects including falling behind in reading development. This can continue to get worse and worse if the students do not learn what they missed. Other schools were closed permanently because of disasters. Hurricane Maria led to the permanent closing of over 250 schools in Puerto Rico. In *The Disappearing Schools of Puerto Rico*, Jonathan Katz details how Puerto Rico was managing a decreasing population following Hurricane Maria writing, “Of the roughly 1,100 public schools left in Puerto Rico at the time of the storms, more than 250 simply didn’t open again “ (Katz, 2019). This closing of schools has led to students not having a close or convenient place to learn. Many students must travel significantly longer to get to the closest school near them to learn. As schools continuously open and close due to poor infrastructure and repeated natural disasters, Puerto Rican students will continue to struggle academically compared to US continental schools.



## Works Cited

- Cardoza, Kavitha. "Are the Challenges of Puerto Rico's Schools a Taste of What Other Districts Will Face?" *The Hechinger Report*, 3 Apr. 2023, [hechingerreport.org/are-the-challenges-of-puerto-ricos-schools-a-taste-of-what-other-districts-will-face/#:~:text=Puerto%20Rico's%20student%20population%20has,steady%20decline%20since%20Hurricane%20Maria](https://hechingerreport.org/are-the-challenges-of-puerto-ricos-schools-a-taste-of-what-other-districts-will-face/#:~:text=Puerto%20Rico's%20student%20population%20has,steady%20decline%20since%20Hurricane%20Maria). Accessed 20 June 2024.
- . "In the 6th-largest U.S. District, Natural Disasters Have Disrupted Schooling for Years." *NPR*, 17 Aug. 2023, [www.npr.org/2023/08/16/1193722562/puerto-rico-schools-education](https://www.npr.org/2023/08/16/1193722562/puerto-rico-schools-education). Accessed 14 July 2024.
- Carney, Kristen. "Richest Puerto Rico Zip Codes (2024)." *Income by Zipcode*, 18 July 2024, [www.incomebyzipcode.com/puertorico/#:~:text=in%20the%20US.,Puerto%20Rico%20Income%20Statistics,make%20over%20\\$200%2C000%20a%20year](https://www.incomebyzipcode.com/puertorico/#:~:text=in%20the%20US.,Puerto%20Rico%20Income%20Statistics,make%20over%20$200%2C000%20a%20year). Accessed 8 Sept. 2024.
- "¿Cual Fue la Tasa de Graduación?" ["What was the Graduation Rate?"]. *Perfil Escolar*, [perfilescolar.dde.pr/dashboard/crosstab/graduationrate.html](https://perfilescolar.dde.pr/dashboard/crosstab/graduationrate.html). Accessed 9 Sept. 2024.
- Diaz, Johnny. "Fiona and Ian Are Retired as Hurricane Names." *The New York Times*, 30 Mar. 2023, [www.nytimes.com/2023/03/30/us/ian-fiona-hurricane-names-retired.html?searchResultPosition=2](https://www.nytimes.com/2023/03/30/us/ian-fiona-hurricane-names-retired.html?searchResultPosition=2). Accessed 7 Mar. 2024.
- "Education in Puerto Rico." *PR51st, Puerto Rican Statehood Council*, 22 Aug. 2022, [www.pr51st.com/education-in-puerto-rico/](https://www.pr51st.com/education-in-puerto-rico/). Accessed 16 May 2024.
- "The facts: Hurricane Maria's effect on Puerto Rico." *Mercy Corps*, 9 Sept. 2020, [www.mercycorps.org/blog/facts-hurricane-maria-puerto-rico/#:~:text=Electricity%20was%20cut%20off%20to,into%20a%20desperate%20humanitarian%20crisis](https://www.mercycorps.org/blog/facts-hurricane-maria-puerto-rico/#:~:text=Electricity%20was%20cut%20off%20to,into%20a%20desperate%20humanitarian%20crisis). Accessed 7 Sept. 2024.
- Fernandez, Gerardo, and Neil Baum. "Mental Health of Puerto Ricans Living On and Off the Island After Hurricane Maria." *The Journal of Medicinal Practice Management*, 8 Aug. 2022, pp. 44-48, [www.physicianleaders.org/articles/mental-health-puerto-ricans-living-off-island-after-hurricane-maria](https://www.physicianleaders.org/articles/mental-health-puerto-ricans-living-off-island-after-hurricane-maria). Accessed 15 Aug. 2024.
- Guth, Lorraine J., et al. "Perceived Impact of Hurricane Maria on Educators and Students in Puerto Rico." *Centro Journal*, vol. 33, no. 3, 2021, pp. 65-98. *Research Library Prep*, [www.proquest.com/scholarly-journals/perceived-impact-hurricane-maria-on-educators/docview/2618441026/se-2?accountid=41092](https://www.proquest.com/scholarly-journals/perceived-impact-hurricane-maria-on-educators/docview/2618441026/se-2?accountid=41092). Accessed 24 July 2024.
- Heissel, Jennifer A., et al. "Testing, Stress, and Performance: How Students Respond Physiologically to High-Stakes Testing." *Education Finance and Policy*, vol. 16, no. 2, 19 Apr. 2021, [https://doi.org/10.1162/edfp\\_a\\_00306](https://doi.org/10.1162/edfp_a_00306). Accessed 16 July 2024.
- "How Trauma Impacts School Performance." *Mental Health America*, [mhanational.org/how-trauma-impacts-school-performance/#:~:text=Social%20Demotional%20Development%3A%20Experiencing%20trauma,rates%20of%20suspension%20and%20expulsion](https://mhanational.org/how-trauma-impacts-school-performance/#:~:text=Social%20Demotional%20Development%3A%20Experiencing%20trauma,rates%20of%20suspension%20and%20expulsion). Accessed 16 July 2024.
- Katz, Jonathan M., and Diana Zeyneb Alhindawi. "The Disappearing Schools of Puerto Rico." *The New York Times Magazine*, 12 Sept. 2019,

- [www.nytimes.com/interactive/2019/09/12/magazine/puerto-rico-schools-hurricane-maria.html](http://www.nytimes.com/interactive/2019/09/12/magazine/puerto-rico-schools-hurricane-maria.html). Accessed 15 July 2024.
- Nelson, Christopher, et al. "The Education Sector in Puerto Rico After Hurricane Maria: Predisaster Conditions, Hurricane Damage, and Themes for Recovery." RAND, RAND Corporation, 30 Sept. 2020, [www.rand.org/content/dam/rand/pubs/research\\_reports/RR2800/RR2858/RAND\\_RR2858.pdf](http://www.rand.org/content/dam/rand/pubs/research_reports/RR2800/RR2858/RAND_RR2858.pdf). Accessed 17 July 2024.
- Orozco, Samantha, and John Whitlock. "Educational Challenges in Puerto Rico." Broken Chalk, 18 Nov. 2023, [brokenchalk.org/educational-challenges-in-puerto-rico/#:~:text=High%20school%20drop%20out%20rates,current%20percentage%20in%20mainland%20US](http://brokenchalk.org/educational-challenges-in-puerto-rico/#:~:text=High%20school%20drop%20out%20rates,current%20percentage%20in%20mainland%20US). Accessed 14 July 2024.
- Pérez Sánchez, Laura N. "May Day Protests in Puerto Rico Show an Economy Still on the Brink." The New York Times, 9 May 2023, [www.nytimes.com/2023/05/09/us/may-day-puerto-rico-protest.html?searchResultPosition=1](http://www.nytimes.com/2023/05/09/us/may-day-puerto-rico-protest.html?searchResultPosition=1). Accessed 7 Mar. 2024.
- "Poverty and the Puerto Rican Family." DISCovering Multicultural America: African Americans, Hispanic Americans, Asian Americans, Native Americans, Gale, 2003. Gale in Context: High School, [link.gale.com/apps/doc/EJ2116207156/SUIC?u=win5026&sid=bookmark-SUIC&xid=96495a01](http://link.gale.com/apps/doc/EJ2116207156/SUIC?u=win5026&sid=bookmark-SUIC&xid=96495a01). Accessed 7 Mar. 2024.
- "Puerto Rico." League of United Latin American Citizens, May 2023, [lulac.org/puertorico/](http://lulac.org/puertorico/). Accessed 17 Aug. 2024.
- "Puerto Rico Earthquakes." Center for Disaster Philanthropy, 3 Dec. 2020, [disasterphilanthropy.org/disasters/puerto-rico-earthquakes/#:~:text=Who%20was%20affected%20and%20in,in%20the%20continuing%20quake%20activity](http://disasterphilanthropy.org/disasters/puerto-rico-earthquakes/#:~:text=Who%20was%20affected%20and%20in,in%20the%20continuing%20quake%20activity). Accessed 10 Sept. 2024.
- Trines, Stefan, editor. "Economic Storm: The Crisis of Education in Puerto Rico." World Education News + Reviews, 8 May 2018, [wenr.wes.org/2018/05/economic-storm-the-crisis-of-education-in-puerto-rico](http://wenr.wes.org/2018/05/economic-storm-the-crisis-of-education-in-puerto-rico). Accessed 20 June 2024.
- Ujifusa, Andrew. "Average Puerto Rican Student Missed 78 Days of School after Maria, Study Finds." EducationWeek, 25 Sept. 2018, [www.edweek.org/education/average-puerto-rican-student-missed-78-days-of-school-after-maria-study-finds/2018/09/#:~:text=The%20calculation%20that%20the%20average,about%20its%20report%20on%20Tuesday](http://www.edweek.org/education/average-puerto-rican-student-missed-78-days-of-school-after-maria-study-finds/2018/09/#:~:text=The%20calculation%20that%20the%20average,about%20its%20report%20on%20Tuesday). Accessed 15 July 2024.
- "Why Is School Attendance Important? The Effects of Chronic Absenteeism." School of Education American University Washington, DC, 14 Jan. 2021, [soeonline.american.edu/blog/importance-of-school-attendance/#:~:text=When%20children%20are%20absent%20from,effect%20that%20impacts%20future%20learning](http://soeonline.american.edu/blog/importance-of-school-attendance/#:~:text=When%20children%20are%20absent%20from,effect%20that%20impacts%20future%20learning). Accessed 6 Sept. 2024.

# **An Exploration of Gender Disparity in Business Classes at Public High School**

## **By Festina Hyseni**

### **Introduction**

Gender inequality is a crucial issue that's prevalent in today's business world. According to Sarah Thébaud, a professor of sociology at the University of California, "gender is understood to be a status characteristic, a categorical distinction based on either a personal attribute (e.g., gender, race) or a role that has attached to it widely shared cultural beliefs about the status worthiness of one category over the other" (Thébaud, 2015, p. 4). Consequently, the roles women and men play in the business world are culturally rooted in conventionalized perceptions that are correlated to gender inequality. Stereotypically, women tend to be underrepresented in aspects of the business field due to a perceived male culture. This is highlighted through Thébaud's study that expresses "US respondents rated men higher than women on a scale that included perceptions of competence, intelligence, confidence, competitiveness, and independence" (Thébaud, 2015, p. 2). In this study, perceptions of gender disparity within the business world, and its contributing factors are analyzed within the enrollment of Public High School's business classes. Gender disparity, for the purpose of this study, is defined as the unequal treatment of an individual based on their gender (male or female). Therefore, the purpose of this project is to use this data to investigate: *To what extent do students' perception of gender disparity exist in enrollment of business classes at Public High School, and what factors contribute to these disparities?*

### **Literature Review**

A review of related literature on gender inequality in the business world is necessary to understand why this paper focuses on the perceptions and opinions of the students. Moreover, it is further necessary to understand why this paper focuses on the factors that contribute to the gender disparities. According to Jennifer Ann Ball, doctor of philosophy, women's interest in attaining an education has increased throughout the years due to the consideration of going into nontraditional fields; however, "they still are underrepresented in science, math, and engineering. A significant reason for this underrepresentation is lower expected returns to these majors for women" (Ball, 2012, p.9). Ball highlights the persisting underrepresentation of women in the business field even when there's an increasing amount of females that have shown interest in attaining an education to further prove the existential gender inequality today. James Morrison, degree in education, expresses "there is a lack of female role models in business schools and also a corresponding lack of opportunity for students to work with women professors" (Morrison, 2015, p. 4). Even within the educational realm of business, gender inequality persistently exists as women aren't represented in positions that allow them to educate and inspire a multitude of women that follow after them in the field of business. In the essence of education, Laura Marini Davis, bachelors in psychology, highlights how behaviors of "sexist use of language, presentation of stereotypic views of women, and instructors favoring male students may dissuade

women from pursuing business classes, and perpetuate the problem of sexism that has long plagued the corporate world” (Davis, 2015, p. 6). By depicting various sexist behaviors, Davis presents one of many causes that contribute to the gender inequality experienced first hand by women in the business world; however, mandates to promote gender equality have been executed as “ the Dodd-Frank Wall Street Reform and Consumer Protection Act called for the establishment of an Office of Minority and Women Inclusion in numerous federal agencies (Davis, 2015, p. 8). This act was implemented based on the belief that greater diversity would create a more effective business; therefore, women were increasingly included in all activities of the agency. Furthermore, the wage gap by gender in graduates has been systematically depicted within the business field. “Most of the gap in these fields can be explained by the different types of jobs women and men in these areas take after college” (Ball, 2012, p.11). Supporting that women are automatically underrepresented once they’ve graduated and are aiming to pursue a career. A contributing factor to the underrepresentation of women in the business realm is the culturally, stereotypical rooted ideas that perceive men better suited for business careers. For instance, many people “strongly associate entrepreneurship with “masculinity” and stereotypically masculine traits” (Thébaud, 2015, p. 5). This connects to Davis’ study as mentioned in which sexist behaviors were analyzed and proven to contribute to gender disparity within the educational aspect of business, and Thébaud’s study highlights how sexist ideas regarding entrepreneurship are instilled in the perceptions of people today. Additionally, “women entrepreneurs will receive lower ability and effort ratings, and their businesses will be rated less worthy of support than men’s” (Thébaud, 2015, p. 8). By presenting the systematic gender inequality, it depicts that women are oppressed in the business world. Another contributing factor to these gender disparities is the lack of discussion of female inclusivity in the business realm. For example, “gender inequality is an issue that continues to be understudied as a topic of importance in both the workplace and in business programs at universities around the globe” (Morrison, 2015, p. 7). Since the underrepresentation of women in the business field isn’t spoken about in educational or work settings, there continues to be a lack of awareness and harm being done to the corporate world in which female talent is systematically being overlooked. Moreover, women businesses continue to be regarded as less important than male businesses as expressed by Katie Abouzahr, a medical doctor, “when women business owners pitch their ideas to investors for early-stage capital, they receive significantly less—a disparity that averages more than \$1 million—than men” (Abouzhar, 2021, p. 9). However, studies have shown that women businesses thrive persistently more than male start up businesses as “businesses founded by women ultimately deliver higher revenue—more than twice as much per dollar invested” (Abouzhar, 2021, p. 9). Women’s potential continues to be underrepresented due to the sexist and unequal gender gaps that are present in today’s educational, social, and work settings in the business world. Many of these studies have proven gender disparities within the business realm; however, extensive research on the perceptions, attitudes, and opinions of specifically students within the educational world of business has not been conducted. The researcher will therefore contribute to the literature on gender inequality within the business world by analyzing the

perceptions of gender disparity within enrollment of business classes at Public High School, and the factors that contribute to these disparities.

## **Gap**

Furthermore, many studies around general underrepresentation of specifically, adult women in the business world have been conducted and thoroughly researched; however, there hasn't been extensive research on students' attitudes or perceptions of gender inequality within the business realm. For instance, in Irene Padavic's study, *What's Really Holding Women Back*, she delves into the gender gap of adult women in the business realm expressing that "High-level jobs require extremely long hours, women's devotion to family makes it impossible for them to put in those hours, and their careers suffer as a result" (Padavic, 2020, p. 6). The gap that hasn't been extensively addressed is specifically student attitudes, experiences, and descriptions with their business classes, why they did or did not enroll in a business class, and how this may or may not lead to a finding of gender disparity. By specifically analyzing the students of Public High School, new research is able to be conducted on the perspectives of students with their business classes and the possible finding of a gender gap through the stated research question.

## **Methodology**

### **Overview**

#### ***Study Design and Method Type***

To answer the research topic, the researcher decided to use a mixed methodologies approach. This necessitated mixing quantitative data—which is based on numbers—with qualitative data, which is not based on numbers. The exploration technique helps the researcher to fill in the gaps by guaranteeing that the investigation will produce data that will allow the researcher to make new conclusions. Moreover, the mixed methods experimental design aids the researcher in analyzing both qualitative and quantitative data that can lead to the finding of gender disparity within the enrollment of business classes at Public High School. Additionally, the use of a mixed methods approach rather than a strictly quantitative one can attain more insight from the data results, while only a qualitative method would enable the researcher to more convincingly defend the findings they came to during their investigation.

#### ***Variables***

The variables of the survey are qualitative and quantitative data sets. Quantitative data is considered to be information of gender, enrollment of business classes, number of business classes taken, interest in taking a business class, and attitudes and encouragement towards the business class or business world. Qualitative data is considered to be the perceptions of gender disparity, factors contributing to disparities, personal opinions and/or experiences, and recommendations for change.

## ***Tools***

A mixed methods approach was implemented through a survey questionnaire. The survey divided the participants into two target groups: general Public High School students and students enrolled in business classes at Public High School. The survey division will allow for the researcher to analyze both qualitative and quantitative factors that shaped enrollment decisions and perceptions of gender disparity within each group- allowing for a comparative analysis of the differing insights, attitudes, and perceptions of general Public High School students and business-enrolled Public High School students that ultimately allows the researcher to conclude if there is a proven gender disparity.

Each target group was required to fill out a consent form that detailed to participants that participation was voluntary and that their information would be kept confidential. They were also asked for age, gender, and if they are enrolled in a business class at Public High School. Once this section of the survey was complete, the survey split into two segments depending on whether the participant is enrolled in a business class at Public High School or not. The content of the questionnaire differed for each target group. Students not enrolled in business classes were asked about their interest in business, reasons for not enrolling in offered business classes, exposure to the content of business classes, encouragement to study in the field of business, their perception of gender inequality within the business world, and depending on whether they believe gender disparity is present, what could be done to solve the issue. On the other hand, students enrolled in a business class were asked the specific class they're taking, their attitude regarding their business class, discussions of female inclusivity in the business field within the classroom, their perception of gender equality in the business world, their reasoning for taking a business class, and depending on whether they believe gender disparity is present, what could be done to solve the issue. While the free-response questions were left as optional, the researcher's multiple-choice survey questions were all mandatory. The researcher examined and contrasted each response in terms of enrollment for this study.

The survey was created through Google Forms and was sent out and completed by students at Public High School High School. These students were introduced to the survey through advertisement of flyers around the school and through the use of social media platforms such as Instagram.

## **Participants**

Before they could participate in the study, the volunteers had certain obligations to meet. The prerequisites included being a student in a Public high school, being under the age of eighteen, and obtaining parental approval.

## **Measures**

### ***Scale***

Quantitative data was interpreted through a five point Likert scale used to measure the extent of disagreement and agreement of the questions and statements asked for general Public

High School students and business enrolled Public High School students. Furthermore, the following table lists the scales that correlate with different questions and statements:

**Figure 2**

***Quantitative Data Scaling: General Public High School Students***

Question/Statement	Scale
To what extent do you feel that the business field is generally male-dominated?	1 - strongly disagree 2 - disagree 3- neutral 4- somewhat agree 5 -strongly agree

***Quantitative Data Scaling: Business enrolled Public High School Students***

Question	Scale
Please rate how you feel in your business class(es)	1 - strongly disagree 2 - disagree 3- neutral 4- somewhat agree 5 -strongly agree
To what extent have there been discussions of increasing female inclusivity in the business industry during your business class(es)?	1 - strongly disagree 2 - disagree 3- neutral 4- somewhat agree 5 -strongly agree
To what extent do you perceive that there's a lack of female representation in your business class(es)?	1 - strongly disagree 2 - disagree 3- neutral 4- somewhat agree 5 -strongly agree
Please rate the overall gender diversity in your business class(es).	1 - strongly disagree 2 - disagree 3- neutral 4- somewhat agree 5 -strongly agree
To what extent do you feel that the business field is generally male-dominated?	1 - strongly disagree 2 - disagree 3- neutral 4- somewhat agree 5 -strongly agree

Regarding the last question, if your answer was yes, how much of an influence did the exposure of Public High School's business courses have towards your decision in enrolling in a class?	1 - strongly disagree 2 - disagree 3- neutral 4- somewhat agree 5 -strongly agree
---	---

### ***Equations***

The two equations given below will be used by the researcher to examine the study's quantitative data. The mean, or average score on the Likert scale, will be determined by multiplying each number on the scale (1–5) by the proportion of replies for that number separately, as shown in the equation below.

Mean = (1 X Percentage of respondents for 1) + (2 X Percentage of respondents for 2) + (3 X Percentage of respondents for 3) + (4 X Percentage of respondents for 4) + (5 X Percentage of respondents for 5)

Multiplying the response option number by its percentage allows for the researcher to interpret data that presents the distribution of responses among different options. Ultimately, the equation is able to reflect the overall perception of respondents while taking into account the frequency of each response.

Furthermore, in order to analyze how far a given response deviates from the average, the researcher will also calculate the standard deviation, which is how dispersed the data is in relation to the mean. Therefore, the following equation will be used to this study's aspect:

Standard Deviation =  $\sqrt{(1 - \text{Mean})^2 \times \text{Percentage of respondents for 1} + (2 - \text{Mean})^2 \times \text{Percentage of respondents for 2} + (3 - \text{Mean})^2 \times \text{Percentage of respondents for 3} + (4 - \text{Mean})^2 \times \text{Percentage of respondents for 4} + (5 - \text{Mean})^2 \times \text{Percentage of respondents for 5}}$

The standard deviation highlights how much respondents' opinions or attitudes differ from the average response (mean); therefore, by squaring the differences of each response option and the mean, negative and positive deviations are taken into account to measure the overall variability of the question/statement asked. A higher standard deviation would mean that responses are more different from each other while a lower standard deviation implies that most students agree with the question or statement that's asked.

The mean and standard deviation that is calculated for each question or statement depicted in *figure 2* will effectively interpret the overall answer/agreement to the statements and ultimately lead the researcher to determine the finding of gender disparity.

### ***Short Answer Responses***

Qualitative data was interpreted through long answer free answer responses that were used to highlight a specific theme to allow the researcher to holistically code the short answers that were provided. Each long answer free response question was optional; therefore, it was a mere extension to the required, multiple choice question that was asked beforehand for the purpose of gaining a greater insight to the participant's answer. Furthermore, the following free



response questions/statements that were asked to general Public High School students and business enrolled Public High School students are shown below:

**Figure 3**

<b>General Public High School Students</b>	<b>Business Enrolled Public High School Students</b>
Following the last question, please describe the answer(s) that you selected regarding your decision to not enroll in business class(es) at Public High School.	Please describe why you feel that way in your participation in a business class(es).
Have you ever been encouraged by a family member or friend to study business? If your answer was yes, please describe the advice/feedback that was given to you and your feelings or concerns regarding it.	To what extent have there been discussions of increasing female inclusivity in the business industry during your business class(es)? Please describe the discussion that occurred.
Have you ever been exposed to materials that introduced/described Public High School’s business courses? If your answer was yes, what aspect(s) of the materials did not or did appeal to you?	Following the last question, please describe why you chose your rating regarding a male-dominated business environment.
To what extent do you feel that the business field is generally male-dominated? Please describe why you chose your rating regarding a male-dominated business environment.	What motivated you to enroll in a business class at Public High School? Please describe the answer(s) that you selected regarding your motivation for enrolling in business class(es) at Public High School.
If you feel that there is a perceived male-dominance and/or underrepresentation of women in the business classes of Public High School, what improvements do you feel could best help resolve this issue?	If you feel that there is a perceived male-dominance and/or underrepresentation of women in the business classes of Public High School, what improvements do you feel could best help resolve this issue?

To interpret the short answer responses, thematic analysis of holistic coding will be utilized. This allows for the researcher to examine patterns in the participants’ answers by categorizing the long answer free responses into several groups. The process is executed in which the researcher 1) familiarizes themselves with the data, 2) identifies codes (words/phrases) within each response, 3) assigns the initial codes to each response depending on the overall content, and 4) organizes the codes into broader themes that depict recurring patterns across responses.

## Results

### *Data Overview*

Over the course of two months, 61 students took part in the study's survey. There were 18 male and 43 female responders. Of these respondents, 14 were business enrolled Public High School students and 47 were general Public High School students.

### *Descriptive Statistical Analyzation*

The study's quantitative data was analyzed through descriptive statistical analysis. The researcher used the mean and standard deviation to examine the key features of the data set in order to prove or disprove the finding of a gender disparity. The following calculations were computed:

#### *General Public High School Students*

##### *Figure 4*

Question 1: To what extent do you feel that the business field is generally male-dominated?

<b>Mean</b>	3.87
<b>Standard Deviation</b>	1.04

*Figure 4* shows the results of question one for general Public High School students. The mean for the answer was 3.87 and the standard deviation was 1.04.

#### *Business Enrolled Public High School Students*

##### *Figure 5*

Question 1: Please rate how you feel in your business class(es).

<b>Mean</b>	3.57
<b>Standard Deviation</b>	1.29

*Figure 5* shows the results of question one for business enrolled Public High School students. The mean for the answer was 3.57 and the standard deviation was 1.29.

##### *Figure 6*

Question 2: To what extent have there been discussions of increasing female inclusivity in the business industry during your business class(es)?

<b>Mean</b>	1.93
<b>Standard Deviation</b>	.92

Figure 6 shows the results of question two for business enrolled Public High School students. The mean for the answer was 1.93 and the standard deviation was .92.

**Figure 7**

Question 3: To what extent do you perceive that there's a lack of female representation in your business class(es)?

<b>Mean</b>	2.57
<b>Standard Deviation</b>	1.21

Figure 7 shows the results of question three for business enrolled Public High School students. The mean for the answer was 2.57 and the standard deviation was 1.21.

**Figure 8**

Question 4: Please rate the overall gender diversity in your business class(es).

<b>Mean</b>	3.71
<b>Standard Deviation</b>	1.17

Figure 8 shows the results of question four for business enrolled Public High School students. The mean for the answer was 3.71 and the standard deviation was 1.17.

**Figure 9**

Question 5: To what extent do you feel that the business field is generally male-dominated?

<b>Mean</b>	3.57
<b>Standard Deviation</b>	1.37

Figure 9 shows the results of question five for business enrolled Public High School students. The mean for the answer was 3.57 and the standard deviation was 1.37.

**Figure 10**

Question 6: Regarding the last question, if your answer was yes, how much of an influence did the exposure of Public High School’s business courses have towards your decision in enrolling in a class?

<b>Mean</b>	2.57
<b>Standard Deviation</b>	1.37

Figure 10 shows the results of question six for business enrolled Public High School students. The mean for the answer was 2.57 and the standard deviation was 1.37.

**Thematic Analysis**

The study’s qualitative data was analyzed using holistic coding of thematic analysis in which long answer free responses were categorized into different groups in order to examine patterns in the participants' answers. The following tables illustrate the codings and themes that were analyzed from participants’ answers.

*General Public High School Students*

<b>Themes:</b>	<b>Codings:</b>
1. Disinterest with Business	- “Lack of interest in business”
2. Lack of Knowledge	-“Lack of knowledge about business”
3. Influence of Career Aspirations	-“I am interested in a different career path” - “Influenced by family and friends”
4. Perception of Male Dominance	- “I believe the classes are extremely male dominant” -“Business is a male-dominated field”

*Business Enrolled Public High School Students*

<b>Themes:</b>	<b>Codings:</b>
1. Gender Inclusivity Discussions	- “Gender-neutral topics are talked about” -“Women can do that at their own pace”
2. Positive Class Environment	-“The class was laid back” -“I am comfortable with the work” -“I have trouble understanding”

3. Motivation for Enrolling in a Business Class	-“I’m interested in the business industry in the future”
4. Perception of Male Dominance	<ul style="list-style-type: none"> <li>- “There are more businessmen than women in real-world situations”</li> <li>- “Higher paying jobs are mostly male”</li> <li>-“People believe males are better than females”</li> </ul>

### Analysis & Discussion

The aim of this research was to understand the extent to which students’ perceptions of gender disparity exist within enrollment of business classes at Public High School, and what factors contribute to the disparities. Furthermore, the literature review's examples of women's gender inequality in the business world, the scales shown in *Figure 2*, the short answer responses shown in *Figure 3*, and the scales themselves were used to find out how much students believed that there was a gender gap in the enrollment of business classes at Public High School, as well as what factors contributed to the gaps.

In regards to the first target group of general Public High School students being asked the extent they felt that the business field is generally male dominated on a likert scale of 1-5, the calculated mean response was 3.87 and the standard deviation was 1.04. This demonstrates that respondents typically feel that men predominate in the economic sector, even though individual replies vary greatly from this average.

The second target group of business enrolled students at Public High School were asked a series of 6 quantitative questions on a likert scale of 1-5. The first one asked to rate how the participants felt in their business classes. The calculated mean was 3.57 and the standard deviation was 1.29. The average response implies that, overall, students' experiences in business studies have been passably positive. Conversely, the very large standard deviation values indicate student variability.

The second question analyzed the extent of discussion of increasing female inclusivity in the business industry during class. The calculated mean was 1.93 and the standard deviation was .92. The low standard deviation implies that students' responses are generally consistent, and the low mean score says that students believe their classes don't cover much on how to make the corporate environment more inclusive of women.

Moreover, the third question analyzed the extent to which the participant perceived there to be a lack of female representation in their business class. The mean was 2.57 and the standard deviation 1.21. Students believe there is a somewhat big absence of female representation in their business lectures, based on the average response. The standard deviation shows that different students have different perceptions, with some seeing a greater lack of female presence than others.

The fourth question analyzed the overall gender diversity in the participants' business class. The mean being 3.71 and the standard deviation 1.17. Students believe there is a good lot of gender diversity in their business lectures, based on the average response.

Furthermore, the fifth question analyzes the extent the participant felt that the business field was generally male dominated. The mean was 3.57 and the standard deviation 1.37. The average response indicates that students think men make up a significant portion of the business sector's workforce. The standard deviation demonstrates the disparities in students' perceptions of one another.

The sixth question analyzes the influence to which promotion of Public High School's business courses have toward enrolling a business class. The mean being 2.57 and the standard deviation 1.37. The average response shows that Public High School's business courses have a moderate impact on students' decisions to enroll in classes. The relatively high standard deviation indicates that the effect levels of the students differ.

Student views suggest that there may be a gender difference in the enrollment of business classes at Public High Schools based on an analysis of data from both enrolled and unenrolled business students. These disparities are a result of beliefs that men predominate in the corporate sector, a lack of female representation, and a lack of gender diversity. Another reason for these disparities is that there is less discussion in the classroom on encouraging gender diversity in the corporate sector.

Furthermore, in regards to the thematic analysis of long free answer responses, the first target group, general Public High School students', long-answer free-response questions were holistically coded with "Lack of interest in business", "Lack of knowledge about business", and "I believe the classes are extremely male dominated" which examined why the participants didn't enroll in a business class. "Influenced by family and friends" highlighted how some students were encouraged to consider business as a career pathway. "Business is a male-dominated field" analyzed participants' opinions of a male-dominated business environment. The themes that arose from these codes are disinterest with business, lack of knowledge and understanding, influence of career aspirations, and perception of male dominance in business classes and fields.

The second target group, business-enrolled Public High School students' and their codes of "The class was laid back", "I am comfortable with the work", and "I have trouble understanding" analyzed participants' comfort in their business class. "Gender-neutral topics are talked about" and "Women can do that at their own pace" analyzed discussion of increasing female inclusivity in the business industry during class. "Higher paying jobs are mostly male", "There are more businessmen than women in real-world situations", and "People believe males are better than females" analyzed the extent to which participants felt the business world was generally male dominated. "I'm interested in the business industry in the future" analyzed participants' motivation to enroll in a business class at Public High School. The themes that arose from these codes are perception of male dominance in the business field, gender inclusivity discussions, positive class environment, and motivation for enrolling in a business class.

The primary conclusions drawn from this thematic analysis are that participants' perceptions of male dominance in the business sector were evident in both their enrollment and non-registration in business classes. A lack of proactive measures to overcome gender disparities

in enrollment in business programs at Public High School is shown by the absence of clear discussions on gender inclusivity or considerations in the classroom and in the business world.

### **Limitations**

There were limitations on this research endeavor. The danger of volunteer bias is raised by the fact that this study was voluntary. Voluntary bias is when people with strong beliefs tend to participate in research. Results can be skewed by volunteer bias since individuals with strong opinions are more likely to engage in studies and have comparable viewpoints.

Secondly, sample bias is another limitation of this study. Sample bias is when one group is systematically more likely to be selected in a sample than others. The survey may not be fully representative of all high school students or even all students in the area because it only includes a sample of Public High School's students. It's probable that students at this particular high school differ from those at other schools, from other regions, or from other demographic and socioeconomic groups in terms of their experiences and points of view. The study depicted themes and insights unique to the experiences and perspectives of the kids at this particular school and isn't representative of all high school students.

Finally, this study had limited parameters based exclusively on the opinions of the students. While it is vital to get students' perspectives, perceptions may not always be a true reflection of the situation at hand or provide a comprehensive understanding of the factors contributing to gender disparities in enrollment. Students' perceptions may be influenced by a variety of factors, such as personal experiences, cultural beliefs, and biases. As a result, student opinions may not always align with actual enrollment statistics or the ways in which social and systemic factors interact to shape enrollment patterns.

### **Conclusion**

This study focused on two aspects: the extent to which students perceived there to be gender disparity within the enrollment of business classes at Public High School, and the factors that contributed to these disparities. The results of the questionnaire from both general Public High School students and business enrolled Public High School students presented that participants' perceptions of male dominance in the business sector were evident and that these disparities were a result of beliefs that men predominate in the corporate sector, there's a lack of female representation, and a lack of gender diversity being discussed. To further understand the gender inequality that's present in the business world, researchers can conduct this exact study and ensure that it's more representative of students' perceptions by executing it to different high schools due to the sample bias that was discussed in the limitations of this study.

### Works Cited

- Abouzahr, Katie et al., “Why Women-Owned Startups Are a Better Bet,” BCG Global (BCG Global, January 8, 2021), <https://www.bcg.com/Publications/2018/whywomen-owned-startups-are-better-bet>.
- Ball, JenniferAnn. “The Gender Gap in Undergraduate Business Programs in the United States.” *Journal of Education for Business*, vol. 87, no. 5, Sept. 2012, pp. 260–65. *EBSCOhost*, <https://doi.org/10.1080/08832323.2011.617792>.
- Davis, Laura Marini, and Victoria Geyfman. “The Glass Door Remains Closed: Another Look at Gender Inequality in Undergraduate Business Schools.” *Journal of Education for Business*, vol. 90, no. 2, Feb. 2015, pp. 81–88. *EBSCOhost*, <https://doi.org/10.1080/08832323.2014.980715>.
- Morrison, James L. “Integrating Gender Equality Into Business and Management Education.” *Journal of Education for Business*, vol. 90, no. 7, Oct. 2015, pp. 410–12. *EBSCOhost*, <https://doi.org/10.1080/08832323.2015.1081867>.
- Padavic, Irene. “What’s Really Holding Women Back?” *Harvard Business Review*, Harvard Business Review, 19 Feb. 2020, [hbr.org/2020/03/whats-really-holding-women-back](http://hbr.org/2020/03/whats-really-holding-women-back).
- Thébaud, Sarah. “Status Beliefs and the Spirit of Capitalism: Accounting for Gender Biases in Entrepreneurship and Innovation.” *Social Forces*, vol. 94, no. 1, Sept. 2015, pp. 61–86. *EBSCOhost*, <https://doi.org/10.1093/sf/sov042>



# **Low Fertility Rates in Developed Nations: Causes, Consequences, and Policy Solutions**

## **By David Lubin**

### **Abstract**

This paper examines the impact of declining fertility rates in advanced economies and offers policy recommendations for addressing the current trend. Fertility rates in advanced economies are falling even more rapidly than in the rest of the world, and while the full consequences may not be felt for decades, governments must take early action to prevent large-scale economic challenges. Based on current data and economic theory, the analysis reveals a strong link between wealth and birth rates, with childcare costs and other social mechanisms playing a key role. By comparing and contrasting successful programs in Sweden with unsuccessful government-mandated practices in Romania, the following study concludes that pro-natalist schemes must provide benefits to families with children instead of taking benefits away from citizens without children.

### **Introduction**

The United Nations refers to a level of total lifetime fertility below 2.1 children per woman as 'low fertility.' When average fertility rates consistently fall below the so-called 'replacement rate' of 2.1, a country's population is unable to sustain itself at the current level (United Nations 5). There are few exceptions of advanced economies that have birth rates above the replacement threshold; a majority of them fall far below the 2.1 figure. ("Home Data World Population Dashboard"). This trend is especially pronounced in South Korea, where the fertility level has fallen from 5.1 births per woman in 1960 to 0.8 in 2022 (World Bank Group). Similar patterns are seen throughout most OECD countries, which have a combined average of 1.5 births per woman (OECD).

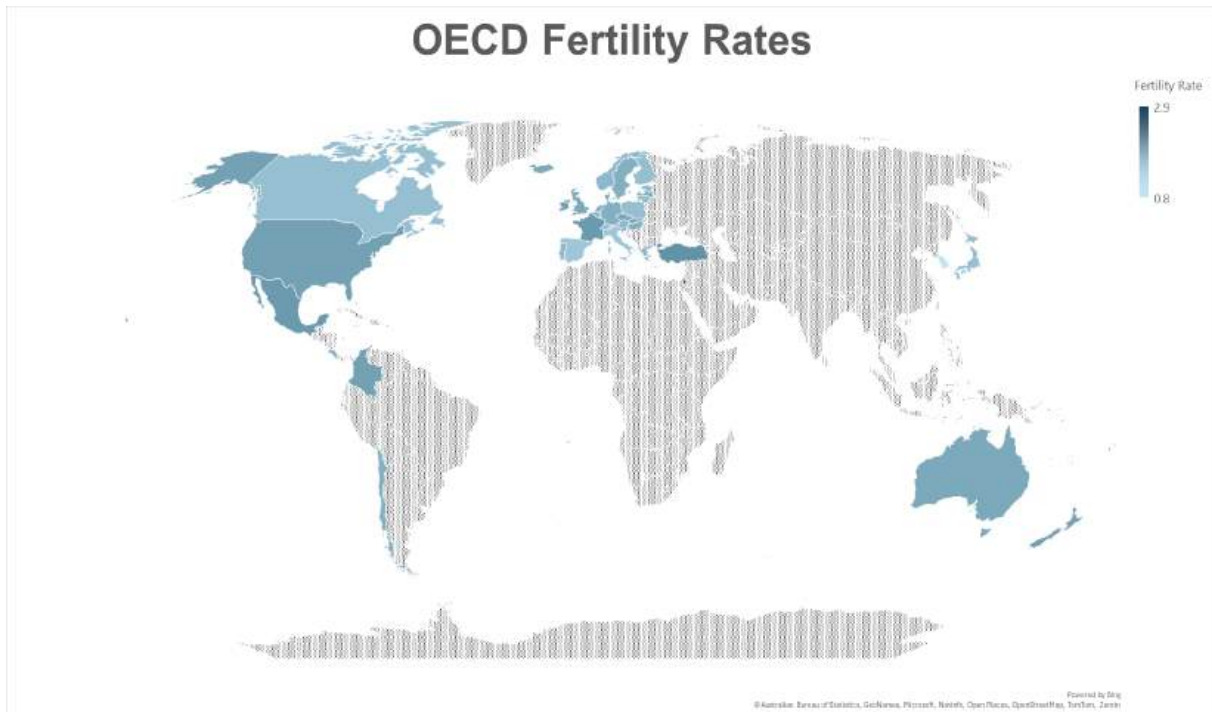


Figure 1: Fertility Rates in OECD Countries (by author). Data sourced from the World Bank Group, *World Bank Databank, World Development Indicators, 2022*.

The sharp reduction in fertility levels within advanced economies has been driven primarily by economic pressures, and presents far-reaching challenges. In order to raise fertility, governments must implement pro-natalist policies and promote state-sponsored programs aimed at creating a supportive environment for procreation and child-rearing.

### **Economic Pressures**

Increasing economic pressures have been the leading cause of declining fertility rates across OECD countries. However, other factors have also contributed to this alarming phenomenon, including structural societal transformations and a variety of governmental policies. This view is supported by social scientists of the likes of Max Roser, who has pointed to rising childcare costs and higher rates of secondary school enrollment and retention amongst young women as the main reasons behind declining birth rates (Roser). In the orthodox economic model, pioneered by Thomas R. Malthus in the 18th century, fertility rates rise as income goes up (Becker et al. 13-14). Since the 18th century, Malthus's theory has largely been disproven, as in the now economically developed parts of the world a significant drop in fertility rates has coincided with substantial income growth starting with the demographic transition that began in the late 1800s (Becker et al. 14).

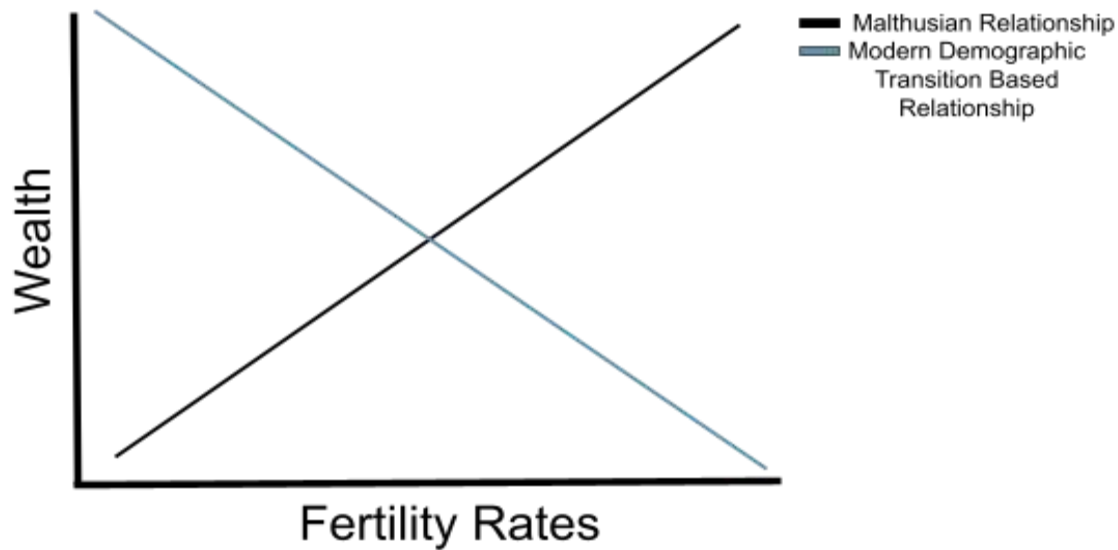


Figure 2: Comparing the Malthusian relationship between fertility rates and income with modern economic theory (by author). *Using the description provided in (Becker et al).*

Malthus’s ideas were superseded by the neoclassical model, a paradigm that incorporates dynamic factors such as wages, unearned income, and education levels to arrive at more accurate demographic estimates (Schultz 599-600). Nobel Prize-winning economist Gary Becker partially agreed with advocates of the neoclassical approach—which directly addressed the failures of the Malthusian model—but thought these explanations to be incomplete, arguing that neoclassical theory focused too heavily on physical capital at the expense of human capital (Becker et al. 13). Becker’s model, by contrast, was grounded in the notion that, with increased investments in education, fertility rates should fall due to the higher costs associated with raising children (Becker et al. 13-15).

### **Economic Conditions**

A shift of focus from theoretical to practical factors supports the argument that reduced birth rates have mainly come from changing economic conditions. Evolutionary biologist Lonnie W. Aarssen’s research into demographic trends within wealthier nations has shown that there exists a strong correlation between lifetime fertility and GDP per capita: when wealth increases, people tend to have fewer children (Aarssen 113).

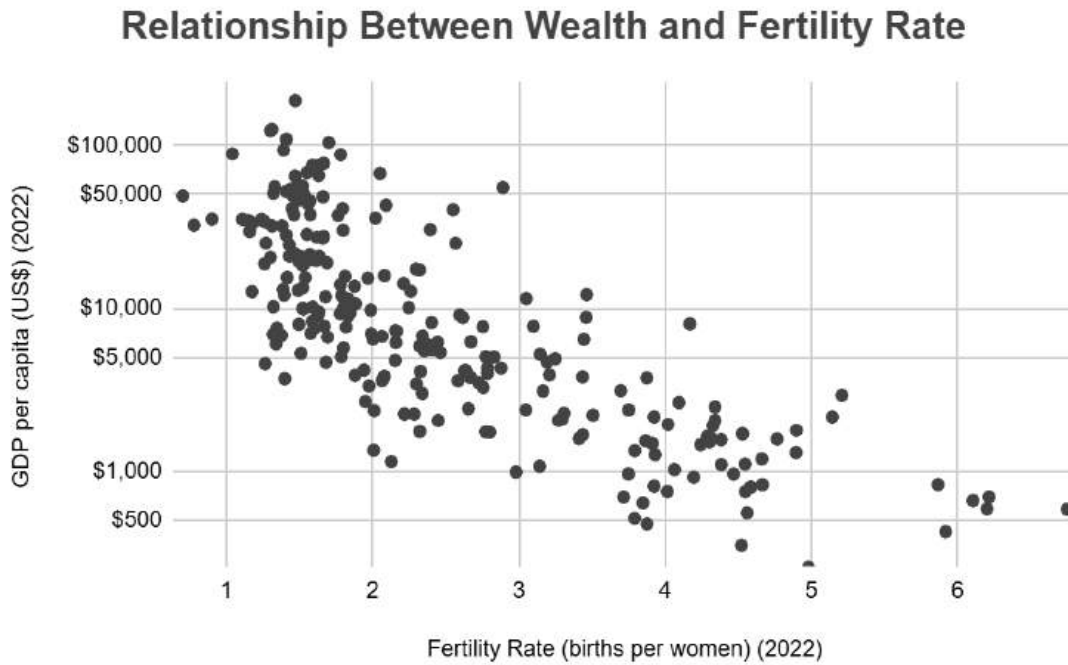


Figure 3: Relationship Between Wealth and Fertility Rate (by author). Data sourced from the World Bank Group, *World Bank Databank, World Development Indicators, 2022*.

Aarssen believes that this relationship should be attributed to a biological phenomenon known as ‘offspring-endowment-selection,’ whereby parental behaviors in advanced economies have promoted larger investments into fewer children (Aarssen 114). Aarssen’s belief that parents’ decision to invest more in their offspring leads to fewer children being born aligns with Becker’s view. The rise in GDP per capita in advanced economies has led to greater access to contraceptives, higher female employment rates, and increased child-bearing costs, all of which have contributed to declining fertility rates over time (Nargund 191). Additionally, anti-natalist governmental policies such as China’s One-Child Policy have played an important role for the past several decades (Leong).

### Downstream Effects

The consequences of low fertility rates range from population decline and aging to a reduction in the labor supply and slower economic growth. Nevertheless, many of the by-products are yet to materialize, as large-scale migration has had a stabilizing influence even as birth rates continue to drop (Bloom et al. 2). Accordingly, the argumentative framework of this paper centers around the changes that have been seen thus far, as well as the likely downstream effects that may arise in years to come.

The share of the elderly in the total population is inversely related to fertility rates, and when birth rates fall, the number of dependents also increases as a percentage of the population, as illustrated in Figure 4. This trend is being further accelerated by innovations in healthcare and

medicine, which have resulted in increased life spans across industrialized nations globally (Georgieva).

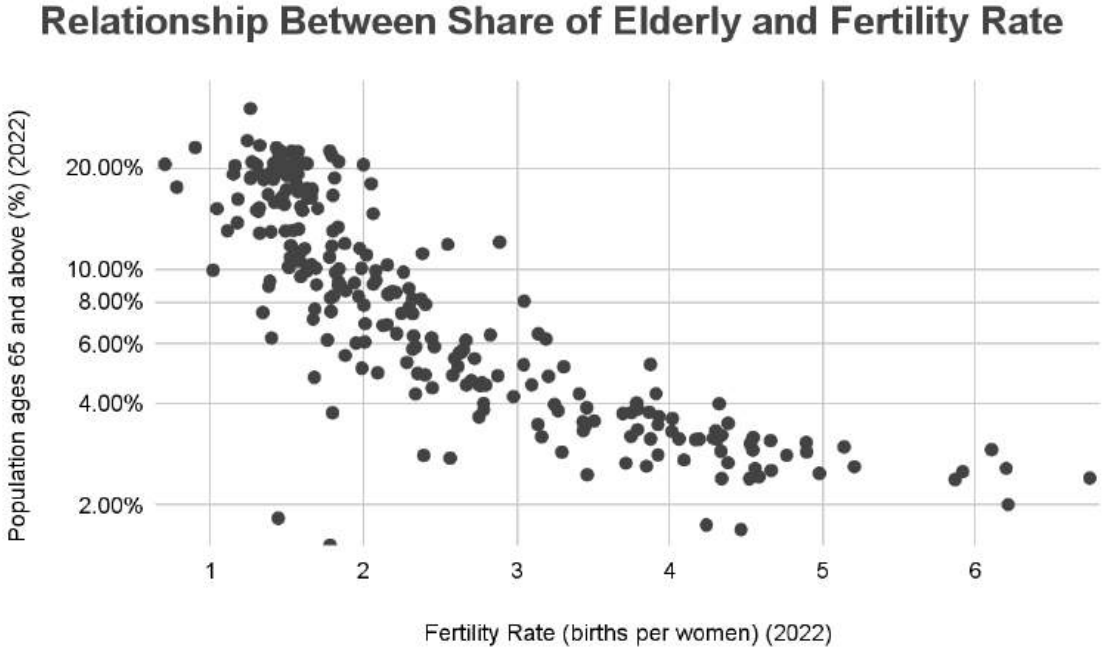


Figure 4: Relationship Between Share of Elderly and Fertility Rate (by author). Data sourced from the World Bank Group, *World Bank Databank*, 2022.

South Korea, a country that currently holds the record for the world’s lowest fertility rate, is projected to have 60% of its population over the age of 65 by 2050 (World Social Report 20). An aging population negatively affects tax revenues via a shrinking workforce and rising pension costs and threatens the viability of social welfare systems (Georgieva). These risks force policymakers to confront the deficit by either lifting the retirement age (as has been seen with the 2023 French pension reform) or increasing taxation (Onrubia & Sánchez 13-14; République Française). In such scenarios, prolonged economic stagnation is a major concern. In fact, American demographer David E. Bloom argues that when countries find themselves below the replacement threshold, economic decline is inevitable in the absence of timely governmental intervention (Bloom et al. 5, 9).

**Examination of Pro-Natalist Policies**

The implementation of state-sponsored pro-natalist programs does help mitigate the risks associated with low fertility rates. Family welfare policies and government incentives that encourage parents to have more children have for the most part delivered mixed results, with some initiatives being successful and others failing spectacularly (Zhang). Extreme pro-natalist policies in Romania during the 1960s, 1970s, and 1980s that ended access to abortions and

contraceptives while placing a tax on unmarried citizens have turned out to be wildly unsuccessful, and have produced multiple unintended consequences (Haupt 3). Specifically, as a result of these measures, Romania saw the highest recorded maternal mortality rate in all of Europe, as thousands of pregnant women turned to illegal abortions (Haupt 4).

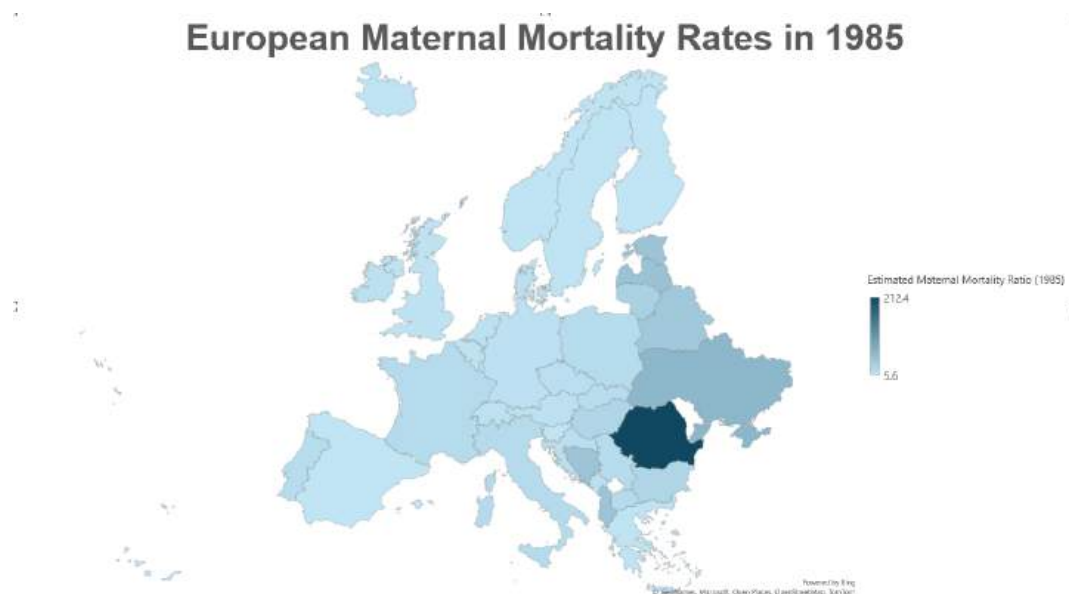


Figure 5: European Maternal Mortality Rates in 1985 (by author). *Data sourced from UN MMEIG (2023) through Our World in Data.*

Romania’s failed experiment with Decree 770 shows how fertility-related policies should be carefully designed to minimize negative impact. On the other end of the spectrum, better-engineered solutions such as the Swedish ‘Speed Premium’ system provide valuable insights into how advanced economies can effectively promote a more supportive environment for families (Miranda). Since the introduction of the speed premium policy in 1980—which allowed parents to retain the parental leave benefits paid after the birth of one child until after the next birth if it arrived within two years—the Swedish birth rate has remained close to constant vis-à-vis other European countries with comparable economic conditions. From 1980 to 1990, Sweden experienced the steepest rise in fertility levels compared to both the European average and other Northern European nations (Miranda).

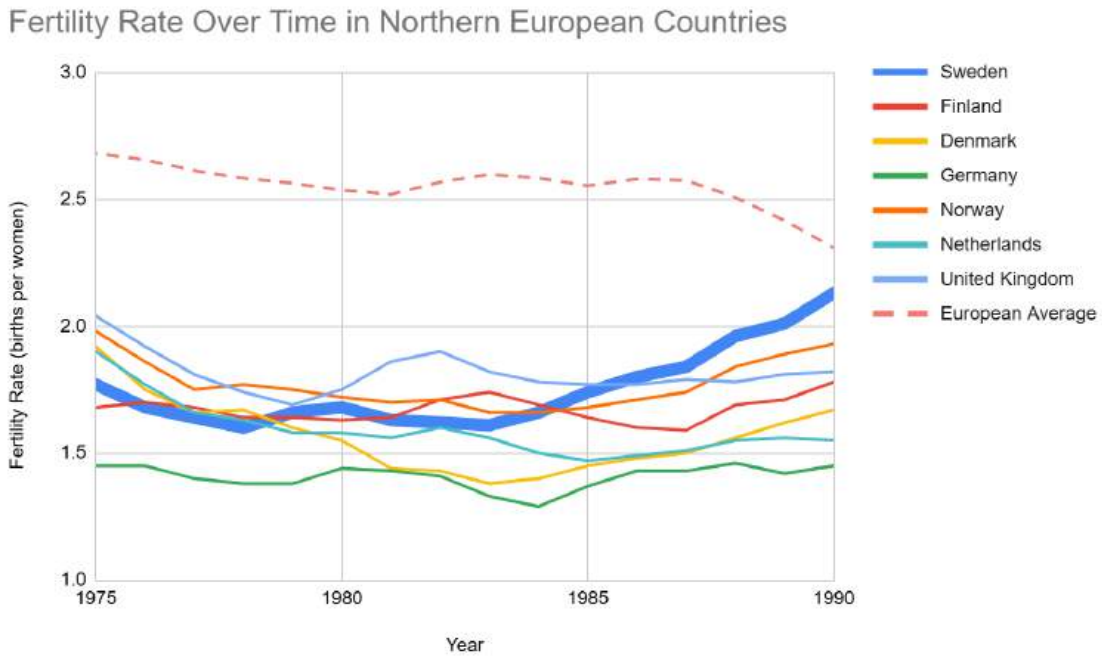


Figure 6: Fertility Rate Over Time in Northern European Countries (by author). Data sourced from the World Bank Group, *World Bank Databank*, 2022

When it comes to state-sponsored fertility-related programs, the key difference between successful and unsuccessful policies is that the former provide benefits to families with children rather than punish those who do not wish to become parents (Frejka 635-36).

## Conclusion

Low fertility rates are driven by economic pressures and social transformations, which modern economic theory attributes, in part, to increases in per capita wealth. This perspective is largely supported by Becker’s research, as well as Aarsen’s findings and data from the World Bank. In the long run, a lower tax base should reduce revenues for the state and cause budget deficits while at the same time increasing demands on social services and placing strain on welfare systems. This paper’s discussion of pro-natalist policy yields several conclusions. The first is that pro-natalist measures do not always produce the intended outcomes, as seen in Romania between 1967 and 1989. Another key finding is that governments interested in boosting birth rates should incentivize parents to have children through benefits instead of penalties, which explains why the Swedish speed premium system was successful and Romania’s Decree 770 produced overwhelmingly negative outcomes.

## Works Cited

- Aarssen, L. W. "Why Is Fertility Lower in Wealthier Countries? The Role of Relaxed Fertility-Selection." *Population and Development Review*, vol. 31, no. 1, 2005, pp. 113–126. JSTOR, <http://www.jstor.org/stable/3401440>.
- Becker, G. S., Murphy, K. M., and Tamura, R. "Human Capital, Fertility, and Economic Growth." *Journal of Political Economy*, vol. 98, no. 5, 1990, pp. S12–S37. JSTOR, <http://www.jstor.org/stable/2937630>.
- Bloom, D., Canning, D., Fink, G., and Finlay, J. "Introduction to Special Issue of the European Journal of Population." *HSPH.Harvard.edu*, 25 Jan. 2010, [https://www.hsph.harvard.edu/pgda/wp-content/uploads/sites/1288/2013/10/PGDA\\_WP\\_54.pdf](https://www.hsph.harvard.edu/pgda/wp-content/uploads/sites/1288/2013/10/PGDA_WP_54.pdf).
- Frejka, T., and Zakharov, S. "The Apparent Failure of Russia's Pronatalist Family Policies." *Population and Development Review*, vol. 39, no. 4, 2013, pp. 635–647. JSTOR, <http://www.jstor.org/stable/23655311>.
- Georgieva, D. "What Do Demographic Changes Mean for Labor Supply?" *World Bank Blogs*, 9 Apr. 2024, <https://blogs.worldbank.org/en/developmenttalk/what-do-demographic-changes-mean-for-labor-supply>.
- Haupt, A. "How Romania Tries to Govern Fertility." *Population Today*, vol. 15, no. 2, 1987, pp. 3-4.
- Leong, P., and Sriramesh, K. "Romancing Singapore: When Yesterday's Success Becomes Today's Challenge." *Public Relations Review*, vol. 32, no. 3, 2006, pp. 246–253, <https://doi.org/10.1016/j.pubrev.2006.05.011>
- Miranda, V. "Recent Trends in Birth Intervals in Sweden: A Decline of the Speed-Premium Effect?" *NCBI*, 8 Oct. 2019, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7363752/>.
- Nargund, G. "Declining Birth Rate in Developed Countries: A Radical Policy Re-think Is Required." *Facts, Views & Vision in ObGyn*, vol. 1, no. 3, 2009, pp. 191–193.
- OECD. "Declining Fertility Rates Put Prosperity of Future Generations at Risk." *OECD*, 20 June 2024, <https://www.oecd.org/en/about/news/press-releases/2024/06/declining-fertility-rates-put-prosperity-of-future-generations-at-risk.html>.
- Onrubia, J., and Sánchez, J. "Taxation Challenges of Population Ageing: Comparative Evidence from the European Union, the United States, and Japan." *Documentos*, Banco de España, 2021, <https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSerias/DocumentosOcasiones/21/Files/do2102e.pdf>.
- République Française. "Raising the Statutory Retirement Age to 64 in France." *Cleiss*, 2 May 2023, <https://www.cleiss.fr/actu/2023/2305-raising-statutory-retirement-age-to-64-france.html>.
- Roser, M. "Fertility Rate." *Our World in Data*, 2024, <https://ourworldindata.org/fertility-rate>.



- Schultz, T. P. "Testing the Neoclassical Model of Family Labor Supply and Fertility." *The Journal of Human Resources*, vol. 25, no. 4, 1990, pp. 599–634, <https://doi.org/10.2307/145669>.
- UN MMEIG (WHO, UNICEF, UNFPA, World Bank Group, and UNDESA/Population Division). *Trends in Maternal Mortality 2020*. Our World in Data, 2023. Dataset, "Estimated Maternal Mortality Ratio in 1985."
- United Nations. *World Population Policies 2021: Policies Related to Fertility*. United Nations, 2021, [https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/undesa\\_pd\\_2021\\_wpp-fertility\\_policies.pdf](https://www.un.org/development/desa/pd/sites/www.un.org.development.desa.pd/files/undesa_pd_2021_wpp-fertility_policies.pdf).
- World Bank Group. "Fertility Rates Korea, Rep." *World Bank Group*, 2022, <https://data.worldbank.org/indicator/SP.DYN.TFRT.IN?end=2022&locations=KR&start=1960&view=chart>.
- World Bank Group. "Fertility Rates" *World Bank Group*, 2022, <https://data.worldbank.org/indicator/SP.DYN.TFRT.IN>.
- World Bank Group. "GDP per Capita" *World Bank Group*, 2022, <https://data.worldbank.org/indicator/NY.GDP.PCAP.CD>.
- World Bank Group. "Population ages 65 and above (% of total population)" *World Bank Group*, 2022, <https://data.worldbank.org/indicator/SP.POP.65UP.TO.ZS>.
- World Social Report. *Leaving No One Behind In An Ageing World*. United Nations, 9 Jan. 2023, <https://www.un.org/development/desa/dspd/wp-content/uploads/sites/22/2023/01/2023wsr-chapter1-.pdf>.
- Zhang, T.-T. "The Effect of Family Fertility Support Policies on Fertility, Their Contribution, and Policy Pathways to Fertility Improvement in OECD Countries." *NCBI*, 8 Mar. 2023, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10049131/>.

## **Cancer in a Global Context: Addressing Risk, Access, and Policy Disparities Between High-Income Countries and Low and Middle-Income Countries By Mali Lord**

### **Abstract:**

This review investigates the global impact of cancer, comparing High-Income Countries (HICs) and Low- and Middle-Income Countries (LMICs), and evaluates how factors such as substance use, environmental risks, and disparities in healthcare access influence cancer outcomes. In high-income countries (HICs), robust public health measures have led to decreases in tobacco and alcohol use, contributing to better cancer control. In contrast, low- and middle-income countries (LMICs) experience rising cancer rates due to weaker regulations and aggressive marketing. Environmental risk factors, such as asbestos exposure and air pollution, impact both HICs and LMICs; however, the responses to these challenges vary significantly. Additionally, disparities in cancer care between rural and urban areas are notable, particularly in LMICs. Rural regions often encounter barriers like cultural misunderstandings and limited resources, negatively affecting cancer outcomes. Effective policy reforms and targeted public health strategies are essential to address these issues. In HICs, efficient pricing and publicly funded healthcare systems enhance treatment affordability, whereas LMICs struggle with high treatment costs and access barriers. This underscores the urgent need for cost-reduction strategies and international aid to improve cancer care in these regions. To mitigate the global cancer burden, a comprehensive approach involving enhanced public health education, expanded healthcare infrastructure, and improved access to affordable treatments is crucial.

### **Introduction**

This study investigates the key factors driving cancer disparities between High-Income Countries (HICs) and Low- and Middle-Income Countries (LMICs). Understanding these disparities is crucial, as they significantly impact global health outcomes and the effectiveness of cancer prevention and treatment strategies.

This review paper is organized into three distinct parts:

1. **Risk Factors:** The first section examines prominent risk factors for cancer, particularly tobacco and alcohol consumption. In HICs, effective public health measures have successfully reduced tobacco use and moderated alcohol consumption, contributing to improved cancer control. In contrast, LMICs experience rising rates of tobacco and alcohol use due to weaker regulations and aggressive marketing strategies, leading to increased cancer risks, particularly for lung and oral cancers.
2. **Access to Healthcare Services in Rural vs. Urban Areas:** The second section explores disparities in cancer care between rural and urban areas, highlighting unique challenges faced by rural regions, such as cultural misunderstandings, limited resources, and fragmented healthcare systems. Urban areas typically benefit from better access to

comprehensive care and resources, underscoring the need for targeted interventions to address these geographic disparities.

3. **Policy & Cancer Disparities:** The final section addresses the impact of policy limitations on cancer outcomes in both HICs and LMICs. Inadequate healthcare infrastructure, fragmented services, and restrictive policies can impede effective cancer care. This section emphasizes the importance of comprehensive policy reforms to improve access to care and address disparities in treatment.

By exploring these three interrelated areas, this study aims to provide a comprehensive understanding of how disparities in access to cancer care contribute to poor outcomes globally. The insights gained will offer potential strategies for mitigating these disparities and advancing cancer prevention and treatment.

## **RISK FACTORS FOR CANCER IN LMICs vs HICs**

### **Tobacco in LMICs**

Smoking tobacco accounts for 21% of cancer deaths worldwide. It causes 70% of lung cancers globally, with incidence varying by geography. Researchers find that tobacco and alcohol have largely been concentrated in high-income countries (HICs). However, there is a growing need to understand and address these issues in low- and middle-income countries (LMICs), given their substantial and increasing cancer burden due to these substances. Tobacco use is a major risk factor for many different types of cancers, including lung, head and neck, esophagus, stomach, pancreas, liver, kidney, bladder, and cervix, as well as leukemia. The incidence of smoking-related cancer cases is significantly higher in LMICs. Globally, just under 7 million smoking-related cancer cases are diagnosed each year. 2.3 million cases occur in HICs and 4.3 in LMICs.

This issue is prominent in India, which faces widespread tobacco use and the resulting health problems. Tobacco use in India is a major cause of oral cancers, with the country having one of the highest rates of this disease globally (Mohan & Lando). Over 50% of oral cancer cases in India are linked to smokeless tobacco, a widespread practice (Pandey et al). Smokeless tobacco is commonly consumed by placing it in the mouth or chewing, and its use is culturally ingrained in India. Additionally, bidi smoking—the smoking of a popular, unfiltered, and often flavored cigarette—further exacerbates the problem. Both forms of tobacco use significantly contribute to India's oral cancer burden. Furthermore, this issue extends to non-smokers in India with secondhand smoke (SHS) being three to four times more toxic per gram than mainstream smoke: secondhand smoke is known to contain over 4,000 identified harmful chemicals, including 250 known to be dangerous (Mishra et al.).

A study on India's tobacco issue found that smoking tobacco caused a 5-fold higher risk for oropharyngeal, laryngeal, and lung cancers, a 3-fold higher risk for hypopharyngeal and esophageal cancers, and a 2-fold higher risk for oral cancer (Sapkota et al.). Tobacco chewing

also has a significant association with esophageal and oral cancers. Further, bidi smoking contributes to a heightened risk for multiple types of cancer including a 6-fold increased risk for lung cancer, a 3.5-fold increased risk for esophageal cancer, and a 3-fold increased risk for oral cancer (Prasad and Dhar 803).

### **Tobacco in HICs**

Tobacco use in high-income countries (HICs) has decreased significantly, with current prevalence rates at 30% for men and 19% for women. In the United States, for instance, tobacco use among men declined from 52% in 1960 to 22% in 2010, while among women, it dropped from 34% to 17% (Lee and Hashibe). In the United States, according to the American Cancer Society, smoking causes about 20% of all cancers and 30% of all cancer deaths. It is responsible for around 80% of lung cancers and lung cancer deaths, making lung cancer the leading cause of cancer death in the country. Smoking also contributes to chronic obstructive pulmonary disease (COPD), a leading cause of death, and significantly increases the risk of heart disease, stroke, and other cardiovascular issues (American Cancer Society).

Similarly, tobacco smoking is the leading preventable cause of cancer in Australia, contributing to 20% of the nation's cancer burden and resulting in 16,610 cancer cases and 9,921 deaths in 2013. Additionally, 90% of lung cancer cases in men and 65% in women are attributed to smoking (Cancer Council Australia).

The United States has recognized this issue and implemented laws to help lessen it through point-of-sale restrictions. An example includes the Family Smoking Prevention and Tobacco Control Act where the FDA regulates tobacco products to protect public health by restricting sales to minors, banning tobacco sales through vending machines, and limiting marketing to youth (National Cancer Policy Forum et al.). This law requires large health warning labels on tobacco products including e-cigarettes and prohibits false health claims. Smoking bans in public places such as workplaces, airports, restaurants, and bars also help reduce exposure to secondhand smoke and improve public health (National Cancer Policy Forum et al).

Furthermore, increasing tobacco taxes is a highly effective strategy for preventing cancer and improving public health. By raising cigarette prices, taxes reduce tobacco use, particularly among young people and low-income groups who are more price-sensitive (Le & Jaffri). This reduction in consumption lowers cancer rates and decreases healthcare costs. The Tobacconomics Cigarette Tax Scorecard measures how effectively countries use cigarette taxes to reduce smoking by assessing tax structure, the share of taxes in retail prices, cigarette affordability, and how frequently taxes are updated; Australia and New Zealand's high score of 4.63 out of 5.0 reflects their strong and regularly updated tobacco tax policies, which have significantly reduced cigarette affordability and smoking rates (Chaloupka et al.).

While smoking rates in HICs such as the U.S. and Australia have significantly declined, countries like France, part of the EU5, continue to grapple with higher smoking prevalence. In 2020, smoking rates among adults in France were reported at 25.5%,(Global State of Tobacco Harm Reduction). France still grapples with tobacco use due to deeply embedded cultural

attitudes towards smoking in public spaces like cafes and bars. The economic and social roles these places play may contribute to the resistance to stronger regulations. These venues serve as important social hubs where people gather for leisure, conversation, and networking, and they also play a significant role in the local economy by attracting tourism and supporting small businesses. This social and economic significance may contribute to the resistance to stronger regulations (The Parliament Magazine). (The Parliament Magazine).

Tobacco prevalence has been decreasing in high-income countries with the rise of anti-smoking campaigns. For example, the United States' adult smoking rate has decreased from 42% to 11.5% between 1960 and 2022 (CDC). On the other hand, tobacco has become more prevalent in LMICs, with an estimated 39% of adult populations in middle-income countries and 24% in low-income countries, due to factors such as weaker tobacco control policies, economic factors making tobacco more affordable, social norms and cultural acceptance of smoking, and limited public health infrastructure (Tauras).

### **Alcohol in LMICs**

LMIC populations undergo a higher burden of alcohol-related harm, worsened by historical and socioeconomic factors. This includes spending a larger percentage of their income on alcohol and facing greater health impacts compared to higher-income countries (Walls et al. e001958). In LMICs the cancer burden due to alcohol is substantial and expected to rise, with alcohol-related cancers predicted to have a 68% increase in LMICs (Lee and Hashibe 380). Alcohol use is increasing in low-income and middle-income countries (LMICs) like South Africa, driven by assertive marketing strategies from the alcohol industry targeting these regions as new growth markets. In South Africa, alcohol consumption significantly contributes to health issues, most commonly cancer as well as other non-communicable diseases, high levels of violence, and increased risk of infectious diseases like HIV and tuberculosis (Walls et al. e001958).

In Uganda, alcohol consumption is notably high, with a yearly per capita consumption of 23.7 liters. A nationwide survey conducted in 2014 revealed that 26.8% of Ugandans are current alcohol users, with 9.8% having alcohol-use disorders (AUD)(Kabwama et al.). The survey highlights important connections between alcohol use and risks, including cancer, but also traffic accidents, HIV infection, sexual oppression, and domestic violence. Public health interventions in Uganda need to address these regional disparities and AUDs to effectively reduce alcohol-related issues (Kabwama et al.). In Uganda, interventions should focus on amplifying legislation and using cultural institutions to promote alcohol-use disorder resources. Increasing awareness and support for early intervention and prevention programs could also help (Kalema et al.).

### **Alcohol in HICs**

A new study published by the *A Cancer Journal for Clinicians* by the American Cancer Society and the International Agency for Research on Cancer says alcohol is a major risk factor

for cancer, accounting for about 40 percent of all cancer cases within the age group of thirty and above in U.S. adults (“Alcohol’s Role in Cancer Risk”). Alcohol is associated with seven types of cancer, including breast and colorectal cancer. In 2019, alcohol was attributed to 44,000 cases of breast cancer, accounting for 16% of all breast cancer cases, and attributed to 18,000 cases of colorectal cancer, accounting for 13% of all colorectal cancer cases. Another study by the CDC reported that even moderate drinking raises the risk of cancer and that adherence to the recommended limits may prevent up to 80% of the deaths occurring due to alcohol-related cancer (“Alcohol’s Role in Cancer Risk”).

In Japan, a study on HIC’s led by Harvard T.H. Chan School of Public Health found that moderate alcohol consumption is associated with an increased risk of cancer among hospital patients in Japan. Consuming one drink daily (equivalent to six ounces of wine, 17 ounces of beer, or two ounces of whiskey) for ten years or two drinks a day for five years raised the relative risk of various cancers by 5% compared to non-drinkers. Zaitso stressed the importance of moderation in alcohol consumption to mitigate cancer risk. The study also noted that genetic differences affecting alcohol metabolism may influence these findings (Harvard T.H. Chan School of Public Health).

Nordic HICs (including Denmark, Finland, Iceland, Norway, and Sweden) implement strict alcohol regulations that have led to some of the lowest consumption rates in Europe. Their approach includes government-run stores, high alcohol taxes, and strict marketing bans. These measures aim to reduce alcohol availability, protect public health, and limit consumption, especially among youth. The Nordic model prioritizes health over economic profits, demonstrating an effective strategy for minimizing alcohol-related harm specifically cancer. ("Reducing Alcohol Consumption")

### **Environmental Risks LMICs**

Environmental risk factors for cancer present substantial challenges in Low- and Middle-Income Countries (LMICs), often characterized by industrial practices, contaminated water sources, and dietary habits unique to these regions.

A 37-year cohort study in China found that asbestos factory workers had more than three times the risk of lung cancer deaths compared to controls (Wang et al.). Similarly, Taiwanese shipbreaking workers showed increased cancer incidence, especially in oral and lung cancers, linked to high asbestos exposure (Wu et al.). In Brazil, the heavy use of asbestos in cement products has led to rising mesothelioma mortality rates, particularly in areas with numerous cement factories (Algranti et al.). Argentinean and Ukrainian data reveal similar trends, with increased mesothelioma deaths following asbestos use despite bans (Nishikawa et al.).

Arsenic contamination in drinking water is another critical issue, with over 200 million people globally exposed to levels above the WHO's safety standard of 10 µg/L. In countries like Bangladesh, where the guideline is 50 µg/L, chronic arsenic exposure is linked to various cancers, including skin, bladder, lung, prostate, and liver cancers. Studies from Taiwan and Chile have identified associations between arsenic exposure and increased cancer risks, with early-life

exposure further compounding these risks. Additionally, arsenic contamination in food, particularly rice, contributes to cancer risk, with young children being particularly vulnerable (Hashim and Boffetta 399).

Future studies of cancer prevention in LMICs should focus on environmental risks, which may differ from those in HICs. Also, understanding how different carcinogens, like HBV infection combined with smoking and drinking, interact could reveal important cancer mechanisms. The high rate of esophageal cancer in Central Asia suggests that environmental factors like certain chemicals and drinking hot beverages might play a role. Large-scale studies in these areas could provide valuable insights for improving cancer prevention worldwide (Kumar et al. 977).

### **Environmental Risks in HICs**

Similar to LMICs, HICs also face significant environmental cancer risks; however, these risks emerge differently based on regional environmental factors and industrial practices. In European HICs, air pollution is a severe health issue, with 91% of the urban population exposed to pollutants above WHO guidelines. This exposure leads to over 300,000 premature deaths annually, many from cancer. Air pollution is linked to about 0.5-1% of all cancer cases and over 7% of lung cancers (European Environment Agency).

Long-term exposure to air pollution is associated with leukemia and esophageal cancer from indoor air pollution. Although emissions of key pollutants have decreased, levels have largely returned to pre-pandemic conditions. As of 2020, 96% of Europe's urban population faces unsafe PM<sub>2.5</sub> (tiny air particles 2.5 micrometers or smaller that can harm health by infiltrating deep into the lungs and bloodstream) levels (European Environment Agency). The EU is managing this problem through policies like the National Emission Ceilings (NEC) Directive and Ambient Air Quality Directives. The NEC Directive sets legally binding limits on the total emissions of certain pollutants, including nitrogen oxides (NO<sub>x</sub>), sulfur dioxide (SO<sub>2</sub>), and particulate matter (PM<sub>2.5</sub>), to improve air quality across member states. It aims to ensure that emissions remain below established ceilings, thereby protecting public health and the environment. (European Environment Agency)

The Ambient Air Quality Directives establish air quality standards for various pollutants, including PM<sub>2.5</sub>, to safeguard human health and the environment. These directives require EU countries to monitor air quality and take action when pollution levels exceed the established limits (European Commission).

In line with these efforts, the EU plans to align more closely with WHO guidance, aiming to reduce early deaths from PM<sub>2.5</sub> by 55% by 2030. This ambitious goal highlights the EU's commitment to improving air quality, reducing health risks associated with air pollution, and promoting sustainable practices across member states (United Nations Regional Information Centre).

In Slovenia, another HIC, indoor radon exposure is a notable risk factor for lung cancer, especially in areas with high radon levels. Studies of lung cancer cases from 1978 to 2017 show

that regions with medium to high radon exposure have a 5-7% higher risk of lung cancer. Radon exposure is responsible for lung cancer cases in other HICs, including 8% in Switzerland and 5% in Germany. WHO guidelines recommend lowering radon levels in homes to below 100 Bq/m<sup>3</sup> to reduce health risks. Slovenia's lung cancer rates and radon exposure patterns highlight the need for ongoing public health efforts to manage radon-related cancer risks (Birk et al.).

### **ACCESS TO HEALTHCARE SERVICES IN RURAL vs. URBAN AREAS**

Qualitative studies identified several barriers to healthcare access in rural areas, including cultural misunderstandings, fragmented communication, and resource constraints. Rural healthcare systems are often hindered by individualism, stigma, and poor follow-up. Effective communication and standardized cultural training are lacking, leading to delays and errors. Limited resources and time constraints harm rural health systems. Additionally, the profit-driven nature of the US healthcare system exacerbates access disparities. Improving rural healthcare access requires new delivery models, better affiliations with larger systems, workforce development, and expanded telehealth services.

#### **Access to Healthcare Services in Rural vs. Urban Areas: HICs**

Although the cancer death rate has been dropping and new cancer treatments have been developed in the U.S., large disparities between rural and urban cancer care still exist. While the cancer incidence rate is lower in rural areas, mortality is higher compared to that in urban areas (National Cancer Institute, n.d.). This gap stems mainly from late diagnosis of the cancer, coupled with poor quality of care and poorer outcomes for patients in rural areas (Sokale, Raza, & Thrift).

A study found that rural cancer patients in California faced significant care quality disparities compared to urban Californian patients. Rural residents had lower rates of adequate pathological examinations for gastric and colon cancers and were less likely to receive radiation therapy after breast-conserving surgery. These issues stem from limited specialized medical resources and logistical challenges in rural areas. Despite these disparities, cancer-specific survival rates were similar between rural and urban patients when adjusted for other factors. The results underscore the need for improved access to specialized care and support services in rural areas (Parikh-Patel et al.).

Furthermore, another study describes how patients in rural areas of the United States face significant challenges compared to their urban counterparts, including limited access to specialized care, transportation obstacles, financial problems, and fewer clinical trial opportunities. Rural areas have fewer oncologists and cancer treatment facilities, which leads to longer travel times for care and increased difficulties in managing co-existing health conditions. This shortage is compounded by socioeconomic disadvantages impacting patients and providers (Coombs, Campbell, & Caringi).

Another study shows that cancer mortality in rural and remote Australia is higher than in metropolitan areas, in part due to greater travel distances to care and limited facilities locally.



While cancer incidence is slightly lower outside major cities, mortality rates are higher, particularly for Aboriginal and Torres Strait Islander populations. The barriers include inconveniences related to the distance of traveling and a lack of advanced diagnostic facilities for treatment. Solutions are being explored such as the expansion of regional cancer centers, integrated care pathways, outreach programs, and improved telemedicine; however, there are still problems concerning staffing and cultural considerations (Smith 38).

Strategies to address these issues include outreach clinics, Tele-oncology, virtual tumor boards, (multidisciplinary meetings where healthcare providers from various specialties, such as oncologists, radiologists, surgeons, and pathologists, come together virtually to discuss cancer cases and collaboratively decide on the best treatment plans), and improved provider recruitment and education. These approaches aim to lessen the burden of long travel distances and limited local resources. Expanding these strategies and developing new solutions are crucial to improving access to high-quality cancer care for rural residents (Charlton et al.).

### **Access to Healthcare Services in Rural vs. Urban Areas: LMICs**

Access to healthcare presents significant challenges in rural areas of India, where 65% of the population faces socio-economic disadvantages and limited health awareness. The costs of cancer treatment are particularly high, especially in advanced cancer stages, often resulting in catastrophic out-of-pocket expenses for many rural families. While the Ayushman Bharat-Pradhan Mantri Jan Arogya Yojana, a health insurance government initiative, has made notable efforts to expand health insurance coverage, gaps persist in early detection and care.

These challenges have driven the development of the Health System Capacity Assessment (HSCA) tool, which has been adapted for the Indian context under the Access Cancer Care India project. This tool focuses on assessing health system readiness for cancer care, with an emphasis on screening and early diagnosis, and aims to identify barriers to improved service delivery at all levels through a multi-level strategy. The Access Cancer Care India (ACCI) tool was designed to model and address systemic issues in cancer care delivery and has the potential for broader application in low- and middle-income countries (LMICs) (Mallafre-Larrosa et al.).

Another study shows a significant gap in cancer care between metro cities and rural areas in India. Radiation therapy is a key first-line treatment for many cancers in India, but the significant shortage of Radiation Therapy Units (RTUs) in rural areas limits access to timely care, exacerbating treatment delays and worsening outcomes for rural patients. In contrast, metro cities, which have only 10.9% of the population, control 38% of the RTU capacity. The rest of India has limited access to cancer treatment. This disparity in RTU availability is representative of problems facing other developing countries (Chaudhuri et al.).

### **Policy & Cancer Disparities**

This section examines the critical role of policy in addressing cancer disparities across both high-income and low and middle-income countries. It explores how structural barriers, social determinants of health, and affordability issues are shaped by and influence health

policies. Analyzing these factors gives us insights into how effective policy reforms can improve cancer care access and outcomes.

### **Structural Barriers and Policy Reform in HICs**

In the United States, Medicaid expansion under the Affordable Care Act (ACA) has greatly improved cancer care by increasing timely screenings and treatment, leading to better survival rates. While 40 states that adopted the expansion have seen reduced uninsured rates and more comprehensive care, the 10 states that opted out continue to face disparities. Research gaps, such as the lack of long-term outcome data, highlight the need for further study (Hotca et al.). Health inequity, which affects cancer outcomes by creating barriers to screening, diagnosis, and treatment, requires efforts from patients, advocacy groups, and healthcare professionals to address these disparities (Richardson-Parry et al.). Community-based interventions, like those from the Tigerlily Foundation providing mobile mammography and educational events, and policies supporting workplace protections for screenings, are crucial in improving access (American Cancer Society). Community Health Workers (CHWs) also enhance cancer screenings in underserved areas through culturally tailored support (O'Donovan et al., 1; Kale et al., 1).

### **Late-stage Diagnoses**

Preventative care is an important component of cancer outcomes. For example, late-stage breast cancer diagnoses drastically reduce survival rates, with 5-year overall survival (OS) at 90.0% for stage I, but dropping to just 22.8% for stage IV. These outcomes are further influenced by race, ethnicity, and socioeconomic status, as survival rates differ significantly among these groups. For example, uninsured patients or those from counties with lower median incomes face worse prognosis due to limited access to timely screenings and advanced treatments. Structural barriers, including healthcare access in rural and underserved communities, exacerbate these disparities. While age-standardized 5-year OS has improved overall, with a 0.97% annual increase observed from 1975 to 2010, and notable shifts during periods like 1983-1987 (a 2.5% increase per year), disparities persist (Yang et al.). Advances in the standard of care and earlier detection have played key roles in these improvements, yet access to such advancements is uneven across populations. Continued efforts to improve screening accessibility and address socioeconomic barriers are essential to closing the survival gap and ensuring equitable outcomes for all breast cancer patients (Yang et al.).

Two key policies that support cancer care and prevention are the ACA's Preventive Services Mandate, which requires insurers to cover cancer screenings without copayments, and the Community Health Center Program, which funds grants for comprehensive health services, including cancer prevention, to underserved populations (U.S. Department of Health and Human Services; Health Resources and Services Administration). Moreover, the National Health Service (NHS) of the UK is responsible for full cancer care that is paid for using general taxation. (NHS England, "Cancer Services")

Cancer treatment services are widely offered in NHS hospitals as well as screening facilities, while private patients (PPPs) are also included in the system. The NHS Cancer Plan aims at enhancing early identification and treatment with equal entry across Britain. Nevertheless, there are apprehensions concerning the burden imposed on publicly financed services and resource allocation along with patient safety when it comes to PPPs' treatment, thereby necessitating continual policy monitoring (Walpole 56)

### **Structural Barriers and Policy Reform in LMICs**

Noncommunicable diseases (NCDs), including cancer, are now recognized as a major public health crisis by the United Nations and WHO, with cancer care in LMICs acknowledged as a global health priority (World Health Organization).

In LMICs, about 80% of cancer patients are diagnosed at advanced stages (Stage III or IV), leading to worse outcomes and higher death rates. This late diagnosis is often due to several factors: low public health awareness, weak healthcare systems, and inadequate training for doctors in recognizing cancer symptoms. Many LMICs also lack comprehensive cancer registries, making it difficult to understand the full extent of the cancer problem to leverage in the creation of effective policies. The World Health Assembly advises setting up national cancer control plans to improve cancer prevention, early detection, treatment, and care. However, these plans are often not fully carried out because of limited resources in these LMICs (Cazap et al. 14). Cultural factors, such as stigma and lack of awareness, particularly in rural areas, hinder early diagnosis and treatment. Women, for example, may be reluctant to report symptoms due to fear or cultural norms. Economic barriers include the cost of transportation and treatment, which prevent many patients from accessing care (Brand et al. e1371).

In Brazil, breast cancer is a major public health problem, and disparities in access to healthcare services worsen the situation. Despite increased funding and initiatives such as the "Viva Mulher" program (meaning "Live Woman," aimed at promoting women's health and increasing access to breast cancer screenings) (Porto & Habib) and the "Plano de Ações Estratégicas" (translated as "Strategic Action Plan," which outlines targeted actions to enhance public health efforts regarding breast cancer), regional disparities persist, especially in poorer areas (Ministério da Saúde). Policies have aimed at improving early detection and treatment, with initiatives like "Outubro Rosa" (or "Pink October," an annual campaign raising awareness about breast cancer and encouraging regular screenings). However, there are still challenges including cultural stigma, economic barriers, and limited infrastructural provision of diagnosis and treatment facilities.

Although some statistics indicate a rise in detection rates due to the introduction of policy efforts, it is necessary to have a more holistic and fair enactment of these policies that will combat the obstacles on the way while decreasing breast cancer death rates in every part of the country (Figueiredo et al. 121).

Another challenge for LMICs is the shortage of trained oncologists. These countries typically have less healthcare workers. Specifically, LMICs have 1.3 physicians and 2.5 nurses

per 1,000 people, compared to the 3.1 physicians and 10.9 nurses per 1,000 people in HICs. This disparity results in a higher patient load for oncology providers in LMICs and an urgent shortage of specialists in fields like pathology, radiation, and oncology. The lack of adequate oncology training programs and focus on cancer care during medical education worsens these challenges (Karim et al.).

### **Social Determinants of Health in HICs**

The cancer research community is increasingly aware of the need to address structural and social determinants of health (SDOH) to improve cancer control and screening equity. Current American policies often create barriers to addressing basic needs like food and housing. For example, if a person's income changes, they might lose access to programs like SNAP (food assistance), impacting interventions that handle basic needs. Similarly, hospital systems and pharmaceutical companies may face restrictions, referring to legal and regulatory barriers that limit the ability of healthcare providers, hospitals, and pharmaceutical companies to address patients' social needs, such as offering more affordable care or assistance programs or lower-cost prescriptions. To improve cancer care, policies must address these barriers. Social needs like food, housing, and transportation are crucial for cancer outcomes, and healthcare providers can push for changes to help meet these needs. Solving problems related to SDOH requires tackling issues at a broad level, including social, political, and economic factors. This requires collaboration between different sectors and levels of government (Tucker-Seeley et al. 2176).

Factors like unemployment, lack of education, poverty, and income disparities- critical social determinants of health- are intertwined with poor health and cancer survival outcomes, the reduction of which calls for interventions that consider the broader social and environmental contexts in which people live. (Goodman and Huang) The American Cancer Society Cancer Disparity Research Team investigates how these factors—along with health insurance and housing access—may bear upon cancer screening, diagnosis, and treatment ("Socioeconomic Factors and Cancer Disparities").

Housing insecurity, including high housing costs, frequent moves, and homelessness, affects millions of Americans, particularly low-income individuals, racial minorities, and unmarried people. For cancer patients and survivors, these housing issues can exacerbate health problems by increasing financial stress, leading to poorer health outcomes, and limiting access to care (Martin P). Programs like those from the Housing and Urban Development (HUD) offer support to address housing instability. However, even with assistance, the high costs of cancer treatment can still be unaffordable for those struggling with housing issues. This financial burden can force patients to choose between paying for housing or health care (National Alliance to End Homelessness).

### **Social Determinants of Health in LMICs**

ASEAN (Association of Southeast Asian Nations) is a regional organization made up of ten Southeast Asian countries, which aims to promote economic cooperation, cultural exchange, and political stability among its members. DALY(Disability-Adjusted Life Year) is a metric that

measures the overall burden of disease by combining years lost due to premature mortality with years lived with disability, providing a comprehensive assessment of health impacts. In the ASEAN region, more specifically in the LMIC countries, marginalized groups such as indigenous peoples, LGBTQ+ individuals, and those from low socioeconomic backgrounds face significant barriers to cancer healthcare. Indigenous peoples often live in remote areas with limited healthcare access, while LGBTQ+ individuals encounter discrimination that hinders their ability to receive proper care. Women in patriarchal societies and economically deprived individuals also struggle with access due to economic and social obstacles (Arevalo et al. 107110).

Policies addressing these issues include national cancer control plans aimed at improving cancer prevention and treatment. An example is The Philippines National Cancer Control Program (NCCP). In the Philippines, the National Cancer Control Program (NCCP), established by the Department of Health, focuses on cancer prevention, early detection, treatment, and care to reduce mortality and enhance patient quality of life. The Universal Health Care Act (2019) aims to ensure all Filipinos have access to comprehensive health services, including cancer care, by expanding medical access and improving insurance coverage. Policy reforms must also address the underlying socioeconomic determinants of health, such as poverty and education, to effectively reduce cancer disparities in LMICs (Dee et al.).

### **Access to Precision Medicine**

In addition to modifiable risk factors such as tobacco use and alcohol consumption, genetic and immunological mechanisms also play a critical role in cancer development. For example, studies have shown that increased activity of activation-induced deaminase (AID), an enzyme involved in antibody diversification, can lead to genomic instability, resulting in chromosomal translocations and mutations that may predispose individuals to cancer (Wang et al.). This highlights the complex interplay between immune function and cancer risk, emphasizing the need for careful regulation of biological processes that, while essential for immune defense, can also pose significant cancer risks when poorly regulated.

The impact of genomic instability on cancer risk may disproportionately affect LMICs compared to HICs. In LMICs, limited healthcare infrastructure and access to advanced genetic testing or treatments can delay the identification and management of cancers caused by such genomic mutations. Additionally, the higher prevalence of infections in LMICs, some of which are associated with immune responses that can trigger genomic instability, may compound cancer risks. In contrast, HICs are better equipped with resources for early detection and advanced treatment options that can mitigate the effects of such mutations, reducing overall cancer outcome disparities between these regions (Drake et al.).

### **Affordability in HICs**

In HICs, cancer medicines are generally more affordable compared to LMICs. Specifically, in HICs, the cost of one Defined Daily Dose (DDD) of medicine requires less than 5 days' wages, with a median of 2.04 days/DDD. This differs drastically with LMICs, where the

cost can exceed 5 days' wages, reaching up to 13 days' wages or more in some cases (Moye-Holz and Vogler). HICs benefit from effective pricing and reimbursement policies, which help manage medicine costs (for individuals) better than LMICs.

In the UK, consisting of all HICs, patients usually do not pay for their cancer treatment directly as it is covered by general taxation in the National Health Service (NHS). This model of affordability ensures that individuals do not pay directly for their cancer care. However, arguments are going on about whether patient charges could be introduced to regulate the increasing expenses related to cancer care while still maintaining its quality standard. The NHS' budget from the UK was cut down and cancer treatment takes up 5.6% of all NHS annual expenditures. There was a boom in cancer financing from 2003 to 2010 but since, the costs have been increasing due to technological advancements, population aging as well as a rise in cases of cancer. These financial pressures cause debates on how access to cancer services of good quality can be preserved. This has sparked discussions on how to sustain access to high-quality cancer care amidst these financial challenges ("Affordability of Cancer Care" 447). Additionally, a key limitation of systems like the NHS is that treatments are sometimes prioritized based on quality-adjusted life years (QALYs), meaning older patients may experience longer wait times for cancer treatment compared to younger patients, which reflects a tradeoff in balancing cost and access (Williams, 1994; Cancer Research UK).

Furthermore, a study in Canada focuses on public engagement and emphasizes the importance of affordable and equitable funding for cancer drugs. Participants advocated for taking into account aspects such as the price of drugs, clinical advantages, and possible side effects while deciding upon funds (Bentley et al.). They highlighted the importance of transparency and accountability in the course of making decisions, supporting a pan-Canadian model of cancer drug funding that would reduce inequalities across provinces. It appears from the findings that policymakers should formulate ways through which the effectiveness of drugs can be reassessed and investments made while ensuring that patients can continue their medications throughout transitions. The study demonstrates how important it is to involve informed citizens in health policy choices to promote equitable and enduring practices regarding payment for cancer treatment (Bentley et al.).

Despite effective policies in HICs like Canada and the UK to manage cancer treatment costs, financial pressures and disparities persist. Continuous policy innovation and public engagement are essential to ensure equitable and sustainable access to quality cancer care for all patients.

### **Affordability in LMICs**

Economic challenges in LMICs are immense, and the costs for cancer treatment are extremely high and likely to influence access and patient outcomes. While cancer incidence and

mortality continue to rise in these countries, many effective technologies remain out of reach due to their high expense and dependence on sophisticated infrastructure.

To reduce these issues, various financial assistance programs, cost-reduction strategies, and international aid efforts are being implemented. The Affordable Cancer Technologies (ACTs) program, funded by the National Cancer Institute (NCI) and the National Institute for Biomedical Imaging and Bioengineering (NIBIB) in the United States, plays a crucial role in developing low-cost, point-of-care technologies for LMICs, created to increase early detection and access to treatment of cancer. ACTs help adapt such technologies by transitioning them from research to practical solutions that can be implemented in the medical field. Among the innovations falling under this program are microfluidics, mobile devices, and 3D printing. (Pearlman et al. 1)

An example of affordability challenges in LMICs is that the cost of medicines often requires between 5 to over 20 days' wages per defined daily dose (DDD), with Mexico standing out due to particularly high medicine prices that make them very unaffordable (Moye-Holz and Vogler). In Mexico, cancer care is largely provided by public sector institutions, including the IMSS (Instituto Mexicano del Seguro Social), ISSSTE (Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado), and various Ministry of Health (MoH) facilities. While innovations in cancer medicines are emerging, their high costs often render new treatments unaffordable for patients and create financial strain on health systems. This pricing issue frequently leads to newer, more effective medicines being excluded from treatment formularies. (Moye-Holz et al.)

## **Conclusion**

This review examined the multifaceted landscape of cancer risk factors and healthcare access disparities, highlighting significant differences between low- and middle-income countries (LMICs) and high-income countries (HICs). Key findings include the impact of tobacco use and alcohol consumption on cancer rates, with varied effectiveness of regulatory measures across different regions. The analysis revealed that while HICs benefit from stringent tobacco regulations and higher alcohol taxes, LMICs face escalating cancer burdens due to less regulated environmental factors and socio-economic challenges. Additionally, disparities in healthcare access between rural and urban areas persist, affecting cancer outcomes globally.

To address these disparities, it is imperative for policymakers and public health organizations to enhance regulatory frameworks and interventions tailored to regional needs. Increased global collaboration to strengthen tobacco and alcohol control measures, alongside targeted environmental and public health initiatives, is crucial. In LMICs, expanding access to healthcare resources and implementing robust prevention programs can mitigate rising cancer rates, while HICs must continue to innovate in reducing environmental cancer risks and improving healthcare delivery in underserved areas.

Future research should focus on evaluating the effectiveness of current tobacco and alcohol policies in diverse contexts, as well as exploring new strategies to address environmental

and socio-economic factors contributing to cancer. Additionally, investigating innovative models for improving healthcare access in rural and low-resource settings can provide valuable insights for enhancing cancer care equity worldwide. By fostering international cooperation and investing in evidence-based interventions, we can work towards reducing cancer disparities and improving global health outcomes.



## Works Cited

- Abolhasani, F., et al. "Gender Differences in Cancer Risk and Outcomes: Insights from the Global Cancer Observatory." *Journal of Global Oncology*, vol. 8, 2022, <https://doi.org/10.1200/JGO.21.00621>.
- Algranti, E., et al. "Analysis of Mortality from Asbestos-Related Diseases in Brazil Using Multiple Health Information Systems, 1996–2017." *Safety and Health at Work*, vol. 13, no. 3, 2022, pp. 302-307, <https://doi.org/10.1016/j.shaw.2022.04.006>.
- Andreeva, V. A., et al. "Assessment of Alcohol Consumption and Its Impact on Cancer Incidence: A Review of the Literature." *Cancer Prevention Research*, vol. 11, no. 2, 2018, pp. 59-65, <https://doi.org/10.1158/1940-6207.CAPR-17-0401>.
- Arana, A., et al. "Tobacco Use and Cancer Incidence: A Comparative Analysis of Smoking Trends in Europe." *European Journal of Cancer Prevention*, vol. 30, no. 4, 2021, pp. 303-309, <https://doi.org/10.1097/CEJ.0000000000000631>.
- American Cancer Society. "Health Risks of Smoking Tobacco." American Cancer Society, 13 Oct. 2022, <https://www.cancer.org/healthy/stay-away-from-tobacco/health-risks-of-smoking-tobacco.html>.
- Arevalo, M. V. P. N., et al. "Social Determinants of Sex Disparities in Cancer in Southeast Asia." *iScience*, vol. 26, no. 7, 2023, 107110, <https://doi.org/10.1016/j.isci.2023.107110>.
- Asiedu, G. B., Wan, M. T., Jacobson, L. E., and Maiese, B. A. "Addressing Global Cancer Care Inequities Using Implementation Science and Community-Engaged Research Approaches." *European Medical Journal Innovations*, 2023, <https://www.emjreviews.com/innovations/article/addressing-global-cancer-care-inequities-using-implementation-science-and-community-engaged-research-approaches/>.
- Bagnardi, V., et al. "Alcohol Consumption and Cancer Risk: A Review of the Evidence and Its Implications." *Journal of Clinical Oncology*, vol. 34, no. 22, 2016, pp. 2727-2738, <https://doi.org/10.1200/JCO.2016.68.5515>.
- Barchitta, M., et al. "The Role of Lifestyle Factors in Cancer Prevention: Evidence from a Population-Based Study." *Cancer Epidemiology*, vol. 77, 2022, 102059, <https://doi.org/10.1016/j.canep.2022.102059>.
- Bentley, C., et al. "Addressing the Affordability of Cancer Drugs: Using Deliberative Public Engagement to Inform Health Policy." *Health Research Policy and Systems*, vol. 17, no. 1, 2019, p. 17, <https://doi.org/10.1186/s12961-019-0411-8>.
- Boffetta, P., et al. "Risk of Cancer Associated with Different Types of Tobacco Smoking: A Systematic Review and Meta-Analysis." *International Journal of Cancer*, vol. 134, no. 5, 2014, pp. 1055-1065, <https://doi.org/10.1002/ijc.28478>.
- Brand, N. R., et al. "Delays and Barriers to Cancer Care in Low- and Middle-Income Countries: A Systematic Review." *The Oncologist*, vol. 24, no. 12, 2019, pp. e1371-e1380, <https://doi.org/10.1634/theoncologist.2019-0057>.

- Cazap, E., et al. “Structural Barriers to Diagnosis and Treatment of Cancer in Low- and Middle-Income Countries: The Urgent Need for Scaling Up.” *Journal of Clinical Oncology*, vol. 34, no. 1, 2016, pp. 14-19, <https://doi.org/10.1200/JCO.2015.61.9189>.
- Cancer Council Australia. “Tobacco and Cancer.” Cancer Council Australia, <https://www.cancer.org.au/cancer-information/causes-and-prevention/smoking-and-tobacco>.
- Chaloupka, F., et al. “Tobacconomics Cigarette Tax Scorecard.” Health Policy Center, Institute for Health Research and Policy, University of Illinois Chicago, 2020, <https://www.tobacconomics.org>.
- Charlton, M., et al. “Challenges of Rural Cancer Care in the United States.” *Cancer Network*, 15 Sept. 2015, <https://www.cancernetwork.com/view/challenges-rural-cancer-care-united-states>.
- Chi, M., et al. “Impact of Tobacco Control Policies on Smoking Prevalence and Cancer Outcomes: A Global Perspective.” *Tobacco Control*, vol. 28, no. 6, 2019, pp. 621-628, <https://doi.org/10.1136/tobaccocontrol-2018-054533>.
- Coombs, N. C., Campbell, D. G., and Caringi, J. “A Qualitative Study of Rural Healthcare Providers' Views of Social, Cultural, and Programmatic Barriers to Healthcare Access.” *BMC Health Services Research*, vol. 22, no. 1, 2022, p. 438, <https://doi.org/10.1186/s12913-022-07829-2>.
- Dee, E. C., et al. “Improving Cancer Care in the Philippines: The Need for Deliberate and Careful Implementation of the National Integrated Cancer Control Act.” *The Lancet Regional Health: Western Pacific*, vol. 28, 2022, 100615, <https://doi.org/10.1016/j.lanwpc.2022.100615>.
- Drake, T. M., Knight, S. R., Harrison, E. M., and Søreide, K. “Global Inequities in Precision Medicine and Molecular Cancer Research.” *Frontiers in Oncology*, vol. 8, 2018, p. 346, <https://doi.org/10.3389/fonc.2018.00346>.
- European Commission. “Air Quality.” European Commission, [https://environment.ec.europa.eu/topics/air/air-quality\\_en](https://environment.ec.europa.eu/topics/air/air-quality_en). Accessed [insert date of access].
- European Environment Agency. “Air Pollution and Cancer.” European Environment Agency, 2021, <https://www.eea.europa.eu/publications/environmental-burden-of-cancer/air-pollution>.
- European Environment Agency. “Beating Cancer: The Role of Europe’s Environmental and Occupational Exposure.” European Environment Agency, 2020, <https://www.eea.europa.eu/publications/environmental-burden-of-cancer/beating-cancer-the-role-of-europes#:~:text=Environmental%20and%20occupational%20exposure%20to,risk%20factors%20are%20largely%20preventable>.
- European Environment Agency. “National Emission Ceilings Directive (NECD).” European Environment Agency, <https://www.eea.europa.eu/en/topics/in-depth/air-pollution/need>. Accessed [insert date of access].

- European Environment Agency. "Population Remains Exposed to Air Pollution." European Environment Agency, 21 Sept. 2023, <https://www.eea.europa.eu/en/newsroom/news/population-remains-exposed-air-pollution>.
- Figueiredo, F. W. d. S., et al. "The Role of Health Policy in the Burden of Breast Cancer in Brazil." *BMC Women's Health*, vol. 17, no. 121, 2017, <https://doi.org/10.1186/s12905-017-0477-9>.
- Hashim, D., and P. Boffetta. "Occupational and Environmental Exposures and Cancers in Developing Countries." *Annals of Global Health*, vol. 80, no. 5, 2014, pp. 393-411, <https://doi.org/10.1016/j.aogh.2014.10.002>.
- Kabwama, S. N., et al. "Alcohol Use among Adults in Uganda: Findings from the Countrywide Non-Communicable Diseases Risk Factor Cross-Sectional Survey." *Global Health Action*, vol. 9, no. 1, 2016, <https://doi.org/10.3402/gha.v9.31302>.
- Karim, S., et al. "Oncology Training and Education Initiatives in Low and Middle Income Countries: A Scoping Review." *ecancermedicalscience*, vol. 15, 2021, p. 1296.
- Khan, N., et al. "Tobacco Smoking and Cancer: A Comprehensive Review of Current Evidence." *Cancer Letters*, vol. 484, 2020, pp. 38-46, <https://doi.org/10.1016/j.canlet.2020.05.035>.
- Kumar, V. S., et al. "Cancer Prevention Research in Low- and Middle-Income Countries." *Annals of Oncology*, vol. 30, no. 6, 2019, pp. 977-979, <https://doi.org/10.1016/j.annonc.2019.03.014>.
- Le, T. T. T., and M. A. Jaffri. "The Association between Smoking Behaviors and Prices and Taxes per Cigarette Pack in the United States from 2000 through 2019." *BMC Public Health*, vol. 22, 2022, p. 856, <https://doi.org/10.1186/s12889-022-13242-5>.
- Lee, Y.-C. A., and M. Hashibe. "Tobacco, Alcohol, and Cancer in Low and High Income Countries." *Annals of Global Health*, vol. 81, no. 3, 2015, pp. 380-382.
- Liu, S., et al. "Environmental Factors and Cancer Incidence: A Global Overview." *International Journal of Environmental Research and Public Health*, vol. 17, no. 18, 2020, p. 6890, <https://doi.org/10.3390/ijerph17186890>.
- Ministério da Saúde. "Plano de Ações Estratégicas para o Enfrentamento das Doenças Crônicas e Agravos Não Transmissíveis (DANT) no Brasil, 2022-2030." Ministério da Saúde, 2022, [https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/svsa/doencas-cronicas-nao-transmissiveis-dcnt/09-plano-de-dant-2022\\_2030.pdf/view](https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/svsa/doencas-cronicas-nao-transmissiveis-dcnt/09-plano-de-dant-2022_2030.pdf/view).
- Misir, M., et al. "Structural Barriers to Cancer Screening among Women in Low-Income Communities." *Journal of Health Care for the Poor and Underserved*, vol. 31, no. 2, 2020, pp. 1012-1028, <https://doi.org/10.1353/hpu.2020.0082>.
- Mohamad, N. I., et al. "The Impact of Health Insurance on Cancer Screening Behaviors in Low- and Middle-Income Countries: A Systematic Review." *Health Services Research*, vol. 57, no. 1, 2022, pp. 114-122, <https://doi.org/10.1111/1475-6773.13683>.
- Mohan, P., and H. A. Lando. "Oral Tobacco and Mortality in India." *Indian Journal of Clinical Medicine*, 2016, <https://doi.org/10.4137/IJCM.S25889>.

- Monami, M., et al. “Trends in Tobacco Consumption and Its Implications for Cancer Incidence: A Review of Recent Data.” *Cancer Epidemiology*, vol. 81, 2023, p. 102080, <https://doi.org/10.1016/j.canep.2023.102080>.
- Moreira, J. P. F., and A. R. Paiva. “Alcohol Consumption and the Risk of Cancer: An Epidemiological Review.” *Revista de Saúde Pública*, vol. 55, 2021, Article 18, <https://doi.org/10.11606/s1518-8787.2021055002952>.
- Morgado, S., et al. “Environmental Pollution and Cancer: A Systematic Review.” *Environmental Pollution*, vol. 259, 2020, Article 113995, <https://doi.org/10.1016/j.envpol.2019.113995>.
- Moye-Holz, D., et al. “Access to Innovative Cancer Medicines in a Middle-Income Country - The Case of Mexico.” *Journal of Pharmaceutical Policy and Practice*, vol. 11, 2018, Article 25, <https://doi.org/10.1186/s40545-018-0153-y>.
- Moye-Holz, D., and S. Vogler. “Comparison of Prices and Affordability of Cancer Medicines in 16 Countries in Europe and Latin America.” *Applied Health Economics and Health Policy*, vol. 20, 2022, pp. 67–77, <https://doi.org/10.1007/s40258-021-00670-4>.
- National Cancer Institute. “Rural Cancer Disparities: Rural-Urban Differences in Cancer Rates.” National Cancer Institute, n.d., <https://gis.cancer.gov/mapstory/rural-urban/index.html>. Accessed 26 Sept. 2024.
- National Cancer Policy Forum, Board on Health Care Services, Institute of Medicine. *Reducing Tobacco-Related Cancer Incidence and Mortality: Workshop Summary*. National Academies Press, 16 Apr. 2013, <https://www.ncbi.nlm.nih.gov/books/NBK206879/>.
- Nishikawa, K., et al. “Recent Mortality from Pleural Mesothelioma, Historical Patterns of Asbestos Use, and Adoption of Bans: A Global Assessment.” *Environmental Research*, vol. 207, 2022, p. 112239, <https://doi.org/10.1016/j.envres.2022.112239>.
- Pandey, A., et al. “Oral Smokeless Tobacco Consumption Pattern among Rural Indian Cancer Patients: A Prospective Survey.” *South Asian Journal of Cancer*, vol. 9, no. 1, 2020, pp. 17–19, [https://doi.org/10.4103/sajc.sajc\\_40\\_19](https://doi.org/10.4103/sajc.sajc_40_19).
- Parikh-Patel, A., et al. “Urban–Rural Variations in Quality of Care among Patients with Cancer in California.” *American Journal of Preventive Medicine*, vol. 61, no. 6, Dec. 2021, pp. E279-E288, <https://doi.org/10.1016/j.amepre.2021.05.021>.
- Parliament Magazine. “Is the French Policy Approach to Tobacco Control Working?” *The Parliament Magazine*, 5 Aug. 2023, <https://www.parliamentmagazine.eu/partner/article/is-the-french-policy-approach-to-tobacco-control-working>.
- Pearlman, P. C., et al. “The National Institutes of Health Affordable Cancer Technologies Program: Improving Access to Resource-Appropriate Technologies for Cancer Detection, Diagnosis, Monitoring, and Treatment in Low- and Middle-Income Countries.” *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 4, 2016, pp. 1-8, <https://doi.org/10.1109/JTEHM.2016.2604485>.

- Porto, M. A., and P. A. Habib. “Viva Mulher: Constructing a Cervical Cancer Control Program in Brazil.” *Dynamis* (Granada, Spain), vol. 34, no. 1, 2014, pp. 101–123, <https://doi.org/10.4321/s0211-95362014000100006>.
- Prasad, J., and M. Dhar. “Risk of Major Cancers Associated with Various Forms of Tobacco Use in India: A Systematic Review and Meta-Analysis.” *Journal of Public Health (Berl.)*, vol. 27, 2019, pp. 803–813, <https://doi.org/10.1007/s10389-018-0992-7>.
- Sarma, S., et al. “Inequities in Access to Cancer Care in India: A Systematic Review.” *Indian Journal of Medical Research*, vol. 153, no. 5, 2021, pp. 627–640, [https://doi.org/10.4103/ijmr.IJMR\\_2723\\_20](https://doi.org/10.4103/ijmr.IJMR_2723_20).
- Sapkota, A., et al. “Smokeless Tobacco and Increased Risk of Hypopharyngeal and Laryngeal Cancers: A Multicentric Case-Control Study from India.” *International Journal of Cancer*, vol. 121, no. 8, 2007, pp. 1793–1798, <https://doi.org/10.1002/ijc.22832>.
- Radic, M., et al. “Tobacco Control Policies and Their Impact on Cancer Incidence: Insights from High-Income Countries.” *Global Health Action*, vol. 14, no. 1, 2021, <https://doi.org/10.1080/16549716.2021.1934982>.
- Shen, M., et al. “Therapeutic Targeting of Metadherin Suppresses Colorectal and Lung Cancer Progression and Metastasis.” *Cancer Research*, vol. 81, no. 4, 2021, pp. 1014–1025, <https://doi.org/10.1158/0008-5472.CAN-20-1876>.
- Smith, T. “A Long Way from Home: Access to Cancer Care for Rural Australians.” *Cancer Forum*, vol. 36, no. 1, Feb. 2012, pp. 42–46.
- Sokale, I. O., Raza, S. A., and A. P. Thrift. “Disparities in Cancer Mortality Patterns: A Comprehensive Examination of U.S. Rural and Urban Adults, 1999-2020.” *Cancer Medicine*, vol. 12, no. 18, 2023, pp. 18988–18998, <https://doi.org/10.1002/cam4.6451>.
- Tauras, J. A. “Tobacco Control in Low-Income and Middle-Income Countries: Findings from WHO FCTC Investment Cases.” *Tobacco Control*, vol. 33, Suppl. 1, 2024, pp. s1–s2, <https://doi.org/10.1136/tc-2024-058717>.
- “Tobacco and Cancer.” *The Lancet Oncology*, vol. 25, no. 8, 2024, pp. 1081–1082, [https://doi.org/10.1016/S1470-2045\(24\)00450-5](https://doi.org/10.1016/S1470-2045(24)00450-5).
- United Nations Regional Information Centre. “EU Air Pollution Remains a Major Health Risk.” 29 Sept. 2022, <https://unric.org/en/eu-air-pollution-remains-a-major-health-risk/>.
- Wang, M., et al. “AID Upmutants Isolated Using a High-Throughput Screen Highlight the Immunity/Cancer Balance Limiting DNA Deaminase Activity.” *Nature Structural & Molecular Biology*, vol. 16, no. 7, 2009, pp. 769–776, <https://doi.org/10.1038/nsmb.1623>.
- Wang, X., et al. “A 37-Year Observation of Mortality in Chinese Chrysotile Asbestos Workers.” *Thorax*, vol. 67, no. 2, 2012, pp. 106–110, <https://doi.org/10.1136/thoraxjnl-2011-200169>.
- Williams, A. “QALYs and Ethics: A Health Economist’s Perspective.” *Social Science & Medicine*, vol. 39, no. 5, 1994, pp. 627–634, [https://doi.org/10.1016/0277-9536\(94\)90035-3](https://doi.org/10.1016/0277-9536(94)90035-3).

- Williams, L. R., et al. "The Impact of Smoking on Cancer Outcomes: A Systematic Review of Recent Data." *Cancer Epidemiology*, vol. 81, 2022, 102079, <https://doi.org/10.1016/j.canep.2023.102079>.
- World Health Organization. "Global Action Plan for the Prevention and Control of Noncommunicable Diseases 2013-2020." 26 July 2022, <https://www.who.int/publications/i/item/9789241506236>.
- World Health Organization. "Tobacco and Its Impact on Public Health." 20 Sept. 2023, <https://www.who.int/health-topics/tobacco>.
- Wu, W. T., et al. "Cancer Incidence of Taiwanese Shipbreaking Workers Who Have Been Potentially Exposed to Asbestos." *Environmental Research*, vol. 132, 2014, pp. 370–378, <https://doi.org/10.1016/j.envres.2014.04.026>.
- Yang, M., et al. "Changing Trends and Disparities in 5-Year Overall Survival of Women with Invasive Breast Cancer in the United States, 1975-2015." *American Journal of Cancer Research*, vol. 11, no. 6, 2021, pp. 3201–3211.
- Yip, C. H., et al. "Advances in Cancer Care in Developing Countries: Challenges and Opportunities." *Cancer Science*, vol. 112, no. 12, 2021, pp. 4702–4712, <https://doi.org/10.1111/cas.15170>.
- Zhang, J., and H. Qin. "The Effects of Alcohol Consumption on Cancer Risk: A Meta-Analysis." *Journal of Cancer Research and Clinical Oncology*, vol. 144, no. 8, 2018, pp. 1637–1650, <https://doi.org/10.1007/s00432-018-2708-6>.
- Zhou, B., et al. "Trends in Tobacco Consumption and Its Implications for Cancer Incidence: A Review of Recent Data." *Cancer Epidemiology*, vol. 81, 2023, 102080, <https://doi.org/10.1016/j.canep.2023.102080>.

# **Empowering Small Businesses and Nonprofits: Effective Digital Marketing Strategies for Growth and Visibility By Aanya Mishra**

## **I. Abstract or Summary**

In the rapidly evolving digital landscape, small businesses face unique challenges in leveraging digital marketing strategies to achieve growth and competitiveness. This research explores the digital marketing landscape for small businesses, focusing on the barriers that prevent these businesses from fully adopting digital marketing strategies and offering actionable recommendations to overcome these challenges.

Small businesses are vital to economic growth, yet many struggle to compete with larger corporations, especially in digital marketing. Despite the potential benefits—such as increased brand awareness, customer engagement, and market reach—many small businesses do not fully utilize digital marketing, often due to financial constraints, lack of technical expertise, and time limitations. The misconception that digital marketing is reserved for larger enterprises also hinders adoption.

The paper integrates insights from existing literature and primary data collected through interviews with small businesses and nonprofits in the Bay Area. Findings indicate that while nine out of eleven organizations rely on digital marketing, especially content marketing and social media platforms like LinkedIn and Instagram, they face challenges in gaining traction, visibility, and allocating resources. However, these organizations also recognize the benefits of content creation, particularly in storytelling and visual content, to engage their audiences and support growth.

The research concludes with practical recommendations to empower small business owners with the knowledge and tools necessary to thrive in a digital-first world. By addressing these challenges and leveraging cost-effective digital marketing strategies, small businesses can enhance their competitiveness and sustainability in the digital era.

## **II. Introduction**

In the rapidly evolving digital age, the importance of digital marketing cannot be overstated. For small businesses, effective digital marketing strategies can be the difference between success and obscurity. Despite the clear advantages, many small businesses either do not fully leverage these strategies or avoid them altogether. This paper aims to explore the digital marketing landscape for small businesses, identify the barriers that prevent these businesses from adopting digital marketing strategies, and provide actionable recommendations to overcome these challenges.

Small businesses form the backbone of many economies around the world, contributing significantly to employment and economic growth. According to the U.S. Small Business Administration, small businesses accounted for 99.9% of all businesses in the United States in 2021, employing nearly half of the nation's workforce (U.S. Small Business Administration, 2021). Despite their critical role, small businesses often struggle to compete with larger

corporations, particularly in the digital arena. Digital marketing offers a level playing field where small businesses can effectively reach their target audience, build brand awareness, and drive sales. However, a study by Clutch revealed that 36% of small businesses do not have a website, and 35% of small businesses feel they are too small to warrant an online presence (Clutch, 2021). Several factors contribute to the reluctance of small businesses to adopt digital marketing strategies. Limited financial resources, lack of technical expertise, and time constraints are often cited as significant barriers. Moreover, many small business owners may not fully understand the potential return on investment that digital marketing can offer. Research from Deloitte highlights that digitally advanced small businesses are three times more likely to create jobs and are also 50% more likely to increase their revenue (Deloitte, 2021). These statistics underscore the critical need for small businesses to embrace digital marketing to remain competitive and sustainable in the modern marketplace.

Financial constraints often mean that small businesses cannot afford to hire dedicated digital marketing professionals or agencies, leading to a reliance on traditional marketing methods that may no longer be effective. Technical expertise is another significant barrier, as many small business owners lack the knowledge and skills to navigate the complexities of digital marketing platforms, analytics, and tools. Additionally, time constraints play a crucial role; small business owners often juggle multiple responsibilities, leaving little time to focus on developing and implementing digital marketing strategies.

Furthermore, there is a pervasive misconception among small business owners that digital marketing is only for larger enterprises with substantial budgets. This misunderstanding can result in missed opportunities for growth and development. For instance, social media platforms like Facebook and Instagram offer targeted advertising options that are both affordable and effective for small businesses. Similarly, email marketing and search engine optimization (SEO) provide cost-effective ways to reach potential customers and build lasting relationships. This paper is structured to provide a comprehensive understanding of small business digital marketing strategies. The first section will delve into the current state of digital marketing adoption among small businesses, examining the various strategies employed and their effectiveness. The second section will identify the primary barriers preventing small businesses from adopting digital marketing, supported by recent data and case studies. The third section will offer practical recommendations for overcoming these barriers, including cost-effective solutions and resources for acquiring digital marketing skills. Finally, the paper will conclude with a summary of key findings and a discussion on the future outlook for small businesses in the digital marketing landscape.

The primary goals of this paper are to shed light on the digital marketing challenges faced by small businesses, provide insights into effective strategies, and empower small business owners with the knowledge and tools necessary to thrive in a digital-first world. By addressing these objectives, this research aims to contribute to the broader discourse on small business sustainability and growth in the digital era.



### III. Literature Review

In today's digital age, digital marketing has become indispensable for businesses of all sizes, playing a pivotal role in enhancing market reach, customer engagement, and overall business growth. This section presents a comprehensive review of existing literature on digital marketing strategies tailored to small businesses, highlighting key findings, identifying gaps, and emphasizing the unique contributions of this paper.

Digital marketing encompasses a diverse array of strategies leveraging digital channels such as websites, social media platforms, search engines, email marketing, and SEO to attract, engage, and convert prospects into loyal customers. Small businesses, in particular, stand to benefit significantly from these strategies due to their agility, ability to target niche markets, and flexibility in adapting to market changes.

Research has extensively explored how digital marketing adoption influences small business performance across various contexts. For instance, "The Effect of Digital Marketing Adoption on Small Businesses Sustainable Growth: Empirical Evidence from Ghana" employs the theory of planned behavior to demonstrate that digital marketing enhances small businesses' sustainable growth by improving brand visibility, customer relationships, and operational efficiency (MDPI). Similarly, studies such as "Social Media Marketing Strategies Used by Small Retail Businesses" delve into specific strategies employed by small retailers on platforms like Facebook and Instagram to enhance customer engagement and drive sales, highlighting the strategic use of social media in fostering brand loyalty and generating direct sales (CORE). The existing literature converges on several key benefits of digital marketing for small businesses. These include enhanced brand awareness, increased customer engagement, improved customer service, and greater market reach at lower costs compared to traditional marketing methods. Studies consistently affirm that small businesses leveraging digital marketing tools effectively can achieve sustainable growth and competitive advantage in their respective markets (The Effect of Social Media Marketing on Small Business Performance, Novelty Journals). Despite the wealth of research, gaps persist in understanding the specific challenges that hinder small businesses from fully embracing digital marketing strategies. Limited digital literacy among business owners, financial constraints, and the dynamic nature of digital platforms are often cited as barriers that require deeper exploration (The Impact of Social Media Usage on the Sales Process in Small and Medium Enterprises, Core). Additionally, a broader focus beyond social media—including SEO, email marketing, and PPC—would deepen our understanding of the diverse strategies small businesses can employ for comprehensive digital growth. Recent studies on post-pandemic digital marketing trends and the rising use of automation tools also merit inclusion, as they reveal how small businesses can adapt to new technological advances and shifting consumer behaviors in a post-COVID landscape.

To ensure clarity, it is essential to define key terms often encountered in discussions of digital marketing strategies for small businesses:

- **SMEs:** Small and Medium Enterprises, typically characterized by their size and revenue thresholds, vary by country and industry.
- **Digital Marketing:** Marketing strategies and tactics that utilize digital technologies such as the internet, mobile devices, social media, and search engines to reach and engage with potential customers.
- **PPC (Pay-per-click):** An online advertising model where advertisers pay a fee each time their ad is clicked. Common PPC platforms include Google Ads and Bing Ads.
- **SEO (Search Engine Optimization):** The practice of optimizing web pages and content to rank higher in search engine results pages (SERPs), increasing visibility and organic traffic.
- **Social Media:** Online platforms and websites that enable users to create and share content or participate in social networking. Popular platforms include Facebook, Instagram, LinkedIn, Twitter, TikTok, and others.

This paper distinguishes itself by focusing on filling existing gaps in the literature regarding the barriers to digital marketing adoption among small businesses. It aims to provide a nuanced analysis of these challenges and offer practical, evidence-based recommendations for overcoming them. Additionally, the paper integrates insights from diverse digital marketing platforms such as LinkedIn, Instagram, TikTok, and beyond, including strategies like SEO and email marketing to offer a comprehensive overview of options available to small businesses. Studies like "The Effect of Social Media Marketing on Small Business Performance" and "Social Media Marketing Strategies Used by Small Retail Businesses" have contributed to establishing a consensus on the transformative impact of digital marketing on small businesses. They highlight the effectiveness of social media and other digital channels in improving brand visibility, customer engagement, and overall business performance.

This paper contributes to the existing literature by providing a comprehensive synthesis of current knowledge on digital marketing strategies for small businesses while focusing on overlooked challenges and barriers. By offering practical recommendations grounded in empirical evidence, recent post-pandemic trends, and best practices, this study aims to empower small businesses to navigate and capitalize on the digital landscape effectively. Ultimately, the insights derived from this research aim to support business owners in fostering a conducive environment for small business growth and competitiveness in the digital era.

#### **IV. Methods**

The methodology for this study involved a comprehensive literature review and primary data collection through interviews. For the literature review, various sources were utilized to gather relevant information on digital marketing for small businesses. Key sources included academic journals, industry reports, government publications, and reputable business websites. Search terms such as "small business digital marketing," "digital marketing strategies," "barriers to digital marketing adoption," "digital marketing ROI," and "technological challenges for small

businesses" were employed. Databases like Google Scholar, JSTOR, ProQuest, and websites of organizations such as the U.S. Small Business Administration (SBA), Deloitte Insights, and Clutch were consulted to extract pertinent data and insights.

Secondary data was sourced from existing studies and surveys. The U.S. Small Business Administration provided statistics on small business contributions to the economy and employment. Deloitte offered insights into the impact of digital transformation on small businesses. Clutch provided survey data on small businesses' digital presence and marketing practices.

Primary data was collected through interviews with small business owners and representatives from student-run nonprofits in the Bay Area. A purposive sampling method ensured a diverse representation of businesses across various industries and sizes, including those with fewer than 50 employees and from sectors such as retail, services, hospitality, and nonprofits. Potential interviewees were contacted via email and phone calls, introducing the research, its purpose, and inviting participation in a 30-45 minute interview, either in person or virtually. To ensure thorough responses, the interview questions were sent to the nonprofits in advance, allowing them to prepare detailed answers.

Interviews were conducted using a semi-structured format, allowing flexibility while covering key topics. Core questions included inquiries about the business and its digital marketing efforts, implemented strategies and their effectiveness, barriers to adopting digital marketing, success measurement, needed resources or support, success stories, challenges, and future outlook on digital marketing for small businesses. The data collected from these interviews was transcribed and analyzed using thematic analysis to identify common themes and patterns related to digital marketing adoption, barriers, and potential solutions. The findings from the literature review were integrated with the primary data to provide a comprehensive understanding of the topic.

This methodology outlines a detailed approach to researching digital marketing strategies for small businesses. By combining secondary data from reputable sources with primary data from interviews, this study aims to offer valuable insights and practical recommendations applicable to similar contexts.

## **V. Results**

This section presents the findings from the study on digital marketing strategies among Bay Area non-profits and small businesses. The data were collected through surveys and interviews with eleven organizations: Code for Cause, BoxMates, BISV DECA, Xeno, STEM Hub, Simple Tutor, Infinite Inspiration Initiative, Blue Flag Security, Project Hoops, Mental Health in Youth, and Med Set Go. The results are categorized into quantitative data on social media usage and qualitative insights into the challenges and benefits experienced by these organizations. The analysis is supported by relevant literature to contextualize the findings. To gather comprehensive data, a mixed-methods approach was employed. Quantitative data were collected through an online survey distributed to various non-profits and small businesses in the

Bay Area. The survey included questions about the frequency and platforms of social media use, engagement strategies, and the effectiveness of various digital marketing channels. Additionally, in-depth interviews were conducted with key personnel from all eleven organizations mentioned above. These interviews provided qualitative data, offering deeper insights into the organizations' digital marketing strategies, challenges, and future plans.

One of the most prominent themes from the interviews was the importance of maintaining a strong digital presence. Many organizations emphasized the necessity of having a user-friendly website and active social media profiles. For example, Code for Cause highlighted their investment in a comprehensive website redesign, which led to increased user engagement and improved visitor metrics. Similarly, Xeno noted that their consistent social media activity helped build a loyal online community.

The use of social media platforms emerged as a critical component of digital marketing strategies. Different organizations utilized various platforms based on their target audiences. Nine out of the eleven businesses and nonprofits indicated that digital media is their main marketing tactic, with a particular emphasis on content marketing and social media platforms like LinkedIn and Instagram. For instance, Project Hoops focused heavily on Instagram and TikTok to reach younger demographics, while Blue Flag Security utilized LinkedIn to connect with professionals and industry peers. A common practice among these organizations was the use of targeted advertising and content tailored to platform-specific audiences. There is also a shared ambition to expand their digital marketing efforts further.

The preference for Instagram and LinkedIn reflects broader trends in social media marketing, where visually driven and professional networking platforms have gained prominence. The relatively lower usage of Facebook and Twitter suggests that these platforms may be less effective for targeted outreach in the non-profit and small business sectors. The high usage of social media can be attributed to its cost-effectiveness and the ability to reach a broad audience, making it an invaluable tool for organizations with limited budgets (Saxton & Guo, 2011).

The qualitative data, derived from interviews with representatives from the participating organizations, revealed several key themes related to the challenges and benefits of using social media.

A common challenge mentioned by several organizations, including Code for Cause and Blue Flag Security, was the difficulty in gaining traction and visibility. The competitive nature of social media, where numerous organizations compete for attention, was a significant hurdle. Aayna Dash from Code for Cause noted, "It's challenging to stand out in a crowded social media landscape, especially with limited resources." This sentiment was echoed by other interviewees, who mentioned the impact of constantly changing algorithms on platforms like Instagram and LinkedIn, which can affect the reach and visibility of posts (Capriotti & Kuklinski, 2019). Resource allocation also emerged as a critical barrier. Many organizations operate with small teams, making it challenging to maintain a consistent and engaging social media presence. For instance, a representative from Project Hoops mentioned, "We don't have a dedicated social

media team, so it's tough to keep up with regular posting and interacting with our followers." This lack of resources often leads to a reliance on volunteers or part-time staff, complicating efforts to establish a cohesive and professional online presence

Despite these challenges, content creation was widely recognized as a crucial component of a successful social media strategy. Organizations like Mental Health in Youth and Xeno reported that creating and sharing engaging content helped them connect with their audiences, raise awareness about their causes, and drive donations. Aayna Dash highlighted, "Creating impactful content has been a game-changer for us; it allows us to tell our story and engage with our community in meaningful ways."

Visual content, such as photos and videos, was particularly valued for its ability to capture attention and convey messages quickly. A representative from Xeno noted, "Our most successful posts are usually videos or photos that tell a story. They get more engagement than text-based updates." This preference for visual content is supported by research showing that images and videos tend to generate higher engagement rates on social media platforms (Sabate et al., 2014).

The study reveals that a significant majority of Bay Area nonprofits and small businesses actively utilize social media, with Instagram and LinkedIn emerging as the most popular platforms. This widespread adoption underscores the importance of social media as a crucial tool for outreach and engagement. However, despite the high usage, many organizations face challenges in gaining visibility and fostering engagement, primarily due to the competitive nature of social media and limited resources. On the positive side, creating high-quality, engaging content is recognized as an effective strategy for building connections with audiences and supporting fundraising efforts. The emphasis on storytelling and visual content aligns with broader trends in social media marketing. These key findings set the stage for a more in-depth analysis in the subsequent Discussion/Analysis section, where the implications will be explored in detail, particularly regarding best practices for nonprofits and small businesses looking to enhance their digital marketing strategies.

## **VI. Discussion/Analysis**

The primary objective of this research was to explore the digital marketing strategies employed by nonprofits and small businesses in the Bay Area, with a particular focus on the usage and effectiveness of social media platforms. The study aimed to identify the most commonly used social media platforms, understand the challenges these organizations face in gaining traction, and assess the benefits of content creation. In doing so, the research also sought to draw comparisons with existing literature to provide a comprehensive understanding of the role of social media in enhancing organizational success and outreach.

The study involved in-depth interviews with eleven Bay Area nonprofits and small businesses: Code for Cause, BoxMates, BISV DECA, Xeno, STEM Hub, Simple Tutor, Infinite Inspiration Initiative, Blue Flag Security, Project Hoops, Mental Health in Youth, and Med Set Go. The key findings reveal a high adoption rate of social media among these organizations, with

Instagram and LinkedIn being the most popular platforms. Commonly reported challenges included gaining visibility in a competitive digital landscape and dealing with resource limitations. However, the benefits of content creation, particularly storytelling and visual content, were highlighted as effective strategies for engagement and fundraising. These insights align with existing literature, emphasizing the cost-effectiveness and wide-reaching potential of social media for organizations operating with limited budgets.

For nonprofits in the Bay Area, the study's findings suggest several practical recommendations. First, prioritizing visually driven platforms like Instagram can significantly enhance engagement, especially when content is tailored to tell compelling stories. Nonprofits should invest in developing a consistent and cohesive visual identity across their social media channels to build brand recognition and loyalty. Additionally, leveraging LinkedIn for professional networking can help organizations connect with potential donors, partners, and volunteers, making it an essential platform for growth. Given the challenges of limited resources, nonprofits should consider adopting a strategic approach to content creation—focusing on quality over quantity and utilizing tools like content calendars to maintain a regular posting schedule without overwhelming their teams. Finally, staying informed about platform algorithm changes and adapting strategies accordingly will be crucial in maintaining visibility and engagement in an ever-evolving digital landscape.

## **VII. Conclusion**

This paper aims to answer the question: What are the key barriers preventing small businesses from adopting digital marketing strategies, and how can these barriers be effectively overcome? Through comprehensive research, including interviews with small businesses and nonprofits, this study identifies limited financial resources, lack of technical expertise, and time constraints as the primary challenges. The concrete takeaways include actionable recommendations for small businesses, such as cost-effective digital marketing solutions, strategies for acquiring necessary skills, methods for optimizing time management, and financial solutions, including specific funding strategies and partnerships, to help nonprofits enhance their digital presence and competitiveness.

To gather insights for this research, we conducted interviews with representatives from eleven different nonprofits and small businesses in the Bay Area. These organizations included a diverse range of sectors, from education-focused initiatives to security services, offering a broad perspective on digital marketing practices. The interview process was designed to explore each organization's unique experiences with digital marketing, the strategies they employed, and the challenges they encountered. By using a semi-structured format, we were able to maintain consistency across interviews while allowing for in-depth discussions on topics most relevant to each participant.

The interviews revealed that most of these organizations heavily rely on digital marketing as their primary outreach tool, particularly through social media platforms like Instagram and

LinkedIn. Despite this reliance, many participants expressed frustration with the difficulties of maintaining an effective digital presence. Common issues included limited budgets, which prevented them from hiring dedicated marketing staff, and the fast-paced nature of digital platforms, which made it challenging to stay current with trends and best practices. These conversations highlighted the need for more accessible resources, training, and financial solutions tailored to help small businesses and nonprofits navigate the complexities of digital marketing.

Analyzing the responses from these interviews provided valuable qualitative data that enriched our understanding of the real-world challenges faced by small businesses and nonprofits in the digital age. The recurring themes of resource limitations and the struggle to keep up with digital trends underscored the importance of practical, tailored solutions that address these specific barriers. The findings from these interviews were instrumental in shaping the actionable recommendations presented in this paper, ensuring they are grounded in the actual experiences and needs of small business owners and nonprofit leaders.

### **Works Cited**

- Capriotti, P., & Kuklinski, H. P. (2019). Managing public relations in the digital era: new strategies for social media engagement. *Public Relations Review*, 45(4), 101829.
- Clutch. (2021). *Small Business Website Statistics: 2021 Survey*. Retrieved from Clutch.
- Deloitte. (2021). *Digital Transformation: How Digitally Advanced Small Businesses Drive Growth and Job Creation*. Deloitte Insights.
- MDPI. (n.d.). *The Effect of Digital Marketing Adoption on Small Businesses' Sustainable Growth: Empirical Evidence from Ghana*. Retrieved from MDPI.
- Saxton, G. D., & Guo, C. (2011). Accountability online: Understanding the web-based accountability practices of nonprofit organizations. *Nonprofit and Voluntary Sector Quarterly*, 40(2), 270-295.
- The Effect of Social Media Marketing on Small Business Performance*. (n.d.). Novelty Journals. Retrieved from Novelty Journals.
- The Impact of Social Media Usage on the Sales Process in Small and Medium Enterprises*. (n.d.). CORE. Retrieved from CORE.
- U.S. Small Business Administration. (2021). *2021 Small Business Profile*. Retrieved from SBA.



## **Truth is Marching On: The Evolution of Afro-American Spirituals and Their Impact on the American Civil Rights Movement By Matthew Lee**

During a grand military review in November, 1861 at Bailey's Crossroads in Virginia, an estimated 75,000 Union soldiers witnessed the unifying power of "John Brown's Body," spreading religious and military fervor throughout the country (Gamble). Among them was Julia Ward Howe, an abolitionist author and poet who, inspired by the song's rallying influence, wrote the lyrics to "Mine Eyes Have Seen The Glory of The Coming Lord," later known as "Battle Hymn of the Republic" (Stauffer). Through its complex evolution, "Battle Hymn" motivated the Northern Civil War effort towards a victory that changed the course of history. In addition to unifying Union soldiers during the Civil War, the "Battle Hymn" helped establish an inclusive tradition of American civil rights movements' protest music that reached an influential peak in the 1960's. "Battle Hymn of the Republic's" historic roots germinate from Afro-American spirituals that shape its message of hope.

The enslaved embedded their messages within the African call-and-response structure, reinforcing their sense of community and individual worth, along with sustaining hope subtly and powerfully (Charlotte Lawrence-McIntyre 379-401). The spiritual "Say Brothers Will You Meet Us,"—first published in an 1807 North Carolina hymnbook and sung during Southern Afro-American Christian camp meetings—embraced a hopeful and celebratory tone, focusing on the anticipation of a joyous reunion in the afterlife (Johnson). The opening lines, "Say, brothers will you meet us," serve as an invitation to join in solidarity on "Canaan's happy shore." The spiritual's biblical references—such as the refrain of "Glory, glory, hallelujah"—emphasize the promise of salvation and Providence under God ("Say Brothers Will You Meet Us"). Furthermore, the spiritual "I Got Shoes" declares, "I got shoes, you got shoes, All God's children got shoes. When I get to Heaven gonna put on my shoes, Gonna walk all over God's Heaven" ("I Got Shoes"). In this line, the enslaved transcends the earthly bonds of slavery, illustrating aspirations for a heavenly existence where freedom—"walk all over God's heaven"—and equality—"all God's children got shoes"—are unconditionally afforded to all. Surprisingly, enslavers often misconstrued these songs, believing that "Blacks who sang spirituals worked more efficiently and obediently" and even forcing them to perform in front of company (Lawrence-McIntyre 383). How could they know that the melody of "Say Brothers Will You Meet Us" would later be adapted and repurposed for the iconic abolitionist anthem, "John Brown's Body," which in turn lent the melody to the even more iconic "Battle Hymn?"

In the middle of the nineteenth century, spirituals even encompassed directions to the Underground Railroad, as with "Follow the Drinking Gourd." The "drinkin' gourd" is a coded reference to the Big Dipper constellation containing the North Star, specifically its ladle-shaped configuration (Kelley). It served as a means of navigating escape routes to freedom, often using natural references like the sun, quail calls, and riverbanks. The converted message of the repeated phrase "follow the drinkin' gourd" embodies the promise of emancipation and provides an auditory roadmap toward liberty (Kelley 262-280). These songs were instrumental in

fostering a covert communication network that eluded enslavers, demonstrating continuous resistance and courage in the face of oppression.

Some spirituals were less subtle than others. "Joshua Fit the Battle of Jericho" was believed to have emerged much earlier than its copyrighted date in 1865 (*"Joshua Fit the Battle of Jericho"*). Its feverous lyrics employ the biblical story of the crumbling walls of Jericho as a metaphor for the collapse of seemingly insurmountable obstacles (Lawrence-McIntyre, 396). After six days of marching around the enemy city of Jericho, the Israelite priests blew their trumpets and the people "shouted with a great shout, the wall fell down flat," achieving a military victory (ESV Bible, Joshua 6.20). This is an example of the call-and-response format from the Old Testament: the trumpets call and the people shout in response. The narrative of Joshua's conquest of Jericho embodies biblical triumph through music preceding military might, as if to foreshadow the national fervor of "Battle Hymn" and the soldiers marching to it. This shift in focus reflects the evolving social and political landscape of mid-1800s, as the abolitionist movement gained momentum and began to assert itself more overtly.

On October 16, 1859 John Brown and a group of abolitionists captured citizens of Harper's Ferry and seized the federal armory and arsenal. The standoff lasted two days, and John Brown was hanged as a result (Foner 37-52). His legacy was immortalized in the lyrics of "John Brown's Body," which adopts a militant and political tone by directly referencing his violent efforts along with the progressing abolitionist movement. The opening lines, "Old John Brown's body lies moldering in the grave, While weep the sons of bondage whom he ventured all to save," depict John Brown as a martyr for emancipation ("John Brown's Body"). Beyond the grave, the repeated phrase "His soul is marching on," emphasizes the legacy of Brown's action and the continuing struggle for freedom ("John Brown's Body"). The imagery of heaven looking down on the Union army, symbolized by its "flag red, white, and blue," suggests Providence supporting the North. The anthem concludes with a call to action for "soldiers of freedom" to continue the fight against oppression, inspired by the shining example of John Brown ("John Brown's Body"). By appropriating the melody of a spiritual hymn for a song with explicit political and militant themes, the adaptation bridged the gap between religious and secular music while simultaneously collecting support for the abolitionist cause.

Following that historic grand military review and rousing chorus in Virginia, Julia Ward Howe made explicit the coded messages of resistance by penning the "Battle Hymn of the Republic" (Gamble). The Hymn became popular beyond belief as it was popularized in lecture halls, pulpits, and leading periodicals for children and adults (Gamble). Even President Lincoln was moved to tears when he first heard "The Battle Hymn of the Republic," sung by Chaplain McCabe, a survivor of an infamous Confederate prison in Virginia: Libby (Zombek). Having read Howe's lyrics in *The Atlantic Monthly*, McCabe sang it at a war meeting and carried the song to the front lines, where its impact continued to bolster morale through its lyrics and rhythm among the troops ("The First Night in Libby"). The song soon evolved into a powerful rallying cry for Union troops, exemplifying the religious nationalism that swept throughout the country.

"Battle Hymn of the Republic" was a crucial link in the evolution of coded messages of resistance and a communal narrative seeking liberation. The song evolved from the "continuums of various group performances" on Afro-American spirituals in that "each [new] participant contributed a bit of harmony, a small rhythmic device, a shade of vocal color" (Lawrence-McIntyre 382). Lines like "Mine eyes have seen the glory of the coming of the Lord" offered the soldiers a sense of righteousness and divine purpose in their fight against a morally reprehensible institution. The line "In the beauty of the lilies, Christ was born across the sea, With a glory in his bosom that transfigures you and me" references the contrast between a free North and a South clinging to slavery (Howe). This connection to the religious spirit of Afro-American spirituals, steeped in themes of hope and perseverance, tapped into the collective consciousness of a nation torn by conflict (Lawrence-McIntyre 382). Where "John Brown's Body" took spirituals from the subtle to overt challenges of the Civil War era, "Battle Hymn" raised their trajectory by synthesizing religious imagery with the harsh realities of conflict. Through vivid descriptions of divine intervention and impending judgment, "Battle Hymn" painted a role for God in the ongoing struggle for righteousness. In the crucible of social upheaval, music became a potent tool for dissent, challenging societal norms and motivating communities to action.

During the Civil Rights Movement of the 1960s, the transformative power of spirituals and the Battle Hymn rose to further prominence, stimulating the spirit of resistance and unity that music had sparked during the Civil War. Just as spirituals had provided a source of strength and secretive means of communication among enslaved Afro-Americans, they now fortified a new generation fighting against systemic racial injustices. The movement harnessed the profound emotional and moral resonance of these songs to galvanize support and articulate the struggles and aspirations of Afro-Americans (Richer). One of the most important examples is Nina Simone, who encouraged Afro-American youth to embrace their identity in the face of adversity in her rendition of "To Be Young, Gifted and Black" (Holden). Her concerts often transformed into rallies, where songs transcended entertainment to become powerful declarations of resistance and calls to action (Holden).

The role of spirituals and other freedom songs in the civil rights movement served not only as spiritual sustenance but also as tactical support during protests and marches. These songs, rooted in the Afro-American experience and adapted to reflect the ongoing struggle for equality, provided a unifying force among activists. For instance, the transformation of "Joshua Fit the Battle of Jericho" into "Marchin' Round Selma like Jericho" illustrates how these spirituals were adapted to address contemporary issues like segregation and voting rights. The title specifically references the 1965 Selma to Montgomery marches for voting rights. Its lyrics, "To the Freedom Fighters call / Black, Brown and White American say / Segregation must fall," appropriate biblical imagery to depict the collective action of protesters marching in Selma, Alabama, demanding equality and justice ("Civil Rights Marchers Attacked in Selma"). By linking the biblical story of Jericho's falling walls with the physical and metaphorical barriers of

segregation, the adapted lyrics underscore the ongoing fight for civil rights and the dismantling of systemic oppression.

Additionally, the communal nature of singing these spirituals, whether in government malls, city streets, or university campuses, created a powerful atmosphere of solidarity. More importantly, it was a practical strategy for organizing and energizing the movement's participants, reminding them of their shared goals and reinforcing their commitment to achieving racial equality (Richer). These songs became anthems of resilience, echoing through the streets and squares where activists gathered, instilling a deep sense of community and reinforcing collective identification with the movement's objectives.

Martin Luther King Jr., in his final speech before his assassination in 1968, also drew inspiration from these songs, notably closing with the opening lines of "Battle Hymn," "Mine eyes have seen the glory of the coming of the Lord," demonstrating the enduring significance of these melodies in galvanizing activists (Limpong). Furthermore, King believed that songs like "Ain't Gonna Let Nobody Turn Me 'Round" not only provided emotional uplift but also fortified activists' resolve in the face of brutal opposition, helping them to stand firm and march together. These songs functioned as "organizational glue," binding the community together in solidarity and shared purpose (Hsiung, 23). Songs like "We Shall Overcome," directly adapted from a gospel hymn, echoed the structure and hopeful spirit of the "Battle Hymn." Its iconic refrain, "We shall overcome," mirroring the "Battle Hymn's" "His truth is marching on," resonated with the message of unwavering determination in the face of oppression.

Spirituals link the historical pain and triumphs of Afro-Americans to the ongoing fight against oppression: a lineage of musical activism that eventually found expression in the "Battle Hymn of the Republic," which, one hundred years later, inspired civil rights protesters. Its lineage and legacy underscores the enduring power of music to uplift and unite communities under a common cause. Still sung world round, "Battle Hymn" stands as a testament to the enduring human spirit, echoing the hopes, fears, and aspirations of generations past. (Hsiung, 23)

## Works Cited

- Foner, Eric. "The Crisis of Sectionalism." *Reconstruction: America's Unfinished Revolution*, W.W. Norton & Company, 2018, pp. 37-52.
- Gamble, Richard M. "'Let Us Die to Make Men Free.'" *American Heritage*, 1 June 2024, [www.americanheritage.com/let-us-die-make-men-free](http://www.americanheritage.com/let-us-die-make-men-free).
- Holden, Stephen. "Music: Nina Simone Sings." *The New York Times*, 11 Mar. 1985, [www.nytimes.com/1985/03/11/arts/music-nina-simone-sings.html?searchResultPosition=11](http://www.nytimes.com/1985/03/11/arts/music-nina-simone-sings.html?searchResultPosition=11).
- Howe, Julia Ward. *Battle Hymn of the Republic*. *The Atlantic Monthly*, 1862.
- Hsiung, David C. "Freedom Songs and the Modern Civil Rights Movement." *OAH Magazine of History*, vol. 19, no. 4, July 2005, pp. 23-26. Oxford University Press on behalf of the Organization of American Historians, [www.jstor.org/stable/25161959](http://www.jstor.org/stable/25161959).
- Johnson, J. H. *John Brown's Original Marching Song: Tune--Brothers, Will You Meet Me*. Johnson, between 1861 and 1876, [www.id.lib.harvard.edu/alma/990133264850203941/catalog](http://www.id.lib.harvard.edu/alma/990133264850203941/catalog).
- Jones, Arthur C. "The Foundational Influence of Spirituals in African-American Culture: A Psychological Perspective." *Black Music Research Journal*, vol. 24, no. 2, Autumn 2004, pp. 251-260. Center for Black Music Research - Columbia College Chicago and University of Illinois Press, [www.jstor.org/stable/4145493](http://www.jstor.org/stable/4145493).
- Kelley, James B. "Song, Story, or History: Resisting Claims of a Coded Message in the African American Spiritual 'Follow the Drinking Gourd.'" *The Journal of Popular Culture*, vol. 41, no. 2, 2008, pp. 262-280. Mississippi State University, [www.scholarsjunction.msstate.edu/msu-meridian-publications](http://www.scholarsjunction.msstate.edu/msu-meridian-publications).
- Lawrence-McIntyre, Charshee Charlotte. "The Double Meanings of the Spirituals." *Journal of Black Studies*, vol. 17, no. 4, June 1987, pp. 379-401. Sage Publications, Inc., [www.jstor.org/stable/2784158](http://www.jstor.org/stable/2784158).
- Limbong, Andrew. "One Song Glory." NPR, 4 July 2018, [www.npr.org/2018/07/04/625351953/one-song-glory](http://www.npr.org/2018/07/04/625351953/one-song-glory).
- Richer, Stephen. "The Protest Legacy of African American Spirituals, Part 2." *The Journal of Wild Culture*, 20 Feb. 2022, [www.wildculture.com/article/protest-legacy-african-american-spirituals-part-2/1953](http://www.wildculture.com/article/protest-legacy-african-american-spirituals-part-2/1953).
- Stauffer, John. *Battle Hymn of the Republic*. C-Span, 15 Nov. 2014, [www.c-span.org/video/?322562-3/battle-hymn-republic](http://www.c-span.org/video/?322562-3/battle-hymn-republic).
- "The First Night in Libby; Singing Mrs. Howe's Battle Hymn of the Republic. Turning out for a Roll Call and Cooking Breakfast -- Some of the Celebrities of the Prison -- How Two Captains Escaped Hanging." *The New York Times*, 15 Feb. 1891, [www.nytimes.com/1891/02/15/archives/the-first-night-in-libby-singing-mrs-howes-battle-hymn-of-the.html](http://www.nytimes.com/1891/02/15/archives/the-first-night-in-libby-singing-mrs-howes-battle-hymn-of-the.html).
- The Learning Network. "March 7, 1965 | Civil Rights Marchers Attacked in Selma." *The New York Times*, 7 Mar. 2012, [archive.nytimes.com/learning.blogs.nytimes.com/2012/03/07/march-7-1965-civil-rights-marchers-attacked-in-selma/](http://archive.nytimes.com/learning.blogs.nytimes.com/2012/03/07/march-7-1965-civil-rights-marchers-attacked-in-selma/).
- Zombek, Angela M. "Libby Prison." *Encyclopedia Virginia*, 7 Dec. 2020, [www.encyclopediaivirginia.org/entries/libby-prison/](http://www.encyclopediaivirginia.org/entries/libby-prison/).

# Mediapipe Fingering Classification for Violin By Spencer Wang

## Abstract

The extent of the potential of Mediapipe in analyzing and classifying complex hand patterns, like those found in violin hand movements, remains elusive. This research explores the feasibility of using Mediapipe, an open-source framework that allows the extraction of finger position key points from images/videos, to detect violin fingering positions for playing a virtual instrument. Further, we explore the limitations and potential of Mediapipe for recognizing complex hand movements involved in playing the violin by using a combination of Google Colab and support vector classification models. The repeated training on misclassified images was able to take the F1 score from an initial value of 0.82 to a final value of 0.95. This initial research shows the potential for a video based virtual violin instrument. Further areas of exploration could include expanding the total image training data and testing viability of other classifiers.

## 1) Introduction

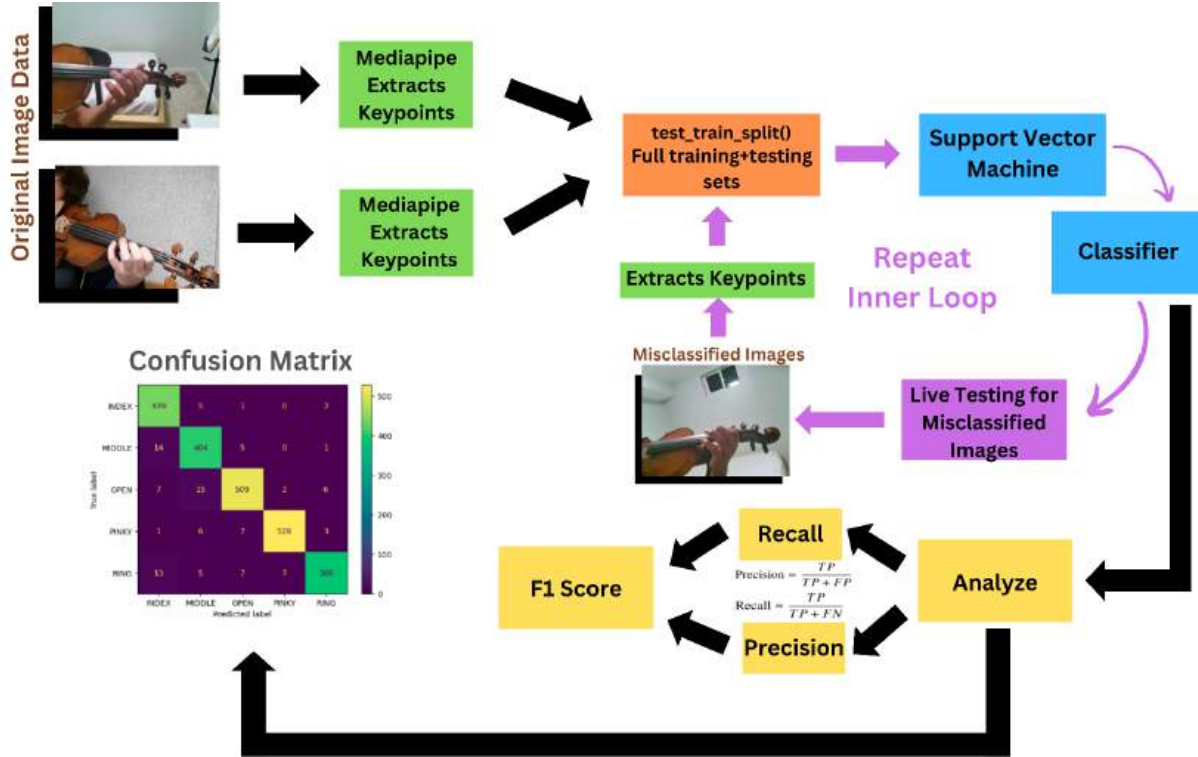
The intersection of artificial intelligence and music technology has opened up new possibilities for innovative applications, from music composition to performance [1-3]. One such application is the creation of a virtual musical instrument that can replicate the nuanced experience of playing a real instrument [4]. This vision involves three critical steps: accurately detecting the fingering positions, recognizing hand placement on the instrument, and capturing the bowing technique. This paper focuses on perfecting the fingering classification process, a fundamental aspect of this vision.

By achieving precise recognition of violin fingering positions through image recognition and machine learning, such as utilizing Mediapipe—an open-source framework developed by Google for real-time hand tracking and gesture recognition [5]—we can open the door to new forms of creative expression and accessibility in music. A virtual instrument capable of mimicking these complex gestures would not only enable musicians to practice and perform without the physical instrument but also allow for new types of musical expression that blend digital and physical realms. This endeavor is worth pursuing because it merges technology with artistic creativity, providing musicians with a powerful tool that expands their creative possibilities while also making music more accessible to a broader audience.

Despite the potential of Mediapipe, its effectiveness in recognizing complex and precise movements, such as those required for playing the violin, remains largely unexplored. Traditional methods of training models for such tasks face challenges related to data collection, computational resources, and varying environmental conditions [6]. This research is a proof of concept that such training is possible for complex motions, like the violin, through just a simple webcam. As shown through Figure 1, we use mediapipe to extract keypoints, simplifying the computer vision process, and then use a support vector machine to classify the key points as this

is a technique that requires minimal data and resources as compared to traditional convolutional neural network and direct classification [7].

This study contributes to the field of AI-driven music technology by exploring a novel application of Mediapipe. The findings can inform future research and development of virtual musical instruments, enhancing both educational tools and performance aids.



**Figure 1:** Overview of algorithm and data processing for Classifying Violin Fingering Positions Using Image Data.

Camera data is collected which Mediapipe then extracts key point data showing the finger positions. A support vector machine classifier is trained. The classifier is tested on live images and any images that are misclassified go back into the training and testing sets. After augmenting the training set with misclassified images, a F1 score of 0.95 is achieved.

### 1.1) Objectives

1. Assess the capability of Mediapipe in detecting violin fingering positions.
2. Evaluate the accuracy of a support vector classifier (SVC) trained on hand key points detected by Mediapipe.
3. Identify the limitations and challenges associated with using Mediapipe for this application.
4. Explore the potential for real-time application and performance under varied conditions.

## 2) Methodology

### 2.1) Tools and Environment

The tools and environment used in this project include software such as Mediapipe, Google Colab, OpenCV, and scikit-learn, along with a webcam for data collection.

## 2.2) Data Collection

### 2.2.1) Enabling Camera Usage in Google Colab

Open-CV does not directly support camera access in Google Colab. To bypass this, a JavaScript code was used to allow camera access in Google Colab [8].

### 2.2.2) Automated Photo Capture

We developed a script in Colab to automate the capture and saving of training photos to Google Drive. The script was designed to take photos by slowly waving the camera around, capturing different angles and positions. The specific camera used was the Polycom EagleEye Mini as it was small and could easily be moved around to take photos at different angles. All photos were taken indoors with the camera running at 30 frames per second and each image had the dimensions of 640 x 480 pixels.

## 2.3) Data Labeling

Each captured photo was labeled and saved, with Mediapipe detecting and recording the hand key points.

## 2.4) Data Processing

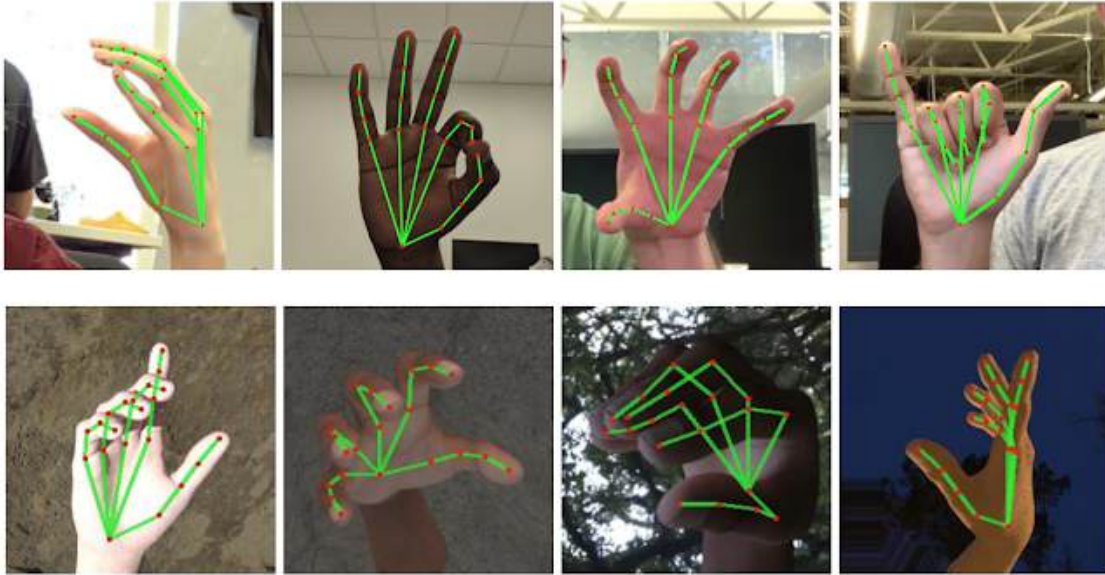
### 2.4.1) Keypoint Extraction

For each photo, Mediapipe was used to extract 16 hand key points (x, y coordinates) as shown in Figure 2. These 16 hand key points can be plotted on the respective image, as shown in Figure 3, showing what Mediapipe would “see” when analyzing a hand position. The extracted key points were saved to a text file for subsequent analysis.



Figure 2: Specific key points “name” and the index to find them when retrieving all data points[6].





**Figure 3:** Given an image of a hand, Mediapipe is able to take that image and extract the location of the key points on the hands [6]. What is shown here are examples of what Mediapipe would see to visualize key points on different hand shapes, lighting, and backgrounds

## 2.5) Model Training

### 2.5.1) Dataset Preparation

The dataset was split into training and testing sets using the `train\_test\_split` function from scikit-learn. The model parameters were fine-tuned to optimize performance (80% training set, 20% testing set, C value = 10).

### 2.5.2) Support Vector Machine

An SVM with an RBF kernel was trained on the training set. SVM typically performs well with high-dimensional and unstructured datasets, which works well in this situation with image data processing. One general downside for using a SVM is computational speed. However, this wasn't a significant issue in this scenario with a considerably low number of training data.

### 2.5.3) Evaluation

The trained model was evaluated using a confusion matrix to measure accuracy on the test set as well as the F1 score on the total combined test set. The F1 score is an evaluation metric for machine learning models that combines precision and recall, providing a balanced measure of accuracy. Unlike simple accuracy, which can be misleading on imbalanced datasets, the F1 score accounts for both false positives and false negatives\*, making it a more reliable metric when class distribution is uneven, which is especially useful during the misclassification backtrack training as the number of images per class gets unbalanced.

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

\*Note: TP = True Positive, FP = False Positive, FN = False Negative [9]

#### 2.5.4) Misclassification Backtrack Training

The trained model was then tested in real time and for every frame that the model misclassified, that frame is then stored and later inputted into the total training set to increase real-time accuracy.

#### 2.5.5) Original Issues

The very first model to test whether mediapipe could classify which finger was being pressed down achieved an accuracy of approximately 90%, likely due to the controlled conditions under which the data was collected (same person, environment, and lighting) with only roughly 200 data images per finger (only one person's image data). After collecting data images that contained a different person, lighting, and background, with roughly the same number of data images per finger, the updated model after training achieved an accuracy of around 94% on both subjects. However, when testing the model in real time classifying, there were discrepancies with the results that were given. Due to inconsistencies with relative data points and how data points were processed during the training process and the live testing process, depending on which set of photos the model was trained on, the model was only producing one output, regardless of the input photo that was on the live screen.

As shown in Figure 4, to fix these issues, each datapoint was normalized relative to the wrist key point (index 0 on the key points list) so that the wrist coordinate would always stay at  $(x, y) = (0, 0)$  and all other key points would be relative to that coordinate axis. The data points were further normalized to stay within the range of 0 to 1.

```
def normalize_landmarks(landmarks):
    wrist = landmarks[0]
    normalized_landmarks = []
    for landmark in landmarks:
        normalized_landmarks.append((landmark.x - wrist.x, landmark.y - wrist.y))

    scale = max(
        ((landmarks[0].x - landmarks[5].x)**2 + (landmarks[0].y - landmarks[5].y)**2)**0.5,
        ((landmarks[5].x - landmarks[17].x)**2 + (landmarks[5].y - landmarks[17].y)**2)**0.5,
        ((landmarks[17].x - landmarks[0].x)**2 + (landmarks[17].y - landmarks[0].y)**2)**0.5
    )
    normalized_landmarks = [(x/scale, y/scale) for x, y in normalized_landmarks]
    return normalized_landmarks
```

**Figure 4:** Python code used to normalize all key points relative to the wrist key point.

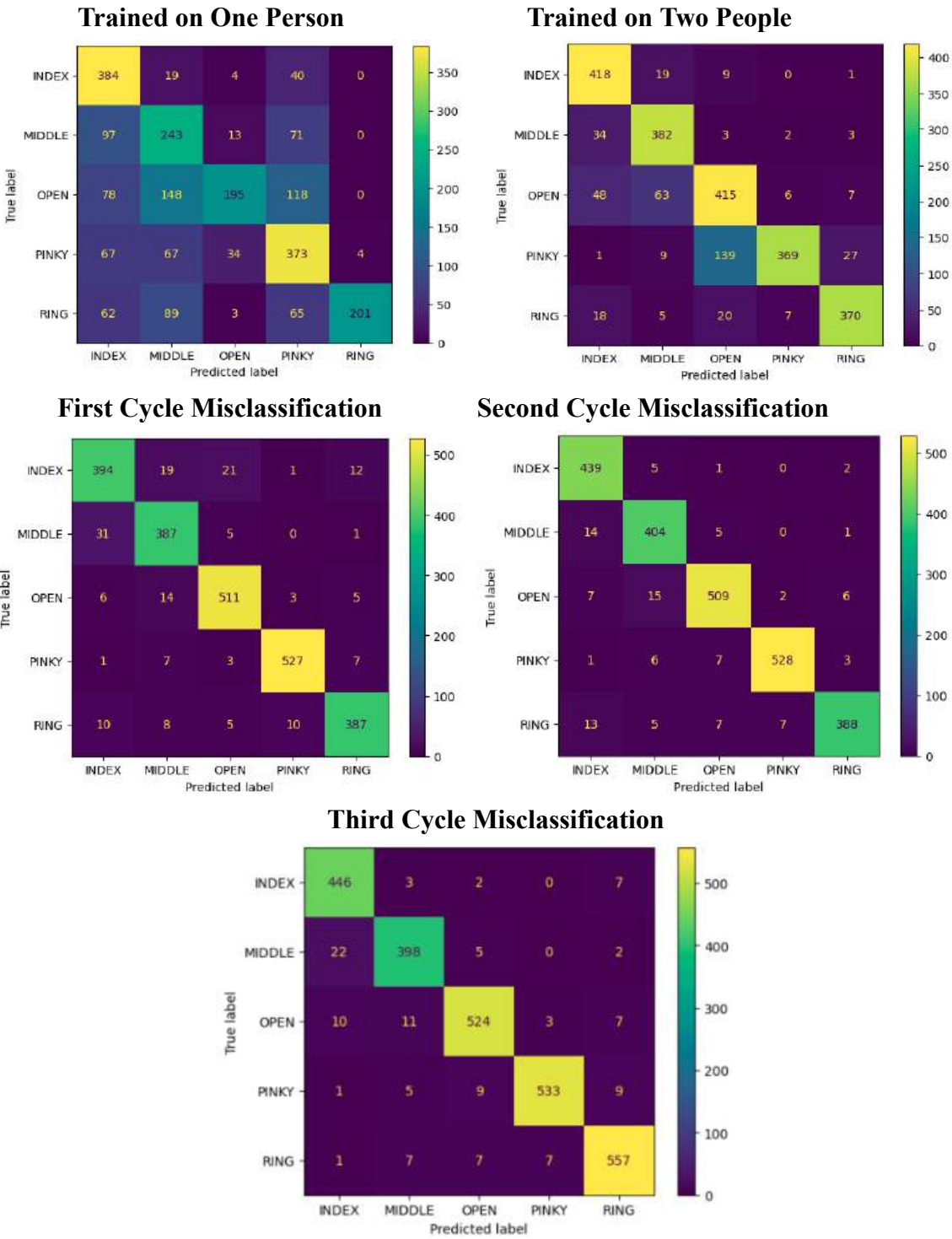
This method ensures consistency in data processing throughout both the training and live classification stages and also makes the SVM work better. The testing set was carefully split and managed (see section 2.5.2) through the `test_train_split()` feature so that no model would be trained on any of the image data in the testing set to ensure consistency throughout the analyzing phase.

### 3) Results

As shown through Table 1, every round of more data inputs significantly increased the overall F1 score as well as improving each individual class's precision and recall. Each class's individual precision and recall was retrieved from testing the classifier on the total test set. Then, the F1 score for that class is calculated through the given formula and the overall classifiers F1 score is calculated by averaging the individual class's F1 scores. Figure 5 shows the resulting confusion matrix for each of the individual classes for each classifier, showing that each class individually improves for every iteration of training.

**Table 1:** Precision, Recall, and F1 Score (P, R, and F respectively) with respect to different classifiers trained on different datasets.

Data Set → Total Images →	Data on 1st Person (200 Per Class)		Data on 1st and 2nd person (400 Per Class)	Add in first cycle of misclassification	Add in second cycle of misclassification	Add in third cycle of misclassification
Class Open	P	0.55	0.80	0.88	0.92	0.93
	R	0.85	0.93	0.86	0.94	0.97
	F	0.67	0.84	0.87	0.93	0.95
Class Index	P	0.42	0.79	0.87	0.94	0.94
	R	0.56	0.89	0.93	0.94	0.93
	F	0.48	0.84	0.90	0.94	0.94
Class Middle	P	0.77	0.70	0.93	0.95	0.96
	R	0.35	0.75	0.94	0.92	0.94
	F	0.48	0.73	0.94	0.93	0.95
Class Ring	P	0.55	0.94	0.96	0.97	0.98
	R	0.67	0.68	0.96	0.96	0.96
	F	0.60	0.79	0.96	0.97	0.97
Class Pinky	P	0.97	0.90	0.93	0.97	0.96
	R	0.47	0.86	0.92	0.93	0.96
	F	0.63	0.88	0.92	0.95	0.96
Macro-Averaged F1 Score:	<b>0.57</b>		<b>0.82</b>	<b>0.92</b>	<b>0.95</b>	<b>0.95</b>



**Figure 5:** Confusion matrices that show the improvement of each specific class as iterations of training go on, the main diagonal clearly increases in accuracy while the remaining components of the matrix slowly decrease.

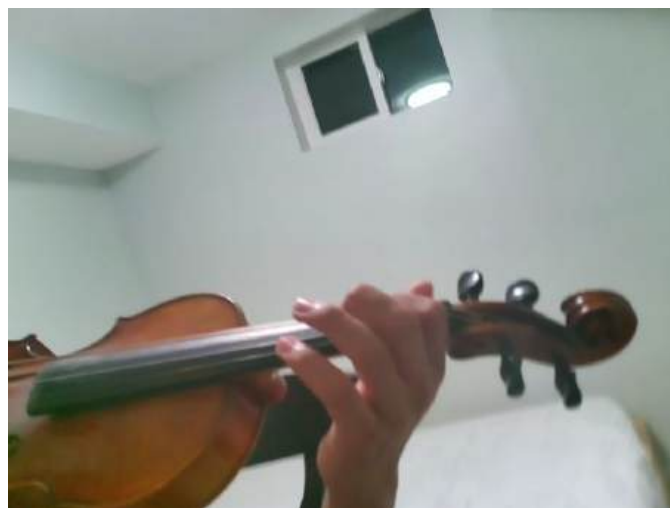
Furthermore, to ensure consistency, each model was never trained on any of the image data in the total testing set so that all of the image data in the total testing set is new to each model. The first model that was only trained on one person's data achieved a F1 score of 0.57 and performed poorly during the live testing. Through the iterations of testing by getting an additional person for more data and then running the misclassification backtracking cycle, the final model achieved a F1 score of 0.95 and significantly improved during live testing.

Therefore, through these results, it is clear that the more data that the model is trained on, the more accurate the overall model will be on a more generalized set of inputs. Although these results were only gathered from two people, this shows that this method could be generalized to a wider spread of individuals given the time and resources needed, and acts as a proof of concept for this method of classifying violin finger positions through webcam based images.

#### 4) Discussion

Perspective and camera angle are crucial in training and classifying images for hand position recognition because they directly affect the visibility and clarity of key features, such as finger placement and hand orientation. Variations in camera angle can lead to inconsistencies in the data, making it challenging for the model to accurately interpret hand positions across different perspectives. This variability can be a significant limitation, as ensuring the model performs well with every possible camera angle requires extensive data collection and training, which could be time-consuming and difficult.

Figure 6 is an example of an image being misclassified, you can see that the middle and index fingers are not fully visible to the camera. The index finger is almost completely blocked by the middle finger and the end of the middle finger is blocked by the ring finger. Because of this, it is reasonable that because mediapipe has to "guess" where those remaining keypoints are, the classifier could easily misclassify from those guesses.



**Figure 6:** Image that was misclassified - classified as a middle, should have been index

Classifying musical expressions such as vibrato, pizzicato for the right hand, and other nuanced gestures like glissando presents unique challenges due to their subtle and dynamic nature. These expressions involve intricate movements that can vary greatly in speed, intensity, and style, making them more complex to detect and classify accurately using image recognition techniques. While these elements are crucial for capturing the full range of musical expression, they were not considered as variables for this paper. Focusing on the fundamental task of fingering classification first lays the groundwork for future studies to explore these more advanced aspects of musical performance.

Future research could explore the scalability of this approach by incorporating a more diverse dataset, including different lighting conditions, backgrounds, and hand sizes. While this study focused primarily on static fingering classification, extending the model to support live playing would be a significant advancement. However, more work is needed to handle the complexities of real-time performance, such as quickly adapting to dynamic changes and maintaining high accuracy. Additionally, there is a wide range of exploratory work that could be pursued, such as determining which string the fingers are on and analyzing the right-hand movements to identify note values and rhythms.

Furthermore, integrating hand position detection could enhance the system's ability to classify different playing positions on the violin. By examining the wrist's location relative to the tuning pegs and calculating distances, the model could potentially identify the position on the fingerboard. Another exciting area for future research involves using right-hand motion to estimate volume based on the speed of bowing, which could add an expressive dynamic element to the virtual instrument.

Although this project did not achieve a complete virtual instrument capable of replicating every aspect of violin playing, it lays the foundation for ongoing development in this area. There remains a clear path forward to refine and expand the system, bringing us closer to a comprehensive virtual instrument that can fully emulate the nuances of live performance.

## **5) Conclusion**

In conclusion, this research demonstrates the potential and limitations of using Mediapipe and support vector classifiers for detecting violin fingering positions to play a virtual instrument. The study initially achieved a promising accuracy of approximately 90% under controlled conditions. By expanding the dataset and iterating through cycles of misclassification backtracking, the model's performance significantly improved, achieving a final F1 score of 0.95. This research contributes to the field of AI-driven music technology by offering insights into practical challenges and opportunities associated with image recognition in musical applications. While the study focused on a specific application of detecting violin fingering, the methodology and findings can be generalized to other types of hand movements and virtual instrument interfaces.

### Work Cited

- [1] Sturm, Bob L. “A survey of evaluation in music genre recognition.” Lecture Notes in Computer Science, 29 Oct. 2014, pp. 29–66, [https://doi.org/10.1007/978-3-319-12093-5\\_2](https://doi.org/10.1007/978-3-319-12093-5_2).
- [2] Labrou, Katerina, et al. “Following the Master’s Hands: Capturing Piano Performances for Mixed Reality Piano Learning Applications.” Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems, 19 Apr. 2023, pp. 1–8, <https://doi.org/10.1145/3544549.3585838>.
- [3] Oudah, Munir, et al. “Hand gesture recognition based on Computer Vision: A review of techniques.” Journal of Imaging, vol. 6, no. 8, 23 July 2020, p. 73, <https://doi.org/10.3390/jimaging6080073>.
- [4] Goodfellow, Ian, et al. “Deep Learning.” Deep Learning, 2016, [www.deeplearningbook.org/](http://www.deeplearningbook.org/).
- [5] Kapur, Ajay, et al. “The electronic tabla controller.” Journal of New Music Research, vol. 32, no. 4, 1 Dec. 2003, pp. 351–359, <https://doi.org/10.1076/jnmr.32.4.351.18854>.
- [6] Lugaresi, Camillo, et al. “MediaPipe: A Framework for Building Perception Pipelines.” arXiv.Org, 14 June 2019, [arxiv.org/abs/1906.08172](https://arxiv.org/abs/1906.08172).
- [7] Cortes, Corinna, and Vladimir Vapnik. “Support-Vector Networks.” Machine Learning, vol. 20, no. 3, Sept. 1995, pp. 273–297, <https://doi.org/10.1007/bf00994018>.
- [8] theAIGuysCode. “Colab-Webcam/Colab\_webcam.Ipynb at Main · Theaiguyscode/Colab-Webcam.” GitHub, 2020, [github.com/theAIGuysCode/colab-webcam/blob/main/colab\\_webcam.ipynb](https://github.com/theAIGuysCode/colab-webcam/blob/main/colab_webcam.ipynb).
- [9] Evidently AI Team. “How to Interpret a Confusion Matrix for a Machine Learning Model.” Evidently AI - Open-Source ML Monitoring and Observability, 2020, [www.evidentlyai.com/classification-metrics/confusion-matrix](http://www.evidentlyai.com/classification-metrics/confusion-matrix).

# **Body Mass Index Mediates the Association Between Physical Activities and Mental Health**

## **By Amay Munjal**

### **Abstract**

Previous research has established the connection between physical activities and positive mental health, with studies showing that individuals who engage in intense and regular physical activities have better mental health outcomes. However, to the best of my knowledge, there is no study on the role of Body Mass Index in the relationship between physical activity and mental health. In this study, the hypothesis tested was that physically active individuals achieved good mental health outcomes primarily because they are in better physical condition (healthy BMI). A sample of 7,932 participants (Male = 3524, Female = 4408; Mean Age = 40.387; SD = 13.299) was analyzed from the Scottish Health Survey. These participants provided information on the frequency of their physical activities and BMI. They also completed a 12-item General Health Questionnaire designed to measure mental health. Results of simple regression analysis showed that the frequency of physical activities is negatively associated with mental health score ( $\beta = -0.172$ ,  $p < 0.001$ ), with a higher frequency of physical activities associated with reduced mental health problems. Similarly, results of simple mediation analysis showed that a higher frequency of physical activity was associated with a better and healthy Body Mass Index which, in turn, predicted a better mental health outcome. Taken together, this study supports previous studies that showed that engaging in regular physical activity has a positive impact on mental health. This study also showed that if the positive impact of physical activity on mental health can be explained by the Body Mass Index, then achieving physical fitness through physical activities not only improves our mental well-being but also aids our physical health.

### **Introduction**

Mental health problems have continued to plague modern society, with huge funding from global organizations and national governments devoted to tackling this condition. As is well-known in society, mental health problems are taken seriously and are an ongoing topic that continues to draw attention due to the problems and costs that it presents to society if not properly addressed. Recent studies have shown that physical activities, such as exercises or sporting activities can improve mood and alleviate symptoms of mental health problems. The purpose of the current study is to examine how the frequency and intensity of physical activity can help improve individuals' mental health.

Several studies have examined the association between physical activities and mental health (e.g., Reardon et al., 2019; Kuettel and Larsen, 2020). For example, one recent study showed that engaging in physical activity or sporting events can boost an individual's mental health, and lower the risk of mental health problems (Trajković et al., 2023). Physical activity is proven to help improve depression and anxiety conditions (Warburton et al., 2006). The benefits to exercise on mental health vary from improved sleep to weight reduction. It can also help improve cognitive function. (Sharma et al., 2006).



Frequent or intense physical exercise is assumed to be particularly effective in reducing mental health problems. Studies have shown that a 30-minute walk can improve mental health more than just a 10-minute walk (Sharma et al., 2006). Consistency of a workout is important because it makes a healthy habit and schedule of working out which is good for mental health. The intensity of a workout can affect just how much work you are actually putting into a workout. These variations of exercise programs can affect mental health. A workout done every day is significantly more effective for one's mental health than a workout done once a week in the way it reduces depression and anxiety (Nie et al., 2021). Studies have shown that the intensity of a workout can improve how a person feels about themselves afterward (Smith and Merwin, 2020).

A person's weight can greatly affect how they feel about themselves. Obesity is recognized to be linked with a higher occurrence of common mental disorders such as depression and anxiety (McCrea et al., 2012). If someone is overweight, they feel less confident in themselves which can lead to depression and low self-esteem (McCrea et al., 2012). Studies have shown that individuals with higher BMI scores reported greater symptoms of common mental health disorders, and this link between BMI and mental health also differed by gender and age. Among young women, the likelihood of experiencing a mental disorder grew as BMI increased, whereas among young men, the relationship followed a U-shaped pattern, with higher probabilities observed for both underweight and obese individuals. These associations became less prominent in older age groups, particularly when considering potential confounding factors like physical health (McCrea et al., 2012).

Although previous studies have established the relationship between physical activities and body mass index or mental health, very few studies, if any, have looked at the interrelations between these variables. That is, how the effect of physical activities on body mass index would likely alter the mental health status of individuals. One likely hypothesis is that engaging in physical activities frequently and intensively would lead to healthy BMI which would, in turn, lead to better mental health of such individuals. On the other hand, individuals who rarely engage in physical activities would likely have an unhealthy BMI which may lead to poorer mental health. In this study, a simple mediation model was used to test the hypothesis that BMI has a mediating effect on the relation between physical activities and mental health. This study hypothesizes that individuals who engage in frequent and intensive physical activities will have a healthy BMI which will, in turn, lead to better mental health. The findings from this study will go along in providing insight into the interrelations between physical activities, BMI, and mental health, and hope that findings from this study will influence sports policies in schools and in larger societies.

## **Methods**

### **Participants**

The data from the Scottish Health Survey was analyzed. In 1995, the Scottish Health Survey ran a study to find data about the health of the working-age population (ages 16-64) in Scotland who live in private households (Hamer et al., 2008). Interviews with participants started in 1995 and were held over 12 months to allow for possible seasonal differences in health and behavior. To choose participants, in stage 1, they used 312 postcode sectors, and each month interviews occurred in 26 sectors. Households within the sectors were chosen at random. After the household was selected, the interviewer would choose a participant from varying ages of 16-64 years old.

In the current study, A sample of 7,932 participants (Male = 3524, Female = 4408; Mean Age = 40.387; SD = 13.299) was analyzed. These participants provided data on physical activities, mental health, and Body Mass Index, during the Scottish Health Survey of 1995.

### **Measures**

Three measures derived from the Scottish Health Survey were analyzed. This includes physical activity, mental health, and Body Mass Index. Participants were asked about their level of physical activities which ranged from engaging in different sporting activities such as basketball, soccer, cycling, etc. The survey asked participants about the intensity of their physical activities measured on a scale of 1-5. The participants were also asked how frequently they engage in physical activities. If participants did not engage in physical activities, responses were recorded 1; responses were coded 2 if participants engaged in physical activities once a week; participants' responses were coded 3 if they engaged in physical activities 2-3 times a week; participants' responses were coded 4 if they engaged in physical activities 4-5 times a week; and finally, if participants engaged in physical activities more than 5 times a week, responses were coded 5. To understand how a participant was feeling and how their mental health was affecting them, participants' data on a 12-item General Health Questionnaire (GHQ-12) was obtained. A GHQ-12 is a widely used survey across studies to determine a person's depression and anxiety levels. A score of over 4 likely indicates ill mental health. It asks participants questions such as, "Have you recently been able to concentrate on what you are doing?" or "Have you recently been thinking about yourself as a worthless person?" These questions combined with the participants' answers give a clear-cut image on how the participants' mental health affects them. Each item in the GHQ-12 was coded 1 if participants reported symptoms or answered "yes" to each of the 12 questions or 0 if no symptom/ "no" was provided for each of the 12 questions. GHQ-12 had a maximum score of 12 indicating the poorest level of mental health. Finally, data on the participants' BMI was also obtained. Body Mass Index is calculated by taking a participant's weight (kg) and dividing it by their height squared (m). A very high BMI indicates that a person is likely obese in an unhealthy way.

### **Statistical Analysis**

All data analysis for this study was performed in R and RStudio. Firstly, A simple regression analysis was performed where the frequency of physical activities performed by the

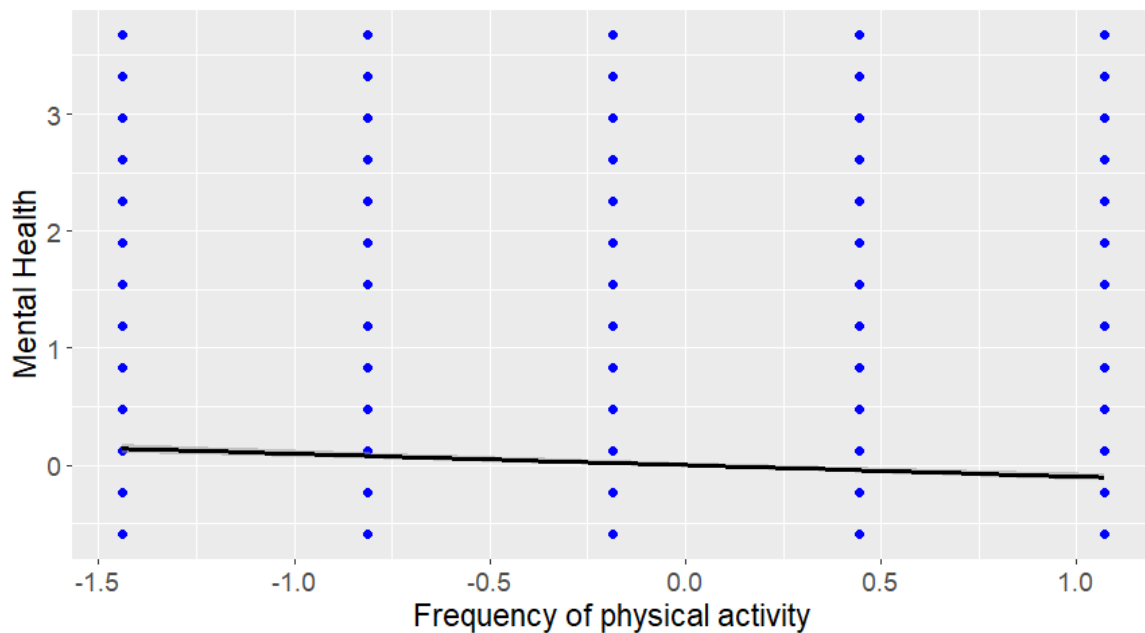
participants was entered as a predictor and mental health score as the outcome. This analysis made it possible to test the hypothesis that engaging in frequent physical activities will lead to an improved or better mental health. Secondly, a simple mediation analysis was also conducted where it tested the mediating effect of Body Mass Index on the association between frequency of physical activities and mental health. To do this, the following aspects were used: frequency of physical activities, Body Mass Index, and mental health score as the predictor, mediator, and outcome variables respectively.

In all analyses, the significance threshold was set at  $\alpha < 0.05$ .

## Results

### Association between physical activity and mental health

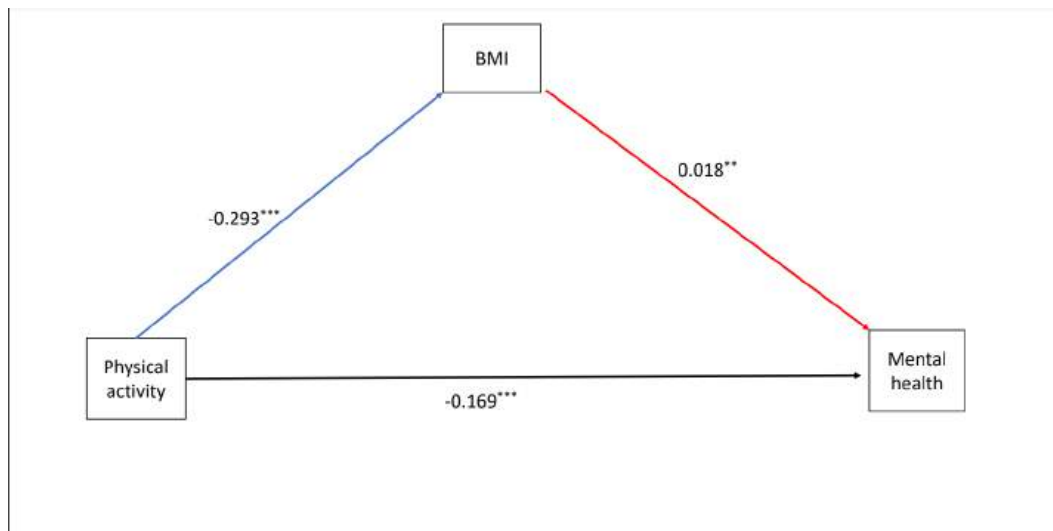
To examine the association between physical and mental health, a simple regression analysis was performed where the frequency of physical activity was entered as the predictor and total mental health score as the outcome in the model. Results showed that sports frequency was negatively associated with mental health problems ( $\beta = -0.172, p < 0.001$ ). This implies that as individuals' frequency of physical activities increases, their mental health problems decrease. This result is consistent with previous studies that have shown that engaging in regular physical activities can lead to better mental health outcomes.



**Figure 1:** Figure 1 shows the association between the frequency of physical activity and mental health problems. The displayed figure showed that individuals with greater frequency of physical activity had lower mental health problems.

## The mediating effect of Body Mass Index

An additional analysis was conducted that tested the hypothesis that Body Mass Index mediates the association between physical activity and mental health using a simple mediation model. In this model, frequency of physical activity, BMI, and mental health score were entered as predictor, mediator, and outcome variables respectively. Results of simple mediation analysis showed that sports frequency was negatively associated with BMI (beta = -0.293,  $p < 0.001$ ). This implies that engaging in regular and frequent physical activity predicted a healthy BMI. In addition, BMI was also positively associated with total mental health score (beta = 0.018,  $p = 0.005$ ). This implies that mental health outcomes become poorer as the BMI score increases to an unhealthy cut-off. A simple test of the indirect effect of BMI showed that BMI significantly mediated the association between frequency of physical activity and mental health score (beta = -0.005,  $p = 0.007$ ). In other words, engaging in regular physical activities predicted a better and healthy BMI which, in turn, predicted a better mental health outcome. These findings support the hypothesis that regular physical activities predict a better mental health outcome through its positive impact on BMI.



**Figure 2:** Figure 2 shows a simple test of mediation analysis with results showing that higher frequency of physical activity was associated with healthy BMI which, in turn, predicted better mental health.

## Discussion

This study used the Scottish Health Survey to examine the association between physical activity and mental health as well as the mediating role of Body Mass Index in the association between frequency of physical activity and mental health. The hypothesis that regular physical activities will predict a healthy Body Mass Index which will, in turn, be associated with better mental health was tested. Results showed that greater regular physical activities predicted a better mental health outcome. Similarly, results of mediation analysis showed that a higher frequency of physical activities was associated with a healthier Body Mass Index which in turn

predicted better mental health. These findings support the hypothesis that physical activity predicts better mental health outcomes through its impact on Body Mass Index.

The findings which showed that physical activity was associated with better mental health are consistent with previous studies (Mahindru et al., 2023). Physical activities several benefits to mental health, such as creating habits in an individual and the body lets out hormones such as endorphins and serotonin which improve mood. (Vafiadis, 2021). Along with improving fitness, it also distracts an individual from ruminative negative thoughts (Vafiadis, 2021). Physical activity helps an individual with body image and feeling better about themselves. It also helps with decreasing stress tendencies (Mahindru, 2023). Additionally, physical activities help improve functions of the hypothalamus-pituitary-adrenal axis and lower cortisol levels (Mahindru, 2023). This may substantially explain why frequent physical exercises improves well being and reduces mental health problems.

The findings also showed that Body Mass Index was associated with mental health outcome, with healthy Body Mass Index predicting better mental health. This finding is consistent with previous studies which showed that an individual's Body Mass Index can affect their emotional well being (Vafiados, 2021). Due to common stereotypes about individuals with excess body weight, largely seen as unattractive, which may cause such individuals to feel bad about themselves. Weight discrimination is frequently associated with situations (e.g., workplace), where, when selecting between two individuals competing for the same job, the decision-maker may opt against candidates simply because of their weight (Vafiadis, 2021). This could potentially lead to lower self-esteem in the affected individuals. Along with this, those who have excess weight may likely have health problems such as diabetes and cardiovascular issues. These health-related issues could also be a potential predictor of mental health problems such as depression and anxiety disorders as well.

More importantly, in a test of our key research hypothesis, the findings in this study showed that regular physical activities predicted better mental health through its impact on Body Mass Index. In other words, engaging in regular physical activities predicted a healthy Body Mass Index which subsequently predicted better mental health. Studies have shown that individuals with excess weight are 55% more likely to have depression than those who are not (Vafiadis, 2021). Along with that, a person's quality of life decreases with the accumulation of excess weight due to the fact their weight limits them from doing things that people with a healthy Body Mass Index could do (Vafiadis, 2021). These findings, therefore, support the hypothesis. Previous studies have shown that engaging in regular and intense physical activities was associated with low body fats and general body fitness (Bradbury et al, 2017). In addition, previous studies also show that individuals who have a high or low Body Mass Index are more vulnerable to different psychiatric disorders including mood and anxiety disorders (Kesavachandran, 2009). The results of the current study support the evidence from these past studies which showed that participants who regularly exercised might have, over time, achieved general good body fitness which subsequently reduced the risk of psychiatric disorders (e.g., mood and anxiety disorders).

## **Conclusion**

The interrelations between physical activities, Body Mass Index, and mental health outcomes using participants who took part in the Scottish Health Survey were examined. After testing the hypothesis that physical activities will affect mental health through its impacts on Body Mass Index. The findings appear to support this hypothesis and provide additional evidence about the importance of regular physical exercise or activity. Engaging in frequent physical activities not only boosts our physical health by maintaining adequate and healthy Body Mass Index but also improves our mental health.

## Works Cited

- Bradbury, K. E., Guo, W., Cairns, B. J., Armstrong, M. E., & Key, T. J. (2017). Association between physical activity and body fat percentage, with adjustment for BMI: A large cross-sectional analysis of UK Biobank. *BMJ Open*, 7(3), e011843. <https://doi.org/10.1136/bmjopen-2016-011843>
- Dorothea Vafiadis. (2023, April 12). The national council on aging. The National Council on Aging. <https://www.ncoa.org/article/excess-weight-food-security-and-mental-health-is-there-a-connection>
- Hamer, M., Stamatakis, E., & Steptoe, A. (2008). Dose-response relationship between physical activity and mental health: The Scottish health survey. *British Journal of Sports Medicine*, 43(14), 1111-1114. <https://doi.org/10.1136/bjism.2008.046243>
- Mahindru, A., Patil, P., & Agrawal, V. (2023). Role of physical activity on mental health and well-being: A review. *Cureus*. <https://doi.org/10.7759/cureus.33475>
- Mathur, N., Kesavachandran, C., & Bihari, V. (2009). Can physical activity maintain normal grades of body mass index and body fat percentage? *International Journal of Yoga*, 2(1), 26. <https://doi.org/10.4103/0973-6131.53839>
- McCrea, R. L., Berger, Y. G., & King, M. B. (2011). Body mass index and common mental disorders: Exploring the shape of the association and its moderation by age, gender and education. *International Journal of Obesity*, 36(3), 414-421. <https://doi.org/10.1038/ijo.2011.65>
- Sharma, A. (2006). Exercise for mental health. *The Primary Care Companion For CNS Disorders*, 8(2). <https://doi.org/10.4088/pcc.v08n0208a>
- Smith, P. J., & Merwin, R. M. (2021). The role of exercise in management of mental health disorders: An integrative review. *Annual Review of Medicine*, 72(1), 45-62. <https://doi.org/10.1146/annurev-med-060619-022943>
- Trajković, N., Mitić, P. M., Barić, R., & Bogataj, Š. (2023). Editorial: Effects of physical activity on psychological well-being. *Frontiers in Psychology*, 14. <https://doi.org/10.3389/fpsyg.2023.1121976>
- Wang, Y., Nie, J., Ferrari, G., Rey-Lopez, J. P., & Rezende, L. F. (2021). Association of physical activity intensity with mortality. *JAMA Internal Medicine*, 181(2), 203. <https://doi.org/10.1001/jamainternmed.2020.6331>
- Warburton, D. E. (2006). Health benefits of physical activity: The evidence. *Canadian Medical Association Journal*, 174(6), 801-809. <https://doi.org/10.1503/cmaj.051351>

## **What Practices Optimize Educational Experiences for Youth in the Juvenile Justice System? By Eliana White**

### **Abstract**

Meeting the educational needs of justice-engaged youth in the United States continues to be a challenge for juvenile justice and educational systems. Half of the young people between the ages of 10-17 involved with the justice system have been involved with the child welfare system, while others have faced homelessness and have challenges with social-emotional skills (Nash and Bilchik, 2009). Estimates from the National Council on Disability suggest that up to 85% of incarcerated youth have learning or developmental disabilities. Since juvenile corrections facilities typically offer schooling in relatively small class sizes of 7-15 students, this means that roughly 5-12 students with disabilities, and a range of developmental and unique learning needs, may be in a single classroom together.

This research explores the most effective strategies to optimize educational experiences for these youth. The focus will be on several key areas: the importance of early screening and intervention for developmental disabilities, the benefits of fostering school bonding for positive outcomes during incarceration and post-release, and the role of the Universal Design for Learning in supporting diverse learning needs. This paper finds that, while all three initiatives are advantageous, early screening for developmental disabilities is the most urgent intervention to implement since it has immediate implications for how young people take in, process, and communicate information.

### **Introduction**

The United States continues to lead the world in youth confinement rates. In 2020, the number of youth, children under the age of 18 years, detained in a juvenile facility on a typical day was 27,600, down 75% from 108,800 in the year 2000 (Rovner, 2024). Still, the U.S. continues to confine youth at rates 11 times higher than countries in Western Europe and Asia, 10 times higher than Eastern Europe, and four times higher than Mexico and Canada (U.N. Global Study, 2019). The juvenile courts processed approximately 437,300 cases in the year 2021 (Hockenberry et al., 2024). In 2018, the average length of stay in a juvenile detention center was 27 days (Annie E. Casey Foundation, 2020).

The negative effects of detention on young people is a widely studied subject. Even short stays in detention can yield poor outcomes with lifelong impact (Annie E. Casey Foundation, 2018). Unintended consequences of detention include increased likelihood of penetrating further into the justice system, exacerbating mental health issues, disrupting academic life, and increasing the likelihood of dropping out of high school (Annie E. Casey Foundation, 2018). Academic experience is pivotal in building the skills youth need to contribute and engage in their community and, for a justice-engaged youth, it plays the critical role in reducing the likelihood of recidivism.



This research focuses on the educational experience of youth in the juvenile justice system.<sup>2</sup> A staggering 85% of incarcerated youth have learning, emotional, and developmental disabilities and only 37% receive special education services(National Council on Disability). Policymakers, legislators, community stakeholders, and educators alike have long struggled with how to best address the educational needs of justice-engaged youth. The needs of this population are complex, but the obligation to ensure that these young people are entitled to a free and appropriate public education, like all other public school students, is straight-forward. The rights of these students are protected by the Individuals with Disabilities Education Act (IDEA) ("Individuals with Disabilities Education Act"), No Child Left Behind Act ("No Child Left Behind Act"), and the Prevention and Intervention Programs for Children and Youth Who Are Neglected, Delinquent or At Risk of the Elementary and Secondary Education Act (ESEA) (U.S. Department of Education "Every Student Succeeds Act (Essa)").

Decades of research has confirmed that educating youth in the juvenile justice system is of great importance . The lack of education among this group increases the likelihood of future delinquency and engagement in negative outcomes, and incarcerating young people is expensive. For instance, the Office of Juvenile Justice and Delinquency Prevention data indicates that just one youth dropping out of school to engage in a life of crime and of drug abuse poses a roughly estimated \$1.7 - 2.3 million annual cost to society (Office of Juvenile Justice and Delinquency Prevention, 2018). The cost of high school dropouts is profound. Compared to a high school graduate, the earning potential of a high school dropout is considerable. Two-thirds of students lacking reading proficiency by the end of fourth grade will be incarcerated or on public assistance (BeginToRead, 2013).

Youth in the justice system generally encounter challenges with education. The young people caught up in the juvenile justice system are often on the fringes of a typical educational experience. Justice-engaged youth generally demonstrate poor educational outcomes, like low grades and standardized test scores, school drop-out, and truancy (Keith et al., 2002). Learning how to read, in particular, is a major problem for juvenile offenders (Vacca, 2008). In fact, 85% of all juveniles who come into contact with the juvenile court system do not have the reading and writing skills to manage daily living and employment functions (BeginToRead, 2013). Young people involved in the justice system also demonstrate high levels of undiagnosed and unmet special education needs, school discipline referrals, and low social emotional competencies (Cottle et al., 2001; Cuellar et al., 2015; Moffitt et al., 1981; Moretti, 2014; Pettit et al., 2004; Wang et al., 2005). Other risk factors include negative attitudes toward school, low school attachment and commitment to school, frequent school transitions, and low academic aspirations (Office of Juvenile Justice and Delinquency Prevention, 2019).

Even though the challenges of educating this complex population are extensive, the upside is enormous. Justice-involved youth who have higher levels of educational achievement while incarcerated are more likely to return to school after release and complete their high school

---

<sup>2</sup> For the purpose of this research, youth in the juvenile justice system are referred to as justice-engaged youth and justice-involved youth.

education or attain a GED. Additionally, the recidivism rate, that is, the re-arrest rate, was less for those youth who re-enrolled in school and attended regularly. At the 12-month interval, youth with higher school attendance had a 26.4% lower probability of re-arrested and 15.3% less likely at 24 months (Blomberg et al., 2011).

Youth who were re-arrested were arrested for notably less serious offenses compared to their counterparts who did not attend school or attended less regularly, providing another reason why understanding and meeting the educational needs of this population is important (Blomberg et al., 2011). One of the goals of the Juvenile Justice and Delinquency Prevention Act (reauthorized 2018) is to prevent juvenile involvement in delinquent behavior and promote public safety, and these outcomes further those goals (Office of Juvenile Justice and Delinquency Prevention, 2018").

## **Background**

### *Education While in Custody Matters*

While researchers have connected poor educational outcomes to higher rates of re-offending, they have also identified that youth engagement with their educational life is associated with positive outcomes for juveniles returning to their communities (Jäggi et al., 2020). During a young person's incarceration, their educational experiences can serve as a critical opportunity to repair the young person's relationship to school and learning; their experiences are associated with a decrease in delinquency after they are released and serve as a predictor for their level of school engagement after they return to school (Jäggi et al., 2020). Returning to school is important in and of itself, but even more so to the extent that it serves as an indicator of community reintegration. Youth who are incarcerated for longer periods of time are also less likely to return to school. An incarcerated youth's attachment to the facility school and their academic achievement while in school are connected to the likelihood that they will return to school and reintegrate to their community. Finally, youth who attended school more regularly following release committed less serious offenses, if re-arrested, when compared to those who did not attend school or attended less regularly. This finding suggests that youth who attended school more regularly were developing prosocial behaviors and attachment (Blomberg et al., 2011).

### *Incarcerating Youth is Expensive*

When the Justice Policy Institute first conducted an analysis of the expense of secure child incarceration in 2014, the results showed that the annual cost per youth in 33 states and the District of Columbia exceeded \$100,000. Even though there has been a decrease in juvenile arrests and incarceration rates over the past five years, 40 states and Washington, D.C. report spending at least \$100,000 annually on every detained child in 2020; some jurisdictions even report spending more than \$500,000 annually per youth. The average state expense for a young person placed in secure custody is currently \$588 per day, or \$214,620 annually, which is a 44%

increase from 2014. These costs over a six-year period show how juvenile incarceration is having an increasingly negative economic impact (Justice Policy Institute, TBD).

### *The Cost of High School Dropouts is Profound*

As a crucial protective factor against delinquency and engagement in the juvenile justice system, education and school attendance are integral for youth to meet developmental milestones. In addition, they may have long-term advantages for future work life and discourage criminal behavior. The cost of dropping out of high school is profound. The economy is hurt by \$272,000 of fewer tax revenue collected, increased use of Medicaid, Medicaid, and welfare, and higher rates of crime for each high school dropout over their lifetime (Levin and Belfield 2007) (National Center for Education Statistics, 2024). For all 23 million high school dropouts in the United States between the ages of 18-67, the annual losses in federal and state income taxes surpass \$50 billion. Health-related losses amount to \$58 billion for high school dropouts in one year, and notably, high school dropouts have a life expectancy that is 9.2 years shorter than high school graduates. Even just a one percent increase in the high school completion rate for men aged 20-60 years would reduce annual crime-related expenses for the U.S. by up to \$1.4 billion. Certain crimes would be reduced, such as murder and assault rates by almost 30%, motor vehicle theft rates by 20%, arson rates by 13%, and burglary and larceny rates by about 6% -- just by increasing the average years of schooling for dropouts by one year (Rouse, 2017). A mere one percent increase in the percentage of men aged 20 to 60 who complete high school would save \$1.4 billion annually in costs associated with crime (Bloomberg et al., 2011).

### *Challenging Educational Journey and Current State of Schooling Behind Bars*

Poor academic performance, truancy, school dropout, undiagnosed and unmet special education needs, school discipline referrals, and literacy difficulties are common characteristics of the educational path for young people involved in the justice system (Cottle et al., 2001; Cuellar et al., 2015; Moffitt et al., 1981; Moretti, 2014; Pettit et al., 2004; Wang et al., 2005). Negative attitudes about education, low school attachment and commitment, frequent school changes, low academic expectations, suspensions and expulsions are additional risk factors (Office of Juvenile Justice and Delinquency Prevention, 2019). Young people involved in the delinquency system often have strained relationships with teachers and school, characterized by distrust, bitterness, and hopelessness (Finn, 1988). These young people also have a poor understanding of themselves as learners and a bleak outlook on their future educational life (Korman, 2019). Healthy youth development is often characterized by the growth of youth developmental assets, like time management skills, conflict resolution skills, resiliency, connection to caring and supportive adults, communication skills -- often practiced and matured in the school setting (Catalano, 1996). Justice-engaged students have not had the benefit of developing these protective factors inherent in a robust learning life and connection to an educational community (Finn, 1988).

A review of the current state of education in juvenile correctional facilities indicates that they are not on par with public schools. For instance, only eight states offer academic and vocational programs for youth in custody in the same manner as regular public schools (The Council of State Governments Justice Center, 2015). The education program in juvenile detention was found to have inadequate special education services and curriculum, uncertified teachers, and poor access to higher science and math learning materials. Most justice-engaged youth arrive at juvenile corrections facilities with a disheartening educational journey and complex learning profile, and while schools have an enormous chance to positively impact their captive audience, students continue to experience poor outcomes, like not passing Algebra 1 at consistently higher rates than their peers in neighborhood schools, not having as many opportunities to access to credit recovery, and not having as many instructional hours (Korman et al., 2019).

### *Disproportionately High Numbers of Justice-Engaged Youth with Disabilities*

Given that up to 85% of youth in juvenile detention facilities have disabilities and only 37% receive the special education service to which they are entitled, community stakeholders may wonder how to close that gap (National Council on Disability, 2015). Disabling conditions include emotional disturbances, specific learning disabilities, and multiple disabilities. Researchers have examined the effect of learning disabilities, in particular, and found it to be a risk factor for many delinquency outcomes (Shandra et al., 2012). Youth with learning disabilities have four times lower college graduation rates, double school dropout rates, double school suspension rates, four times higher middle school expulsion rates, nearly three times higher high school expulsion rates, and double juvenile and adult incarceration rates compared to those without learning disabilities. The undesirable outcomes are the forerunner to arrest on and off of school grounds. Young people of color with learning disabilities are disproportionately affected by these outcomes (Mallett et al., 2023). Among youth with learning disabilities, suspensions and expulsions in middle school lead to suspensions and expulsions in high school, which lead to school dropout. Early school disciplinary experiences, therefore, can have a lingering impact on a student with learning disabilities.

One category of disabilities, called developmental disabilities, include conditions like Attention Deficit Hyperactivity Disorder (ADHD), and impacts behavior, language, learning, or physical functioning. These disorders can impact daily living for both children and adults (CDC, 2024). ADHD is the most commonly diagnosed developmental disorder. Among U.S. children between the ages of 3-17 years, approximately seven million, or 11.4%, had ever been diagnosed with ADHD, according to a 2022 survey of parents (Danielson et al., 2024). The number of U.S. children who had ever been diagnosed with ADHD grew by one million in the preceding six years leading up to the survey (Bitsko et al.). Research indicates that at least 30% of young people in detention meet the criteria for ADHD (Retz et al., 2021).

### *Educational Rights*

Just like their peers who are *not* engaged in the criminal justice system, juvenile offenders, under the federal Individuals with Disabilities Education Act (IDEA, 1990), have a right to a free and appropriate education (FAPE, 1973) (U.S. Department of Education "Individuals with Disabilities Education Act"). IDEA was first passed nearly fifty years ago, and gives states and local communities access to billions in federal funds to help them provide special education options for the 7.5 million or more students who participate in special education with a wide array of disability needs, representing 15% of all public school students (National Center for Education Statistics, 2024). The foundation of IDEA is the concept of FAPE. According to IDEA, kids with disabilities should have the same educational opportunities as their classmates without disabilities, if at all feasible. IDEA also acknowledges that it is the public's duty to pay for the costs involved in meeting the particular requirements of children with disabilities. Thus, the FAPE concept serves as the law's central idea. FAPE generally refers to a child's right to a publicly funded education that is suitable for their age and skill level.

The Elementary and Secondary Education ACT (ESEA, 2015) aims to: enhance educational opportunities so students with special education needs can reach the rigorous State academic content; provide these students with the supports necessary to effectively move from the institutional setting to more school or the workforce; and ensure that youth who are returning to community from an institution and are at-risk of dropping out, have the support system they need to continue their education.

## **Potential Solutions**

*Screen all youth entering the juvenile justice system for developmental disabilities.*

Youth in the juvenile justice system have high rates of developmental disabilities, including ADHD which has a high correlation to delinquency (National Council on Disability, 2015). Youth in the juvenile justice system would greatly benefit from a developmental disabilities screening, given the prevalence. During the four-year period from 2014-2018, approximately one in six children and teens in the U.S. age 3-17 years had a developmental delay, disorder, and disability (Cogswell et al., 2022). Young people with developmental disabilities often have difficulty taking in, processing, and communicating information, and, like others, they may make behavioral choices based on how they understand this information and related expectations (Handler et al., 2011). Educators, developmental psychologists, reading specialists, and pediatricians agree that children who exhibit signs of learning disabilities, in particular, should be referred for educational, psychological, neuropsychological, and/or medical diagnostic assessments as soon as possible. Acquiring literacy skills is one important and well-established indicator of later school success (Keith et al., 2022). Poor verbal reasoning and low vocabulary skills are strong predictors of delinquency (Loeber et al., 1983). Numerous studies have demonstrated the positive effects of early identification and intervention. Experts advise against a wait and see approach noting that it is in the child's best interest to assess and intervene early (National Joint Committee on Learning Disabilities, 2024). Early reading

intervention methods, like Response-to-Intervention (RTI), immediately place struggling students in multi-tiered intervention programs as soon as the student demonstrates learning and behavioral needs (Fuchs et al., 2003). Continued academic difficulty for students produces ongoing discomfort with the school setting and leads to the student internalizing their inadequacy as a learner (Finn et al., 1988). Early assessment and intervention of students who have entered the juvenile justice system will help individualize services and further experience unwanted outcomes.

The American Academy of Pediatrics recommends a thorough assessment in every sphere related to the suspected disability (Handler et al., 2011). These evaluations usually entail interviews with the child and family, rating scales and child medical, social, and developmental history, classroom observation, and analysis of test data ("Dyslexia"). The testing can be conducted by trained school or outside specialists. The composition of testing by a school psychologist varies according to state and school district. This process produces usable information about, including whether attentional and/or emotional issues stem from learning challenges or are contributing to them (Willcutt et al., 2000). Developmental disabilities include an array of and are present in groups across racial and ethnic lines, as well as socioeconomic groups (CDC). Developmental disabilities include ADHD, Autism Spectrum Disorder, Cerebral Palsy, and hearing loss, to name a few (Zablotsky et al., 2019). Partnering with local universities that have graduate-level clinical psychology programs may help offset the cost of a private evaluator who, generally, charges \$1 - \$5 thousand per evaluation. The cost of an evaluation also depends on the geographic location and the specialist providing the assessment (Glassman, 2024).

Individualized services have also proven important in the effective support of justice-engaged youth. An evaluation of 80 rehabilitation programs indicated that those programs with matched services to individual learning needs experienced greater efficacy than those that did not. The analysis showed that when programs aligned interventions with individual needs, recidivism was reduced by as much as 50 percent (Anderson, 2001). Program staff need to have the information necessary, like that which contains youth developmental screenings, to individualize services. In the medical model of screening, a developmental screening is essential to having a clear path forward in the intervention. Information about learning styles is especially important for this age group as program staff are working with youth to improve academic, social, emotional, and behavioral skills. Educational, mental health, and all facility staff are, together, responsible for improving skills for a high-need, complex population, so equipping these practitioners with insight into how each individual client takes in, processes, and communicates information is critical.

One of the challenges to screening all youth entering the juvenile justice system is to get parental input on parent-required measures. Parents of youth in the juvenile justice system, themselves, often have their own undiagnosed learning challenges, are transient, and have challenges related to poverty (Morelli et al., 2014). Facility and educational staff also have to puzzle through training and workflow procedures, and navigate the complexities of

communication on a multi-disciplinary team. Identifying a developmental screening “champion” at each facility helps ensure that communication and workflow are managed efficiently (Zimskind et al., 2009). As with any complex, multi-stakeholder operation, the effort will require political will, program staff buy-in, and the regular production of usable and meaningful information.

*Target facility school climate and school bonding to optimize educational experience for justice-engaged youth.*

School bonding pertains to the connectedness that students have with their school environments, and has implications for a variety of developmental and adjustment outcomes like delinquency and antisocial behavior (Maddox et al., 2003). School bonding is an important component to achieving positive developmental outcomes, higher levels of education while incarcerated, and positive outcomes in the community once released (Maddox et al., 2003). School bonding, or school attachment, deters negative outcomes. The construct reflects student-teacher relationship, as well as school involvement and school commitment (Maddox et al., 2003). School attachment is also characterized by a sense of pride in school, sense of belonging, and belief in the school’s values (Oelsner et al., 2011).

Strategies to operationalize school bonding focus on furthering academic and behavioral objectives, ensuring student and teacher feelings of physical and emotional safety, and supporting the school's emphasis on teaching and learning (Durlak, 2015). The attachment allows for students to understand and adopt prosocial norms and control over their own behavior (Catalano, 1996). A facility school climate that nurtures school bonding has implications for all juvenile offenders returning to their communities.

Analyses of strategies to improve school attachment indicate that school bonding influences youth outcomes, like academic achievement and improved behavior (Sabatine et al., 2017). School attachment is also characterized by bonds with multiple adults working in partnership to help youth better internalize what behavior is expected of them, as well as to effectively internalize prosocial norms. These are important factors to protect against delinquency (Spoth et al., 2008).

*Develop a flexible learning environment with a variety of opportunities for meaningful student engagement.*

Building on the need to understand the unique developmental needs of each student and the importance of school attachment, provide flexible learning environments with multiple and plentiful opportunities for students to engage in meaningful ways. Universal Design for Learning (UDL) is an instructional strategy that seeks to remove any needless obstacles in the way of learning and enhance the educational experience, while meeting the requirements and skills of every student. An essential component of UDL is to create a flexible learning environment where students can participate in multiple ways of learning, material is given in different formats, and students have alternatives for how they want to demonstrate what they have learned (Karger et

al., 2013). UDL was adapted from the Universal Design development in product and architectural spaces by Dr. David Rose in the 1990s (Heelan, 2015). UDL lends itself to an inclusive learning environment where students feel that their contributions are seen and respected, and in turn, value the perspectives of classmates (Cornell University, 2018).

By creating a course with a variety of opportunities for meaningful student engagement, students are able to perceive and understand information in a way that best suits their learning style, ultimately allowing students to show what they know more effectively (Karger et al., 2013); learners are afforded more opportunities to become active participants rather than passive learners. The three principles of UDL, multiple means of representation, multiple means of student action and expression, and multiple means of student engagement, can be applied to curriculum and instruction, assessment, behavior, transition planning, and teacher preparation (Karger et al., 2013). UDL practices are also consistent with the U.S. Department of Education and U.S. Department of Justice's guiding principles for providing high-quality in juvenile justice secure care settings. The first guiding principle emphasizes priority on individual needs of all youth. The principle focuses on setting students up for success by ensuring that the learning environment is supportive and engaging -- conducive to learning (U.S. Department of Education et al, 2014). UDL teaching practices help shift the focus to the individual identities of the learners and recognize the strength and relevance in their lived experiences as tools for interacting with information. In UDL, instruction is straight-forward, accommodates individual abilities, including pacing, and elevates effective communication, regardless of the student's condition (Lunasco, 2018). UDL-based strategies have helped to increase access to curriculum for diverse learners and indicate promising results in efficacy (Ok et al., 2017). In one study, researchers noted improved student storytelling ability after they read books altered in a UDL framework (Browder et al., 2008). In world history curriculum, students who used UDL-formatted podcasts as a way to gain content scored higher on evaluation measures than those who did not (Kennedy et al., 2014).

Shifting to a UDL framework for existing educational institutions may appear to necessitate a fair amount of additional work for the educators. However, once the advanced preparation is completed and the UDL methods are woven into the program, both the teaching and learning experiences are well-positioned to be more rewarding. Center for Applied Specialized Technology (CAST), a non-profit education research and development organization offers a wide array of professional learning opportunities, designed for individuals and small groups to schoolwide implementation. CAST supports educators and administrators to design, implement, and evaluate UDL strategies through coaching and consultation, as well as courses and classes. Also, Harvard Graduate School of Education (GSE) offers an online workshop called Universal Design for Learning: Explore, a course to support the design of UDL learning environments. The \$295, three-week course provides 15 hours of instruction to gain tools and approaches to incorporate UDLA practices into teaching. The course is intended for school staff, like teachers, paraprofessionals, principals, and curriculum specialists, as well as district staff, like superintendents, union leadership, and instructional coaches. When educators are ready to



move onto the application phase of their learning, Harvard's GSE offers a subsequent course for \$495, and encourages teams of four to eight from the same school to attend. The three-week online course will support teams to understand and address pain points in practice.

As with any school system that desires to practice continuous improvement, educational experiences involving UDL will require ongoing professional development. Another consideration in the juvenile justice system is security. Implementing the three core principles of UDL: provide multiple means of teaching, multiple means of engagement, and multiple means of demonstrating skills and knowledge, actually lends itself well to a secure facility because inherent in the design is flexibility. For instance, with regard to multiple means of engagement, Probation and other detention staff, in partnership with school staff, have the opportunity to provide options for self-regulation that make sense for the facility (Karger et al., 2013).

### **Recommendation and Discussion**

Youth in the juvenile justice system have often had educational journeys fraught with learning challenges, under-resourced schools, and poor academic outcomes. Justice system practitioners, education leaders, and other community stakeholders have long grappled with how best to meet the learning needs of these young people, especially since providing a free and appropriate public education is required by law.

The benefits to identifying and addressing developmental and learning disabilities are clear. Youth who have entered the juvenile justice system and have higher levels of educational achievement are anticipated to return to school after release and finish their high school education or GED at higher rates than those who do not (Jäggi, 2020); these youth are also much less likely to re-offend.

This research reviewed the literature about what practices optimize the educational experiences for justice-engaged youth. The educational needs of juvenile delinquent youth are complex. These students have generally demonstrated low grades and standardized test scores, high rates of truancy and dropping out, unmet special education needs, and low school attachment. A justice-engaged student's educational journey is often characterized by undiagnosed disabilities, lack of connection to caring and supportive adults, exclusionary disciplinary practices, and school-based arrest. Examining strategies to assess and support the educational needs of this special population needs to begin with how disabilities are identified and addressed. Compared to their peers without disabilities, disabled students have an approximately three times higher chance of being detained (Whitaker et al., 2019).

The most urgent and fundamental of three recommendations based on this literature review is that all youth entering the juvenile justice system should be screened for developmental disabilities. Based on the fact that a large percent of the target population is shown to be at risk, an intervention path is clear, and, when unaddressed, resulting outcomes are negative, youth who have entered the juvenile justice system would benefit from a developmental screening. Justice-engaged youth have high rates of developmental disabilities, including ADHD, which highly correlates to delinquency.

The American Academy of Pediatrics recommends a complete evaluation of any suspected disability. A private evaluator charges between \$1,000 - \$5,000 per evaluation, depending on the geographic location. However, considering the previously provided information regarding the societal costs associated with high school dropouts, the benefits derived from these evaluations become evident in relation to their costs.

While targeting facility school climate and school bonding may be less pressing, it is still recommended by the literature review in order to optimize the educational experience for justice-engaged youth. Given the prevalent punitive and stigmatizing perceptions surrounding the juvenile justice system, it may seem unlikely that a youth would feel attachment or pride in their facility. When in reality, this sentiment makes it all the more important that youth find connectedness with their teachers, school staff, and classmates. The burden is on the facility and its staff to counteract the prevailing narrative of punishment. There is a fundamental difference between accountability and mere punishment. While accountability is a component of the justice system, it is essential to convey to these youths that they are supported by a network committed to their personal growth and comeback. At the most basic level, this involves access to a trusted, caring adult and a sense of belonging. School bonding serves as a protective factor and deters negative outcomes. It is an important metric to predict academic achievement, higher levels of education while in custody, and desired outcomes once released into the community. A positive facility school climate can thus play a pivotal role in transforming the lives of young people, offering them a significant and enduring opportunity for positive change.

Feasibility of this solution is complex. Prioritizing school bonding and school climate in all aspects of the school operation requires political will, strategic planning, and staff leadership. To successfully implement this strategy, all facility staff must be fully committed to creating an environment that deviates from established norms and breaks existing habits.

The last recommendation to optimize the educational experience for justice-engaged youth is to develop a flexible learning environment with a variety of opportunities for meaningful student engagement. The Universal Design for Learning is an instructional strategy that recognizes diversity in every group of learners, whether challenged with learning differences or involved in the juvenile justice system (Heelan, 2015); it recognizes that there are differences in progression rate, prior subject knowledge, and interests. The strategy focuses on removing needless obstacles that get in the way of learning, while not compromising on meeting the requirements and skills of students. The UDL is particularly well-suited for a juvenile justice facility because it asks the educators and practitioners to reflect on what aspects of the classroom create or aggravate the disability.

One of the guiding principles of the juvenile justice system has traditionally been control. However, currently, the system is not producing the kinds of outcomes that society would hope to see. It has become clear that a shift is necessary. UDL offers an opportunity for the system to reevaluate this principle, and to recognize that the focus should be on rebuilding trust between youth and educational institutions, as well as trust in themselves. It's about getting at the true nature of rehabilitation. The UDL process requires students to see and value the perspectives of

their classmates, and as a result, can expect to have their contributions seen and respected as well (Cornell University, 2018). In order to ensure the safety of both staff and students, Justice system officials may want to explore what student-to-teacher ratio is optimal for maintaining safety while implementing UDL.

## **Conclusion**

This literature review is about the educational experiences of young people involved in the juvenile justice system, based on the widely acknowledged and well-established theories that academic success and school connectedness are important protective factors that promote well-being -- and that healthy cognitive and social-emotional growth is essential to youth development.

While all three practices that were discussed are recommended, screening all youth for developmental disabilities is most advised because it is the most acute and critical strategy. This urgency stems from the fact that a significant portion of these youth are at high risk for these disabilities, which is strongly correlated to delinquency. In addition, the expense of this practice is outweighed by the long-term benefits, including improved educational outcomes and reduced societal costs associated with high dropout rates and future criminal activity. Addressing developmental disabilities early on not only addresses the root causes of justice-engagement, but it provides a foundation for more effective rehabilitation, making it the most vital step in optimizing outcomes for justice-involved youth.

The recommendations are important because education while in custody is not just a customary practice -- it's a lifeline. Researchers find that, when justice-involved youth engage meaningfully with their education, they do not only avoid reverting to old behaviors, they transform the trajectory of their future (Jäggi, 2020). Evidence shows that a flexible learning environment can provide a vital chance to mend a justice-engaged student's relationship with education while they are incarcerated, allowing for personal growth and self-discovery in the process. Adolescents who demonstrated superior academic performance while incarcerated were notably more inclined to resume their education. This phenomenon does not merely reduce recidivism, but it positively impacts the young person's life, and by extension, benefits society as a whole.

## Works Cited

- "Dyslexia." JAMA, vol. 261, no. 15, 1989, pp. 2236-39, doi:10.1001/jama.1989.03420150086042.
- "Individuals with Disabilities Education Act." vol. 20 U.S.C., 2000, 1400–1482
- "No Child Left Behind Act." vol. 20 U.S.C., 2001, 6301–7941 (Supp. V 2005) and Id. § 6315(b)(2)(D).
- Anderson, David. "When Should Kids Go to Jail?" 2001.
- Annie E. Casey Foundation. "Juvenile Detention Explained." Annie E. Casey Foundation, 2020.
- . "Kids Deserve Better: Why Juvenile Detention Reform Matters." Annie E. Casey Foundation, 2018.
- BeginToRead. "Literacy Statistics." <https://www.begintoread.com/research/literacystatistics.html>.
- Bitsko, Rebecca H et al. "Mental Health Surveillance among Children—United States, 2013–2019." MMWR supplements, vol. 71, 2022, pp. 1-42.
- Blomberg, Thomas G. et al. "Incarceration, Education and Transition from Delinquency." Journal of Criminal Justice, vol. 39, no. 4, 2011, pp. 355-65, doi:<https://doi.org/10.1016/j.jcrimjus.2011.04.003>.
- Browder, Diane M. et al. "A Meta-Analysis on Teaching Mathematics to Students with Significant Cognitive Disabilities." Exceptional Children, vol. 74, no. 4, 2008, pp. 407-32, doi:10.1177/001440290807400401.
- Catalano, RF. The Social Development Model: A Theory of Antisocial Behavior. Delinquency and crime: Current theories/Cambridge University Press, 1996.
- CDC. "Developmental Disability Basics." <https://www.cdc.gov/child-development/about/developmental-disability-basics.html>.
- Cogswell, Mary E et al. "Health Needs and Use of Services among Children with Developmental Disabilities—United States, 2014–2018." MMWR. Morbidity and Mortality Weekly Report, vol. 71, no. 12, 2022, pp. 453–58.
- Cornell University. "Universal Design for Learning." <https://teaching.cornell.edu/teaching-resources/assessment-evaluation/inclusion-accessibility-accommodation/building-inclusive-5>.
- Cottle, Cindy C. et al. "The Prediction of Criminal Recidivism in Juveniles:A Meta-Analysis." Criminal Justice and Behavior, vol. 28, no. 3, 2001, pp. 367-94, doi:10.1177/0093854801028003005.
- Cuellar, Alison Evans and Sara Markowitz. "School Suspension and the School-to-Prison Pipeline." International Review of Law and Economics, vol. 43, 2015, pp. 98-106, doi:<https://doi.org/10.1016/j.irle.2015.06.001>.
- Danielson, M. L. et al. "Adhd Prevalence among U.S. Children and Adolescents in 2022: Diagnosis, Severity, Co-Occurring Disorders, and Treatment." J Clin Child Adolesc Psychol, vol. 53, no. 3, 2024, pp. 343-60, doi:10.1080/15374416.2024.2335625.
- Durlak, J. Handbook of Social and Emotional Learning: Research and Practice. The Guilford Press, 2015.

- Finn, Jeremy D. et al. "School Performance of Adolescents in Juvenile Court." *Urban Education*, vol. 23, no. 2, 1988, pp. 150-61, doi:10.1177/0042085988023002003.
- Fuchs, Douglas et al. "Responsiveness-to-Intervention: Definitions, Evidence, and Implications for the Learning Disabilities Construct." *Learning Disabilities Research & Practice*, vol. 18, no. 3, 2003, pp. 157-71, doi:https://doi.org/10.1111/1540-5826.00072.
- Glassman, Suzie. "How to Find Free Testing for Learning Disabilities." <https://www.parents.com/kids/development/learning-disabilities/cost-of-diagnosing-child-learning-difficulty/>. Accessed 2 September 2024.
- Handler, Sheryl M. et al. "Learning Disabilities, Dyslexia, and Vision." *Pediatrics*, vol. 127, no. 3, 2011, pp. e818-e56, doi:10.1542/peds.2010-3670.
- Heelan, Ann. "Universal Design for Learning (Udl): Implications for Education." *Universal Design in Education Conference*, 12-13 November 2015.
- Hockenberry, Sarah and Charles Puzanchera. "Juvenile Court Statistics 2021." *National Center for Juvenile Justice*, 2024.
- Jäggi, Lena et al. "Schooling While Incarcerated as a Turning Point for Serious Juvenile and Young Adult Offenders." *Journal of Adolescence*, vol. 78, 2020, pp. 9-23, doi:https://doi.org/10.1016/j.adolescence.2019.11.002.
- Justice Policy Institute. "Sticker Shock 2020: The Cost of Youth Incarceration." 2020.
- Karger, Joanne and Rachel Currie-Rubin. "Addressing the Educational Needs of Incarcerated Youth." *Journal of Special Education Leadership*, vol. 26, no. 2, 2013.
- Keith, Jill Marie and Audrey Davis McCray. "Juvenile Offenders with Special Needs: Critical Issues and Bleak Outcomes." *International Journal of Qualitative Studies in Education*, vol. 15, no. 6, 2002, pp. 691-710, doi:10.1080/0951839022000014385.
- Kennedy, Michael J. et al. "Using Evidence-Based Multimedia to Improve Vocabulary Performance of Adolescents with Ld:A Udl Approach." *Learning Disability Quarterly*, vol. 37, no. 2, 2014, pp. 71-86, doi:10.1177/0731948713507262.
- Korman, Hailly T.N. et al. "Patterns and Trends in Educational Opportunity for Students in Juvenile Justice Schools." *Bellweather Education Partners*, 2019.
- Loeber, R. and T. Dishion. "Early Predictors of Male Delinquency: A Review." *Psychological Bulletin*, vol. 94, no. 1, 1983, pp. 68-99, doi:10.1037/0033-2909.94.1.68.
- Lunasco, Jeremy. "The Right to Be Human: Universal Design for Learning and Literacy Sponsorship as Liberatory Pedagogy." 2018.
- Maddox, S. J. and R. J. Prinz. "School Bonding in Children and Adolescents: Conceptualization, Assessment, and Associated Variables." *Clin Child Fam Psychol Rev*, vol. 6, no. 1, 2003, pp. 31-49, doi:10.1023/a:1022214022478.
- Mallett, Christopher A. et al. "The "Learning Disabilities-to-Prison" Pipeline: Evidence from the Add Health National Longitudinal Study." *Crime & Delinquency*, vol. 69, no. 13-14, 2023, pp. 2643-77, doi:10.1177/00111287221081024.
- Moffitt, Terrie E. et al. "Socioeconomic Status, Iq, and Delinquency." *Journal of Abnormal Psychology*, vol. 90, no. 2, 1981, pp. 152-56, doi:10.1037/0021-843X.90.2.152.

- Morelli, D. L. et al. "Challenges to Implementation of Developmental Screening in Urban Primary Care: A Mixed Methods Study." *BMC Pediatr*, vol. 14, 2014, p. 16, doi:10.1186/1471-2431-14-16.
- Moretti, Enrico. "Does Education Reduce Participation in Criminal Activities." Symposium on Social Costs of Inadequate Education, Teachers' College, Columbia University, 2005.
- National Center for Education Statistics. "Students with Disabilities. Condition of Education." U.S. Department of Education, Institute of Education Sciences <https://nces.ed.gov/programs/coe/indicator/cgg>. Accessed 2 September 2024.
- . "Trends in High School Dropout and Completion Rates in the United States." <https://nces.ed.gov/programs/dropout/intro.asp#ref4>.
- National Council on Disability. "Breaking the School-to-Prison Pipeline for Students with Disabilities." National Council on Disability, 2015.
- National Joint Committee on Learning Disabilities. "Learning Disabilities and Young Children: Identification and Intervention." <http://ldonline.org/article/11511?theme=print>. Accessed 2 September 2024.
- Oelsner, J. et al. "Factors Influencing the Development of School Bonding among Middle School Students." *J Early Adolesc*, vol. 31, no. 3, 2011, pp. 463-87, doi:10.1177/0272431610366244.
- Office of Juvenile Justice and Delinquency Prevention. "Education for Youth under Formal Supervision of the Juvenile Justice System." 2019.
- . "Juvenile Justice and Delinquency Prevention Act Reauthorization 2018." Translated by U.S. Department of Justice, 2019.
- Ok, Min Wook et al. "Universal Design for Learning in Pre-K to Grade 12 Classrooms: A Systematic Review of Research." *Exceptionality*, vol. 25, no. 2, 2017, pp. 116-38, doi:10.1080/09362835.2016.1196450.
- Pettit, Becky and Bruce Western. "Mass Imprisonment and the Life Course: Race and Class Inequality in U.S. Incarceration." *American Sociological Review*, vol. 69, no. 2, 2004, pp. 151-69, doi:10.1177/000312240406900201.
- Retz, Wolfgang et al. "Attention-Deficit/Hyperactivity Disorder (Adhd), Antisociality and Delinquent Behavior over the Lifespan." *Neuroscience & Biobehavioral Reviews*, vol. 120, 2021, pp. 236-48, doi:<https://doi.org/10.1016/j.neubiorev.2020.11.025>.
- Rouse, Cecilia. "The Economics of Education and Policy: Ideas for a Principles Course." *The Journal of Economic Education*, vol. 48, 2017, pp.229-237, doi: 10.1080/00220485.2017.1320611.
- Rovner, Joshua. "Youth Justice by the Numbers." <https://www.sentencingproject.org/policy-brief/youth-justice-by-the-numbers/>.
- Sabatine, E. et al. "The Unique and Interactive Effects of Parent and School Bonds on Adolescent Delinquency." *J Appl Dev Psychol*, vol. 53, 2017, pp. 54-63, doi:10.1016/j.appdev.2017.09.005.

- Shandra, C. L. and D. P. Hogan. "Delinquency among Adolescents with Disabilities." *Child Indic Res*, vol. 5, no. 4, 2012, doi:10.1007/s12187-012-9135-9.
- Spoth, R. et al. "Increasing School Success through Partnership-Based Family Competency Training: Experimental Study of Long-Term Outcomes." *Sch Psychol Q*, vol. 23, no. 1, 2008, pp. 70-89, doi:10.1037/1045-3830.23.1.70.
- tc.columbia.edu. "Cost of Inadequate Education to Society Is Hundreds of Billions of Dollars, Researchers Say." Teachers College, Columbia University <https://www.tc.columbia.edu/articles/2005/october/cost-of-inadequate-education-to-society-is-hundreds-of-billi/>.
- The Council of State Governments Justice Center. "Reducing Recidivism and Improving Other Outcomes for Young Adults in the Juvenile and Adult Criminal Justice Systems." 2015.
- U.N. Global Study. "Regional Imprisonment Rate of Children." Translated by Chart appearing in Nowak, 2019.
- U.S. Department of Education. "Every Student Succeeds Act (Essa)." <https://www.ed.gov/essa?src=rn>.
- . "Individuals with Disabilities Education Act." <https://sites.ed.gov/idea/>.
- U.S. Department of Education and U.S. Department of Justice. "Guiding Principles for Providing High-Quality Education in Juvenile Justice Secure Care Settings." 2014.
- Vacca, James S. "Crime Can Be Prevented If Schools Teach Juvenile Offenders to Read." *Children and Youth Services Review*, vol. 30, no. 9, 2008, pp. 1055-62, doi:<https://doi.org/10.1016/j.chilyouth.2008.01.013>.
- Wang, Xia et al. "Comparison of the Educational Deficiencies of Delinquent and Nondelinquent Students." *Evaluation Review*, vol. 29, no. 4, 2005, pp. 291-312, doi:10.1177/0193841x05275389.
- Whitaker, Amir et al. "Cops and No Counselors: How the Lack of School Mental Health Staff Is Harming Students." 2019.
- Willcutt, E. G. and B. F. Pennington. "Psychiatric Comorbidity in Children and Adolescents with Reading Disability." *J Child Psychol Psychiatry*, vol. 41, no. 8, 2000, pp. 1039-48.
- Zablotsky, B. et al. "Prevalence and Trends of Developmental Disabilities among Children in the United States: 2009-2017." *Pediatrics*, vol. 144, no. 4, 2019, doi:10.1542/peds.2019-0811.
- Zimskind, Limor et al. "Orange County Developmental Screening Pilot Project." 2009.

## **The Impact of Hay Fever on Public Health in Japan By Sara Tiwari**

### **Abstract**

Hay fever, or allergic rhinitis, significantly impacts public health in Tokyo, Japan due to high levels of airborne pollen from cedar and cypress trees. This study examines the perceptions of 169 individuals with hay fever living in Tokyo regarding the health complications and psychological effects of hay fever. Results indicate that most respondents have experienced symptoms for over five years. At least one in five participants reported a negative impact on their daily lives, and nearly one in three experienced reduced productivity. Despite the widespread use of over-the-counter medications, several respondents found them ineffective or problematic due to side effects. The study highlights the need for personalised treatment strategies and improved public health measures. Integrating awareness campaigns and health technology, such as mobile applications, could enhance symptom management and patient outcomes. These findings emphasise the importance of developing targeted interventions and further research to address the broad impact of hay fever in urban settings like Tokyo.

### **Introduction**

Hay fever, also known as allergic rhinitis, is a common allergic condition characterised by symptoms such as sneezing, runny nose, and itchy eyes, among others that usually affect the eyes, nose, and throat. In Japan, the prevalence of hay fever has been steadily increasing, with seasonal outbreaks that affect a substantial portion of the population—over 40% (Klein and Nippon Research Center). Tokyo, as a densely populated urban province, experiences high levels of airborne pollen from various sources. These sources particularly include cedar and cypress trees, planted in abundance after World War II to reforest the country rapidly and boost timber production, thus making the country, and especially the Kanto region, a prime location for studying the impact of hay fever on public health (Otake).

The importance of understanding the implications of hay fever extends beyond the physical discomfort it causes. There is reason to believe this condition has broader public health implications, including increased healthcare utilisation, decreased productivity, and adverse effects on mental health. Despite the availability of various treatment and prevention strategies in Japan, hay fever remains a significant burden for many individuals, particularly during peak pollen seasons from February to April, which is one-fourth of a year, every year.

Inomata et al. performed a cross-sectional study using a smartphone application called AllerSearch to understand the part that nasal and ocular symptoms play in hay fever. The study found a clear relationship between the severity of hay fever symptoms and reduced quality of life. Symptoms like nasal congestion, facial itchiness, redness of eyes, and tearing were strongly associated with decreased quality of life, with the research demonstrating that there is a wide range of variability in how hay fever affects an individual's quality of life. Patients experienced different combinations of symptoms, and these variations had a substantial impact on their overall well-being, indicating that hay fever should not be treated as a one-size-fits-all condition;



personalized intervention strategies are needed to address these differences. This can lead to more efficient use of healthcare resources and better health outcomes on a national level. The research highlights the importance of multidisciplinary efforts in managing hay fever, as it affects a significant portion of the Japanese population and can lead to substantial medical costs and economic losses. The Inomata et al. study concludes that health apps like AllerSearch can facilitate comparably better clinical care by providing personalised data and customised regimens for patients with hay fever.

Another cross-sectional study, by Muzalyova and Brunner, used a self-reported survey to analyse the choices made by those suffering from hay fever concerning their condition. From the survey, individuals could be grouped into two sets, with one being those who actively seek medical support and the other being those who self-manage symptoms. From their surveyed sample, over 70% of individuals prefer to manage the allergy themselves with only 40% opting for being under medical supervision and 30% utilising immunotherapy. The two aforementioned groups are not mutually exclusive, though, with four in five individuals with hay fever using anti-allergic medicines, regardless of whether or not they self-manage or have consulted a medical professional. A majority of survey participants believed that the drawbacks associated with managing their hay fever were minimal in comparison to the benefits of doing so; however, those with more severe symptoms, as perceived by themselves, were more likely to believe that the effort required to manage their symptoms was greater than if their symptoms were milder. Overall, the perceived severity of one's own symptoms, the perceived seriousness of hay fever as a whole, and the effectiveness of perceived self-efficacy in managing symptoms had great impacts on whether or not an individual takes any kind of action to manage their hay fever. The paper's concluding note was that despite a significant number of allergic individuals opting for self-reliant allergy management rather than seeking medical supervision, given the widespread use of over-the-counter anti-allergy medication, seeking advice from medical professionals, even at the pharmacy level, is a valuable alternative. Educational interventions to improve allergy management by increasing awareness of risks, especially for individuals with inadequate allergy management, can potentially improve one's quality of life.

Australia, another country in the Asia Pacific region, faces seasonal hay fever at levels comparable to Japan, with allergic rhinitis affecting 17% of the population there (Medek et al.). The Medek et al. study investigated the impact of climate change on human health, specifically focusing on hay fever sufferers in Canberra, Australia. The authors found a significant relationship between the severity of nasal and itch symptoms reported by hay fever sufferers and the levels of pollen in the environment in Canberra. After a certain threshold of pollen concentration was reached, symptom scores remained high, indicating a strong association between pollen load and nasal symptoms. This study suggests, similarly to the Inomata et al. study, that online methods could be beneficial for managing allergic rhinitis and could be adapted to monitor other health conditions affected by environmental variables, advocating for the potential of online tools to provide patient-specific information and clinical advice. The authors

recommend this strategy as being particularly useful in Australia since many hay fever sufferers do not consult healthcare professionals there, which is similar to the circumstances in Japan.

An older study, carried out in Nottingham, England by Richards et al., observed that over-the-counter treatments were used by 54% of sufferers, with drowsiness being a common side effect reported by one-third of the users. More people perceive themselves to be hay fever sufferers than those who have been diagnosed, suggesting that many sufferers self-treat without consulting a doctor, and, thus, it is harder to obtain accurate statistics.

It is evident that conducting novel research in the field of hay fever is of utmost importance, especially research focusing on Japan and the work productivity of the urban population. This research aims to investigate the impact of hay fever on public health in Tokyo, focusing on its prevalence, associated health complications, and psychological implications.

Understanding the full scope of hay fever's impact can prove to be useful when developing public health strategies and policies. By identifying the specific challenges faced by individuals, this study aims to gain deeper insight into how the burden of hay fever can be alleviated and how quality of life can be improved for affected individuals. Furthermore, the findings of this study have the potential to contribute to the global body of knowledge on allergic conditions and their management, providing insights that may be applicable to other regions with similar environmental and demographic characteristics.

## **Methodology**

### **Research Design**

This study utilised a mixed-methods approach, combining quantitative research techniques with some qualitative features to gain a comprehensive understanding of the impact of hay fever on public health in Tokyo. The primary research involved administering a semi-structured, self-reported online survey, and the responses were analysed to understand how people in Japan with hay fever perceive the disorder.

### **Data Collection Methods**

#### **Survey Design**

The primary data collection tool was a semi-structured self-reported survey, designed to capture both quantitative and qualitative data. The survey included multiple-choice questions to gather statistical information on the prevalence of hay fever, the severity of symptoms, the side effects experienced, medication usage, mitigation strategies, and the psychological impact of hay fever. The survey also featured short-answer questions and one optional long-answer open-ended question, all of which to collect qualitative data on individual experiences with hay fever. Thus, this design aimed to comprehensively provide statistical information while capturing the unique, personal stories of those affected by hay fever, reflecting the varied intensity and nature of the condition.

## Survey Distribution

To ensure a broad and diverse sample, the survey was distributed via various social media platforms, informing all viewers of the posts or broadcast messages that this research has a predominant focus on Tokyo and hay fever. The survey was open to anyone, whether or not they had been diagnosed with hay fever, to capture a wide range of experiences and perceptions. The results are to only be used for research purposes, and all participants were made aware of this before they could fill out the survey.

## Results

The survey was open between 24 April 2024 and 26 May 2024, and 357 responses were received during this time. A total of 190 individuals (53.2% of total respondents) reported having hay fever; this included those who have either been diagnosed or suspect that they have it. Of the 190 people, 169 currently reside in Japan. This dataset of 169 was used to conduct analysis as this group solely consists of Japanese residents who have hay fever, with the term “have” being used loosely as described above. The demographic details of the individuals who reported having hay fever symptoms can be found in Table 1.

**Table 1: Demographics (n=169)**

### Age

Under 18	8
18-27	33
28-37	73
38-47	47
48-57	8
Over 57	0

### Highest level of education completed

Less than high school diploma	4
High school diploma	19
Associate's degree	6
Bachelor's degree	91
Master's degree	38
Professional / Doctoral degree	11

### Employment status

Full-time worker	112
Part-time worker	21
Student	28
Not working	8

Category of residential area

Rural	7
Suburban	40
Urban	122

**Table 2: Presentation of hay fever (n=169)**

Diagnosis of hay fever

Formally diagnosed by a doctor	94 (55.6%)
Suspected, but not diagnosed	75 (44.4%)

Time elapsed since first onset of symptoms

Less than 1 year	22 (13.0%)
1 to 5 years	37 (21.9%)
More than 5 years	110 (65.1%)

Common symptoms by severity (1= low; 5 = high)

	1	2	3	4	5
Sneezing	13 (7.7%)	25 (14.8%)	55 (32.5%)	39 (23.0%)	37 (21.9%)
Runny nose	9 (5.3%)	33 (19.5%)	35 (20.7%)	41 (24.2%)	51 (30.2%)
Itchy eyes	20 (11.8%)	23 (13.6%)	35 (20.7%)	46 (27.2%)	45 (26.6%)
Sore throat	109 (64.5%)	13 (7.7%)	29 (17.1%)	9 (5.3%)	9 (5.3%)

Other symptoms reported by at least 5% of respondents

Itching / Rashes	22 (13%)
Difficulty in breathing	17 (10%)
Cough	15 (8.8%)
Headache	15 (8.8%)
Fatigue	11 (6.5%)

As Table 2 shows, the majority of respondents have been experiencing symptoms of hay fever for over 5 years. Symptoms with the greatest incidence of high severity are runny nose and itchy eyes, with over half the respondents reporting a severity of 4 or 5. Sore throat does not appear with severe symptoms in this group, with over 70% of the respondents reporting a severity of 1 or 2.

Some other commonly reported symptoms include itchiness or rashes, difficulty in breathing, cough, headache, and fatigue. Some uncommon but concerning symptoms experienced by respondents were nausea, dizziness, and fever.

**Table 3: Perceived impact of hay fever symptoms on the quality of life (n=169)**

Impact of hay fever symptoms on overall quality of life

Negative	138	(81.6%)
No impact	30	(17.7%)
Positive	1	(0.7%)

Affects ability to perform daily activities

Always	16	(9.5%)
Frequently	58	(34.3%)
Occasionally	60	(35.5%)
Rarely	35	(20.7%)

Effect on productivity at work or school

Reduced	112	(66.3%)
Increased	5	(3.0%)
No impact	52	(30.8%)

Days of work/school missed in a year due to hay fever symptoms

0	129	(76.3%)
1-5	35	(20.7%)
Over 5	5	(3.0%)

Increase in stress or anxiety due to hay fever symptoms

Yes	88	(52.1%)
No	81	(47.9%)

When asked how they perceive the impact of hay fever symptoms on their overall quality of life, the overwhelming majority of respondents (81.6%) reported a negative impact. 66.3% reported lower productivity in school or work, and 43.8% reported that their hay fever symptoms always or frequently interfered with their ability to do their daily activities. In addition, 52% of respondents reported an increase in stress or anxiety due to their hay fever symptoms. While most respondents did not report missing any days of work due to their symptoms, one person did report losing as many as 25 working days in a year due to the severity of their symptoms.

**Table 4: Steps to mitigate hay fever symptoms**

Measures to reduce exposure to pollen (n = 169)

	Always	Sometimes	Never
Wearing a mask	37 (21.9%)	90 (53.3%)	42 (24.8%)

Wearing protective glasses	17 (10.1%)	21 (12.4%)	131 (77.1%)
Avoiding outdoor activities	37 (21.9%)	83 (49.1%)	49 (29.0%)
Using air purifiers	13 (7.7%)	7 (4.1%)	149 (88.2%)

Usage of over-the-counter (OTC) medication (n = 169)

Never tried	16	9.5%
Did not have a significant effect	11	6.5%
Could not manage the side effects	8	4.7%
Symptoms reduced slightly	67	39.6%
Symptoms reduced by a lot	67	39.6%

Types of OTC medication tried (n = 153)

Antihistamine tablets	118	77.1%
Nasal sprays	62	40.5%
Eye drops	36	23.5%

Reasons for not trying/discontinuing OTC medication (n (Never tried) = 16; n (Discontinued) = 17)

	Never tried	Discontinued
Worried about side effects	1	3
Symptoms not severe enough	5	2
On prescription medication	3	2
Financial constraints	2	0
OTCs not effective	1	7
Prefer non-medical methods	5	4

Using masks and avoiding outdoor activities during periods of high pollen in the environment were the most common measures used by respondents to minimise their exposure to pollen. Nearly 90% did not use an air purifier to filter out pollen in their homes or vehicles. Some other methods to minimise pollen exposure that respondents mentioned were frequent washing, especially after returning home (9), using a face or nasal spray that traps pollen (6), drying laundry indoors (5), and removing outer garments like overcoats, before entering the house (5).

Over 90% of respondents report having tried over-the-counter (OTC) medication, however only 39.6% report seeing a substantial reduction in symptoms from them. While 16 respondents reported not using OTCs, from the responses it would appear that an additional 17 respondents have discontinued OTCs. The most commonly reported reason for not trying OTCs was not having sufficiently severe symptoms and preferring non-medical methods of controlling symptoms. The most commonly reported reason for discontinuing OTCs was that they were not effective enough.

Among the OTCs used, antihistamine tablets were the most common (77.1%), followed by nasal sprays (40.5%).

Finally, only seven respondents reported undergoing immunotherapy to manage their hay fever symptoms. Five received sublingual treatments and two received antigen injections.

## **Discussion**

The findings of this study show that hay fever has fairly severe effects, with over half of the respondents reporting severe runny noses and itchy eyes, and small but significant numbers reporting other more impactful symptoms like itchiness, difficulty in breathing, and headaches. Unsurprisingly, hay fever also has a negative impact on people's stress levels, productivity and quality of life. Oddly, five respondents reported increased productivity due to hay fever symptoms; this is likely an error.

The fact that only 94 people (55.6%) had a formal diagnosis seems to be in line with the findings by Muzalyova and Brunner. One surprising finding was that although 94 of the 169 had received a formal diagnosis from a medical professional, as many as 110 respondents reported having hay fever symptoms for over five years. A closer examination of the data revealed that of the 110, only 75 (68.1%) had a formal diagnosis. While this is significantly higher than the overall surveyed population, about one-third of those with long-term symptoms were trying to manage them on their own. This indicates the need to build awareness among the public to seek a diagnosis and find appropriate measures to manage symptoms. Awareness programs targeting general physicians to encourage their patients to seek appropriate treatments may also be useful.

One unfortunate finding is that at least two respondents reported not taking OTC medications due to financial constraints. This observation suggests that the economic burden of hay fever is an overlooked issue. Given the chronic nature of the condition, recurring costs for medications may create financial strain, particularly for those without access to insurance or subsidies. This finding aligns with broader discussions in public health regarding equitable access to healthcare resources, highlighting the need for effective education and improving access to the tools needed to manage hay fever.

A significant finding in the study was the direct relationship between hay fever severity and psychological impacts. A substantial majority (81.6%) reported hay fever having a negative effect on overall quality of life. This aligns with the research conducted by Inomata et al., which noted that symptoms of hay fever exceeding a personal threshold of tolerance were strongly associated with decreased quality of life. The Inomata et al. study also reached the conclusion that there is a wide range of variability in how hay fever affects an individual and their quality of life. This is reinforced by the results of this study as well. Although only 20.7% of respondents claimed that hay fever symptoms rarely affect their ability to perform daily activities, among the other 79.3%, there was great variability in how often their symptoms affect their ability to perform daily activities, with an insubstantial 9.5% claiming that hay fever "always" affects their ability to perform daily activities. Ultimately, this aspect of the study reinforces the hypothesis

that hay fever's impacts extend beyond physical discomfort and have tangible effects on mental health.

One of the most significant findings of this study is that 66.3% of respondents reported reduced productivity due to hay fever symptoms. This result is particularly concerning when considering Japan's strong work culture, where long working hours and expectations of overtime are extremely common due to this being regarded as the hallmark of a hardworking employee in Japanese culture. In the 2015-2016 fiscal year, over one thousand cases of *karōshi* were registered by the Japanese government (Duarte). This phenomenon of *karōshi* (過労死), translated to "death from overwork" in English, and the normalisation of long hours may mean that workers feel pressured to maintain high levels of productivity, even while suffering from conditions like hay fever. This is further emphasised by the fact that, despite experiencing significant symptoms, 76.3% of respondents do not miss any days of work or school during hay fever season each year. This suggests that many may feel compelled to continue attending work or school, even when unwell, to meet societal and workplace expectations.

Hay fever's debilitating symptoms as well as the side effects of OTC medication being used to minimise the impact of a particular symptom but consequently worsening the effect of others such as fatigue and headaches, as observed by the survey responses, most likely compounds the issue, especially during the spring pollen season. Reduced productivity during regular working hours may lead employees to compensate by working longer hours, thereby creating a vicious cycle where physical well-being is sacrificed to meet job demands. This is supported by the observation that 66.3% of respondents reported lower productivity, yet overtime is a common solution to make up for such shortfalls. In turn, longer work hours may also exacerbate the stress and fatigue caused by hay fever, worsening both mental and physical health outcomes. Moreover, the concept of presenteeism, when employees are physically at work but not fully functioning due to illness, could also be prevalent. This stresses a key point for public health intervention: improving awareness of hay fever's impact on productivity, especially by encouraging businesses to adopt more flexible or accommodating measures during peak hay fever seasons, such as offering remote work options or allowing flexible work hours. As of the Japanese spring season in 2024, the Japanese government has already begun urging companies to offer remote work during hay fever season, which is a positive step that should be expanded nationwide (Okada).

At the end of the survey, respondents were asked to add any other observations or perceptions they had about their hay fever symptoms. The responses in this section had some noteworthy points for further exploration:

- Many respondents who had moved to Japan from elsewhere reported that their symptoms were mild or non-existent in their home countries. However, three respondents reported the reverse: their symptoms were milder in Japan than in their home countries.
- One person reported fewer symptoms since they adopted a pet dog and were forced to walk outdoors more, thus increasing their exposure to pollen. This is perhaps a case of accidental immunotherapy.



- While most respondents were allergic to pollen from Sugi trees (*Cryptomeria japonica*), some mentioned that they were more allergic to pollen from Sakura (*Prunus serrulata*), or other fruit trees.
- At least two respondents reported that their symptoms became milder after they began consuming probiotic foods regularly, indicating a possible immunomodulatory effect of probiotic bacteria on anti-pollen immunoglobulins.
- Drowsiness was reported as a common side effect among those who took antihistamines.

## Conclusion

It must be noted that this survey was conducted entirely online and in English. Both of these factors limit its reach and applicability to the wider population of Japan or even the Tokyo region. Additionally, since this was a self-reported survey, the respondents' claims regarding the severity of symptoms or specific mitigation measures cannot be verified. However, the number of people reporting severe symptoms, reduced quality of life, and increased stress all strongly indicate that most people suffering from hay fever perceive significant negative impacts and would benefit from public health measures that help manage and mitigate the disorder. Awareness campaigns could help individuals recognise their symptoms earlier and seek medical treatment, with health technology in the form of mobile applications, a commodity available to a majority of the Japanese urban population. As Inomata et al. showed, smartphone applications have great potential to benefit people suffering from hay fever by providing personalised regimens to mitigate and manage their symptoms. Moreover, public health strategies should focus on expanding access to affordable OTC medications and promoting environmental reforms, such as reducing air pollution in urban areas, which may help alleviate symptoms for many sufferers. The above findings regarding the relationship between hay fever symptoms and productivity highlight the critical situation of Japan in a broader cultural context. Without proper interventions, hay fever could continue to significantly affect work-life balance and overall public health.

This small-scale study ultimately emphasises the need for further research to understand the broader prevalence and perception of hay fever in Japan and find effective methods to mitigate its effect on the population. Future research should also look deeper into the impact of hay fever on productivity and the potential reasons behind it. Corporations, in particular, need to conduct internal analyses to determine why employees continue to work despite reduced productivity—whether due to societal pressure, fear of job loss, or lack of support systems. There could also be important differences in how people of different age groups, genders, or employment levels respond to these pressures. Understanding these dynamics could lead to more targeted and effective workplace policies and interventions that better support employees suffering from hay fever.

## Works Cited

- Klein, Catharina and Nippon Research Center. "Share of people suffering from hay fever and/or pollen allergies in Japan from 2003 to 2019." *Statista*, 9 Jan. 2024, [www.statista.com/statistics/831561/japan-hay-fever-pollen-allergy-share](http://www.statista.com/statistics/831561/japan-hay-fever-pollen-allergy-share).
- Otake, Tomoko. "As hay fever season approaches, here's what you need to know to prepare." *The Japan Times*, 4 Feb. 2024, [www.japantimes.co.jp/news/2024/02/04/japan/science-health/otake-national-pollen-allergy-qa](http://www.japantimes.co.jp/news/2024/02/04/japan/science-health/otake-national-pollen-allergy-qa).
- Inomata, Takenori, et al. "Individual multidisciplinary clinical phenotypes of nasal and ocular symptoms in hay fever: Crowdsourced cross-sectional study using AllerSearch." *Allergology International*, vol. 72, no. 3, July 2023, pp. 418–27, doi:10.1016/j.alit.2023.01.001.
- Muzalyova, Anna, and Jens O. Brunner. "Determinants of the utilization of allergy management measures among hay fever sufferers: a theory-based cross-sectional study." *BMC Public Health*, vol. 20, no. 1, 1876 (2020), Dec. 2020, doi:10.1186/s12889-020-09959-w.
- Medek, Danielle Eve, et al. "Hay Fever in a Changing Climate: Linking an Internet-Based Diary with Environmental Data." *EcoHealth*, vol. 9, no. 4, Oct. 2012, pp. 440–47, doi:10.1007/s10393-012-0787-1.
- Richards, S., et al. "How many people think they have hay fever, and what they do about it." *PubMed Central (PMC)*, vol. 42, no. 360, July 1992, pp. 284–86. *British Journal of General Practice*, [www.ncbi.nlm.nih.gov/pmc/articles/PMC1372087](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1372087).
- Duarte, Fernando. "Which country works the longest hours?" *British Broadcasting Corporation*, 8 May 2018, [www.bbc.com/worklife/article/20180504-which-country-works-the-longest-hours](http://www.bbc.com/worklife/article/20180504-which-country-works-the-longest-hours).
- Okada, Ei, editor. "Japan Government Urges Teleworking to Counter Hay Fever." *The Mainichi*, 5 Feb. 2024, [mainichi.jp/english/articles/20240205/p2a/00m/0na/010000c](http://mainichi.jp/english/articles/20240205/p2a/00m/0na/010000c).

# **Analysing the Transformative Effects of Chinese Government Border Policies on Hong Kong's Tourism Industry: Insights from COVID and Post-COVID Analysis**

**By Yu Wei Lin William**

## **Abstract**

This study empirically investigates the relationship between Hong Kong's economic performance and Chinese policy actions concerning COVID-19, with a specific focus on the tourism sector. The study encompasses several key interdependent variables: the influx of visitors from mainland China, economic indicators (notably GDP), and the enforcement of border policies between mainland China and Hong Kong. According to our empirical findings, the COVID-19 pandemic appears to have had a major detrimental effect on the number of incoming tourists, particularly because a significant proportion of the source population of tourists from mainland China has been lost. Additionally, our data point to a direct negative correlation between Hong Kong's overall economic performance and the implementation of border policies. The decline in GDP, coupled with the implementation of border policies, is found to be statistically significant, underscoring the critical role of the tourism industry in Hong Kong's economy. Taken together, these findings demonstrate the devastating effects of COVID-19 and offer a comprehensive overview of Hong Kong's economic recovery following the pandemic's management. Our study highlights the relationship between tourism recovery and pandemic control, with a particular focus on visits from mainland Chinese tourists.

## **Introduction**

Leveraging its unique geographical advantages and international harbour, tourism has consistently served as a strategic pillar of the Hong Kong economy. Prior to the COVID-19 pandemic, the city had broad international appeal, attracting 58.47 million and 65.15 million overseas tourists in 2017 and 2018, respectively (Hong Kong Tourism Board (HKTB), 2024). Among the substantial influx of tourists, visitors from mainland China constituted nearly 80% of the total in 2018, facilitated by the opening of the Hong Kong-Zhuhai-Macao Bridge and the high-speed rail in Hong Kong. Tourism contributed to around 3.6% of Hong Kong's GDP and employed 232 700 persons, accounting for about 6.0% of total employment. In light of the tourism sector's significance, the COVID-19 pandemic has resulted in substantial losses within this industry. In early 2020, the total number of visitor arrivals experienced a year-on-year decline of 88.2%, plummeting to 3.50 million in the first five months. The tourism sector in Hong Kong has suffered greatly as a result of COVID-19 as of 2021. Its employment rate has decreased to roughly 22,500, or 0.6% of total employment, and its contribution to the city's GDP has decreased to 0.1%. The annual average number of visitors from the Mainland and the rest of the world declined by 14.2% over 2018 to 55.91 million. We propose that the viral control measures taken by the Chinese and Hong Kong governments in the COVID-19 pandemic are a significant contributing factor, and are responsible for the shrinkage of these figures. (*Hong Kong: The Facts - Tourism*, 2020) (*Tourism*, n.d.)

During COVID-19, the Chinese government had drawn up a set of rules and various travel restrictions, notably cross-border regulation, social distancing requirements, and quarantine rules upon the borders between Hong Kong and Mainland China. The quarantine policy was implemented early and remained in effect until the conclusion of the pandemic response. This policy, along with cross-border restrictions, mandated that individuals entering mainland China undergo quarantine for 14 to 21 days. To visitors, the cost of time and money returning back to the mainland was huge. Additionally, frequently used control points, such as the Lo Wu Control Point and the Lok Ma Chau Control Point, were deactivated. Travellers were required to make reservations in advance to navigate the limited available slots at other control points. Apart from this, other COVID-19 policies such as scanning health codes in different venues, regulations on nuclei tests and masks are the primary causes of the loss of tourists from mainland China. After three years of the control of COVID-19, in 2023, the Chinese government and Hong Kong government have announced the reopening of borders and easing relevant policies (Healthcare, n.d.). This paper aims to analyse how changes in COVID-19 policies have impacted Hong Kong's tourism sector. Specifically, this research addresses the impact of the Chinese government's border policy adjustments on Hong Kong's economic performance. By utilising data and relevant information, this study seeks to provide insights into the extent and manner in which tourism may recover and stimulate Hong Kong's economy.

Our empirical approach is set upon a thorough analysis of a combination of primary and secondary data sources. By collecting the data from pre-COVID, during the pandemic, and post-COVID, we could have a broad view on the significance of the virus worldwide, and to know the effectiveness of the Chinese government's policies. By utilising and employing econometric modelling techniques such as the Regression model, a close relationship between border policy changes, visitor arrivals from mainland China, and Hong Kong's economic performances is unveiled. These policies and restrictions not only limited the travel aspirations of many visitors to Hong Kong but also adversely affected the local economy.

This study contributes to the existing body of knowledge on the interaction between quarantine laws and tourism dynamics. Existing papers mainly focus on the virus itself or the overall and forecasted impact of COVID. For example, Zhang et al. (2021) forecasts the recovery of tourism demand in response to the unanticipated effects of crises by integrating qualitative and quantitative models (Zhang et al., 2021). This paper, instead, will specifically focus on the mainland China group of tourists and the specific effect of the border policies implemented earlier, considering that this group of tourists has made up a large portion of Hong Kong tourism. The emerging topic of post-COVID life and its effects on Hong Kong tourism has generated high expectations for recovery. The unique context of Hong Kong, under the "One Country, Two System" framework, provides a rich area for research into variations in data and trends before and after the abolition of border policies. Through an examination of post-COVID patterns, we provide perspectives on how policy modifications between Hong Kong and the Chinese mainland might be used to promote economic recovery, particularly expansion in the tourism industry. This study enhances the discourse on policy by elucidating the potential

benefits of modifying border regulations to foster economic resilience. In addition, readers from relevant businesses from the tourism sector are able to gain a deeper understanding on consumers' post-COVID behaviours, which could provide them references on drawing up business strategies.

## **Literature Review**

### **Introduction to COVID-19 and an overview of its impact**

On December 19, a series of acute atypical respiratory diseases occurred in Wuhan, China. The disease rapidly spread throughout China and to other countries worldwide. The virus exhibited approximately 80% homology with SARS-CoV, leading to its designation as severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2, 2019-nCoV) (Yuki et al., 2020). The disease caused by the virus was called Coronavirus disease 19 (COVID-19) and the World Health Organisation (WHO) declares a Public Health Emergency of International Concern (PHEIC) on 30 January 2020 and to characterise the outbreak as a pandemic on 11 March 2020 (World Health Organization, 2024). This global emergency has had a profound impact, affecting approximately 200 countries and territories. As of 10/03/2023, total cases of 676,609,955 are recorded and 6,881,955 deaths had been reported according to the Coronavirus Resource Centre from John Hopkins University and Medicine (John Hopkins University, 2023).

As COVID-19 rapidly spread globally, developing effective responses and precautions against the epidemic became a significant concern for governments worldwide. China, characterised by its unique political environment, is one of the few countries that opted for a persistent approach to combating the virus through proactive risk control and regulatory measures. Some of the reasons for China's successful control of the new coronavirus are timely, effective, and transparent epidemic response capacity, successful case identification and large-scale surveillance, duly city lockdown and social distancing regulations, and impeccable healthcare facilities and medical team preparations (Altakarli, 2020). Nevertheless, although the strict government regulations might be considered successful in controlling the outbreak of the virus, the strike on the overall economy is inevitable.

### **Tourism in the context of Hong Kong and the pandemic**

As a key economic pillar, tourism has significantly influenced Hong Kong's economy. Consequently, the industry experienced substantial challenges due to the pandemic. Prior to the pandemic, international tourist arrivals in Hong Kong increased from 13.6 million in 2000 to 55.9 million in 2019, according to the Hong Kong Tourism Board (*Tourism Statistics | Hong Kong Tourism Board*, n.d.). Revenue from such tourism receipts rose from US\$7.9 billion in 2000 to US\$32.8 billion in 2019 (*UNWTO World Tourism Barometer*, n.d.). The signing of the Individual Visitor Scheme between the Mainland and Hong Kong in 2003 exacerbated these trends. Mainland tourist arrivals increased fivefold, from 8.4 million in 2003 to 2019. Mainland tourists accounted for about 78% of Hong Kong's tourism market share. Tourism—especially

from mainland China—became to form one of Hong Kong’s central economic pillars (C. L. Tai et al., 2022). However, the unprecedented impact of COVID-19 resulted in heightened restrictions on borders, social interactions, and the scope of activities. This situation resulted in a loss of not only international tourists but also the largest group of visitors from mainland China.

Among the anti-epidemic policies implemented by the Hong Kong and Chinese governments, border restrictions, including quarantine measures, were among the most influential factors leading to a sharp decline in visitor numbers, significantly impacting tourism. Border control points between mainland China and Hong Kong are more stringent than those for foreign entries, reflecting mainland China's rigorous closure and control measures in response to COVID-19. The implementation of Hong Kong’s border restriction policies with mainland China occurred in three distinct stages. In the early stage, after the first infection was identified, the HK government tightened its border controls with one city in mainland China. In accurate, travellers from Hubei Province or with the history of visiting the place in the past 14 days were not allowed to enter. Subsequently, visa and entry permits for all mainland cities were suspended, and border crossing methods were closed. Both the frequently used control points, the Lo Wu Control Point and the Lok Ma Chau Control Point, were deactivated, and daily passage quotas were established for the Shenzhen Bay Control Point and the Hong Kong-Zhuhai-Macao Bridge Hong Kong Port. The maximum number of reservations allowed per day is 1000 for both of the control points at the early stage of city lockdown. (The Government of the Hong Kong Special Administrative Region - Press Releases, 2022) These actions mark the early stage of border control (between January 23 and February 7). The second stage is enforcing strict border restrictions and banning the entry of most mainland travellers. On February 8th, compulsory 14-day quarantine for all visitors with travel history to mainland China within the past 14 days was set (ZHU & TAN, 2022). At this juncture, the tourism industry experienced an effective halt. The total number of visitors, including a significant proportion of tourists from mainland China, declined markedly. Data from the Immigration Department shows that, in just one and a half months, the number of people entering Hong Kong through border control points and the airport fell by 73%, from 121,828 on February 4 to 32,216 on March 18, following the implementation of the strict entry restriction and mandatory home-based quarantine measures (入境事務處, n.d.).

Affected by the pandemic of COVID-19, Hong Kong’s economy has fallen into a great recession. The global economy has also been severely impacted. According to statistics, the global Gross Domestic Product (GDP) at current price shrinks to US \$84,960.9 billion in 2020, which is US \$2,364.64 billion less than the previous year, 2019, before COVID-19 struck the world (Statista, 2010). Hong Kong was particularly affected due to its open economy, which heavily relies on the service sector. In the first half of 2020, its GDP fell by about 9% in real terms, the worst half-year drop on record. Throughout 2020, the annual GDP shrinkage is over 6%, which is more severe than the 2.5% during the Global Financial Crisis in 2009 and the 5.9% during the Asian Financial Crisis in 1998 (*HKTDC Research*, n.d.). Local retail sales fell 27% in the first 10 months, unemployment rate increased from 3.3% in October-December 2019 to 6.3%

in September-November 2020, the significant drop of both inbound and outbound tourism from 2019's 2.8% and 0.8% to 0.2% and 0.2%, nearly 89% drop from US \$26.37 billion to US \$2.86 billion between 2019 and 2020 in the revenue generated from international tourism... (Statista, n.d.) (Statista, n.d.-a) These statistics illustrate the severe negative impact of the epidemic and associated anti-epidemic policies on the Hong Kong economy, particularly the tourism industry.

With the continued development of vaccines and people slowly adapting to the virus, the world chooses to coexist slowly with the virus as it embarks into roughly late 2022 and early 2023. In early 2023, the governments of Hong Kong and the Mainland officially announced the resumption of cross-border clearance at the two border gates, which signalled the slow recovery of Hong Kong's economy. With inbound tourism being the main key driver, in the fourth quarter of 2023, real GDP increased by 4.3% over the previous year, following a 4.1% increase in the preceding quarter. On a seasonally adjusted quarterly basis, real GDP increased by 0.5% in the fourth quarter, following 0.3% growth in the previous quarter. For 2023 as a whole, real GDP growth restarted at 3.2% in 2023, after contracting by 3.7% in 2022. The anticipated recovery of inbound tourism and private consumption is expected to support labour demand, potentially reducing unemployment in the future. The reopening of Hong Kong's borders is likely to result in a substantial increase in visitor arrivals, thereby boosting the retail market and tourist consumption. Statistics have shown that the value of total retail sales grew notably by 16.2% in 2023 (Hong Kong Economy, 2019). In conjunction with government measures aimed at stimulating Hong Kong's economy, most sectors are expected to recover and strengthen compared to their performance during the epidemic.

In short, the COVID-19 pandemic reshaped Hong Kong's tourism pattern, highlighting the important role played by mainland Chinese tourists in Hong Kong's economic prosperity. The stringent border control policies enacted in response to the pandemic have resulted in a significant downturn, underscoring the dependence of Hong Kong's tourism industry on mainland visitors who traditionally formed a substantial majority of the market share.

## **Data and Method**

As previously introduced, this research paper centres on analysing the effects of alterations in border policies on Hong Kong's tourism sector and its impact on economic performances. Primarily, through the utilisation of a regression model, an estimation and consideration of the complete balanced data pertaining visitor arrivals from mainland China, border policy status, and Hong Kong's GDP spanning from January 2020 to September 2023, covering both pre-COVID-19 and post-COVID-19 eras, will be conducted. The data and information collected are mainly from three sources: online trends, documents officially issued by the Hong Kong government, and online databases.

Fig 1: Baidu Index data covers from Jan 2021 to Jan 2024 on a monthly basis: the chinese terms of "quarantine"(light blue) "tourism"(orange) "Hong Kong"(grey) "Pandemic in Hong Kong"(yellow) "Hong Kong government"(dark blue)

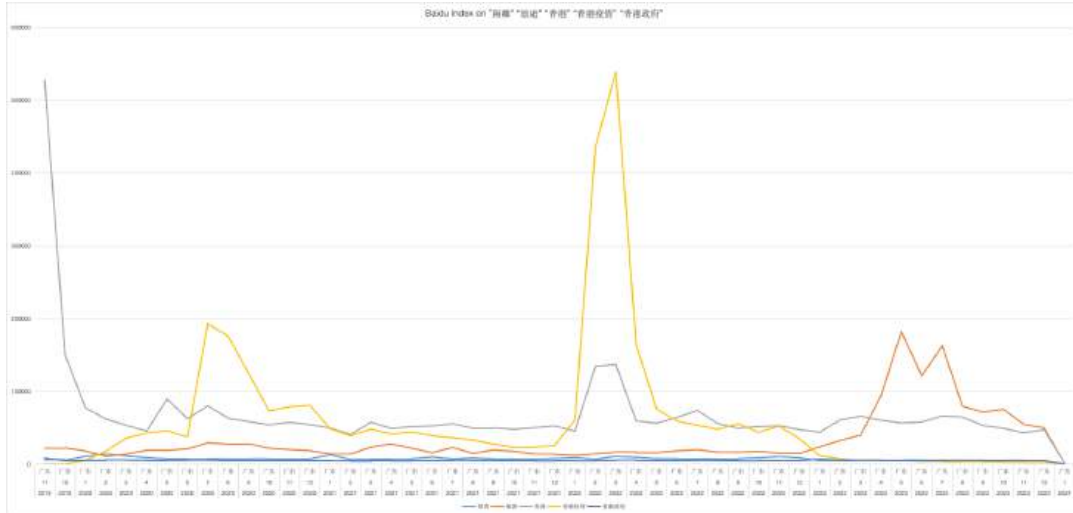


Fig.1 is an example of an online trend collected through the Baidu Index (also known as Baidu Zhishu, allowing users to look up the search volume and trend for certain hot keywords and phrases), considering five key words that are most relevant to the topic of research. As the daily cases rose significantly during February to April 2022 (Wong et al., 2022), the corresponding search frequency of keyword “Pandemic in Hong Kong” reached its peak during the period between March to June 2022. With the rise of uncertain news and people’s strong appeal to reopen the border since 2022, keywords such as “Hong Kong” (grey colour) had remained in a relatively high frequency compared to other keywords in people’s online research until late 2023. These high numbers of frequencies provide evidence that the pandemic in Hong Kong is an issue that people paid quite a lot of attention to during the period. Based on the fact that Baidu is the most common searching engine used in mainland China, they not only represent the large base of potential tourists or people that seek to travel to Hong Kong but also indicate the significance of changes in border policies. These key informations also provide insights to researchers on which time period we should focus on, in alignment with real world trends.



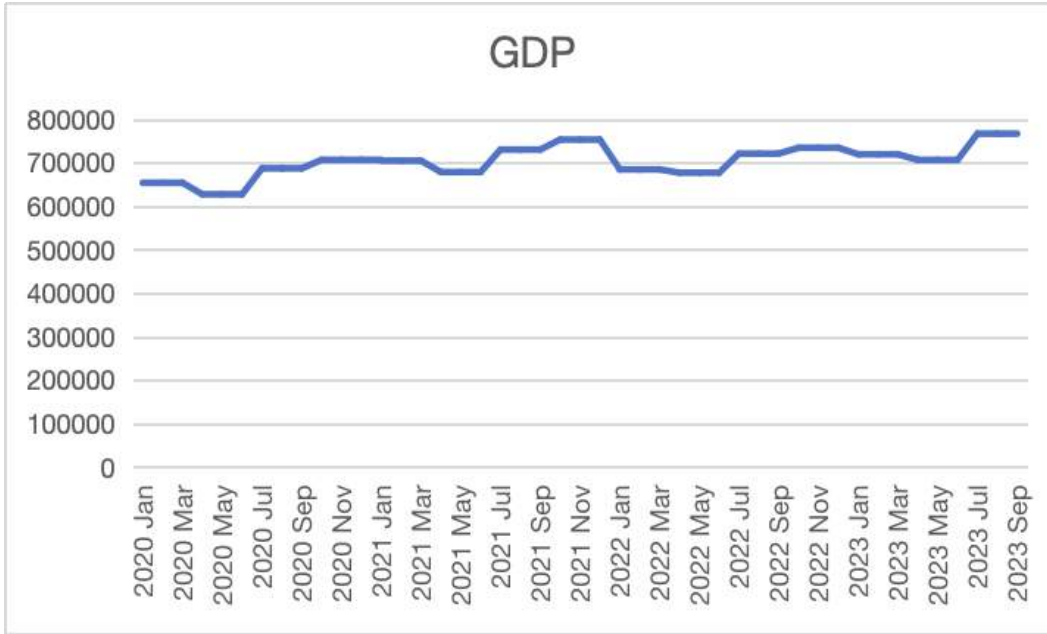


Fig 2: Hong Kong's GDP: The data covers from Jan 2020 to Sep 2023 on a monthly basis.



Fig 3: Visitor arrivals to Hong Kong from mainland China Note: The data covers from Jan 2020 to Sep 2023 on a monthly basis

Fig.2 and Fig.3 gathered sample data from The Census and Statistics Department of the government of the Hong Kong Special Administrative Region and Hong Kong Tourism Board (Census and Statistics Department, 2024). Fig.2 represents Hong Kong's Gross Domestic Product (GDP) from the start of 2020 to September 2023 on a monthly basis. The data collected is based on the units of million HKD and is an effective direct indication of Hong Kong

economy's performance. Fig.3 represents the number of visitor arrivals from mainland China over the same time period. The data shown on the figure is the exact number of visitors and is part of an integrated dataset of visitor arrivals sorted with a range of regions of residence found on Hong Kong Tourism Board. These data provide us a clear trend of how GDP and visitor arrivals from mainland China are affected by the pandemic in specific key times: early February 2020 and in the start of 2023. With the trends and significant timing suggested by the figures and the numbers, in the interpretation in empirical approaches, researchers are able to focus and analyse mainly between important variables that could successfully reflect the potential impact along with the change in government policies.

In the following data and method section, three regression models will be discussed, with a combination of mainly three variables, Hong Kong's GDP, availability of Hong Kong-mainland China cross border, and the number of visitors arrivals from mainland China. The section aims to offer valuable contributions and insights into the interdependencies among these specified variables.

### Regression Model 1

```
. reg visitorarrivalsmainlandchina borderpolicystatus
```

Source	SS	df	MS	Number of obs	=	45
Model	3.4456e+13	1	3.4456e+13	F(1, 43)	=	207.29
Residual	7.1475e+12	43	1.6622e+11	Prob > F	=	0.0000
Total	4.1603e+13	44	9.4553e+11	R-squared	=	0.8282
				Adj R-squared	=	0.8242
				Root MSE	=	4.1e+05

visitorarrivalsm~a	Coefficient	Std. err.	t	P> t	[95% conf. interval]
borderpolicystatus	-2104763	146189.3	-14.40	0.000	-2399581 -1809944
_cons	2122205	128926.8	16.46	0.000	1862199 2382211

Fig 4

In this model(as shown in Fig 4), visitor arrival from mainland China will be set as the dependent variable, with the border policy status as independent variable. This regression seeks to demonstrate the direct influence of border policies on the number of tourists to Hong Kong and find out the correlation between the independent variable and visitor arrivals, specifically from mainland China. In addition, the regression might also help to corroborate the supposition that the presence of border policy is one of the major reasons that affect visitors' decisions in visiting Hong Kong.

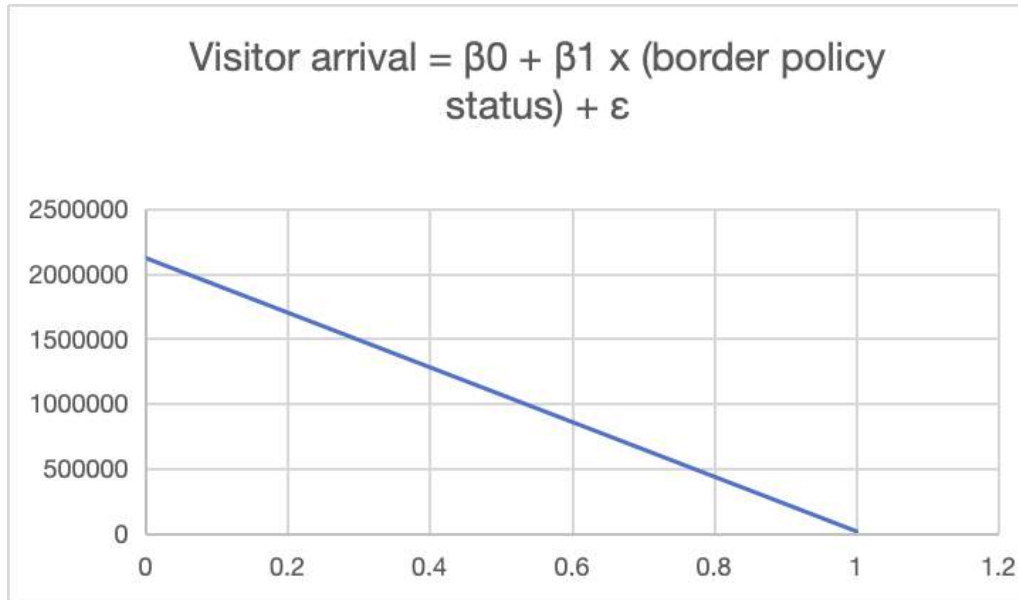


Fig 5: The data used in the model is the same with Figure 3.

According to the model, the coefficient of border policy is negative 2104763. This represents the change in the mean value of visitor arrivals when Border Policy Status changes from 0 to 1. The predicted mean visitor arrival value is 2122205, during the times when the policies were being absent. However, given that the coefficient is negative, it is concluded that there would be a significant predicted decrease in the mean visitor arrival value when the policies were present. The results of this model are in line with our previous conjecture - that border policies have a significant impact on visitors, and it can be hypothesised that this is one of the most important considerations in their decision. The claim is supported by the fact that there would be a 99.2% decrease in overall visitors, with the border policies being implemented. Fundamentally, this substantial decrease is impacted by the policy of allowing a limited number of visitors to cross the border. In fact, the complicated procedures in reaching the border, time-consuming quarantine process, both departing Hong Kong and entering Hong Kong, along with factors such as a high cost in hotel room fees and poor living standard during the quarantine were also the other external factors that had reduced the willingness of most people to leave and enter the region.

In conclusion, this regression helped researchers to confirm the significant negative relationship between the number of visitor arrivals and the status of the border policy, which can be explained by the consequences coming up with visitors' decisions in crossing the border. However, certain limitations were involved in the regression interpretation. A main limitation would be the simplification of the model. The model contained only one independent variable and did not take other influential factors into account. This omission may lead to omitted variable bias, where the excluded variables might distort the estimated relationship between border policies and visitor arrivals. Besides, while regression analysis can reveal associations between variables, it does not necessarily imply causation. Although the regression indicates that

change in border policy status could be associated with great changes in visitor arrivals, other factors could influence both variables independently. For example, government appeals and warnings could be one of the main reasons that caused the significant decrease in number of visitor arrivals.

## Regression Model 2

```
. reg visitorarrivalsmainlandchina gdp borderpolicystatus
```

Source	SS	df	MS	Number of obs	=	45
Model	3.4567e+13	2	1.7284e+13	F(2, 42)	=	103.17
Residual	7.0362e+12	42	1.6753e+11	Prob > F	=	0.0000
				R-squared	=	0.8309
				Adj R-squared	=	0.8228
Total	4.1603e+13	44	9.4553e+11	Root MSE	=	4.1e+05

visitorarrivalsm~a	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
gdp	1.465972	1.798522	0.82	0.420	-2.163593	5.095536
borderpolicystatus	-2067450	153736.7	-13.45	0.000	-2377703	-1757196
_cons	1061631	1307583	0.81	0.421	-1577178	3700440

Fig 6

This model(fig 6) aims to assess the impact of Hong Kong's GDP and border policy status on visitor arrivals from mainland China. Similar to the first model, dummy variables were created in order to make the classification of "border policy status" simpler. By estimating the model, we seek to find out to what extent do the presence of border policies reflect on tourists' willingness to visit Hong Kong. Therefore, by adding another independent variable "GDP" to the model, readers can have a clearer insight on the direct relationship between tourism and economic performance.

The above are the estimated results from conducting an Ordinary Least Squares regression (OLS). First, with the associated T statistics 0.81 and the p-value associated with the T-statistics 0.421 for the constant term, it is suggested that the term is not significant. Additionally, with a positive coefficient for the "GDP" variable, it is indicated that Hong Kong's economic performance is directly proportional to the amount of visitors from mainland China. This answers the question raised in our research question, the impact of the size of tourists on Hong Kong's overall economic performance. The resultant turns out to be consistent with our expectation that the greater the number of GDP, the more the inbound visitors. However, the associated T-statistics for the variable is relatively small, with only 0.82 (<2), which indicates that the coefficient is not statistically significant at the 5% or 10% level. Furthermore, the "border policy status" variable has a negative impact on the number of visitor arrivals, as for the presence of border policy, the number of visitors is expected to decrease by 2067450.

Considering the T-statistics as a large absolute negative number, the value of -13.45 suggests that the coefficient for the “border policy status” variable is statistically significant at 1% level. Before the investigation, we anticipated a link between the number of visitors and the border policy, but we were not expecting one that was this robust. A significant coefficient like this further confirms the Hong Kong government’s determination to strictly control the pandemic and the discouraging effects of various preventive measures on anyone wishing to enter the region.

From the interpretation of the model, researchers are able to obtain a positive relationship between economic activity and the number of visitors arrival. Although a causation relationship could not be concluded from the regression model, it is supported that the number of visitor arrivals from mainland China is one of the factors that contribute to Hong Kong’s economic performances.

While the regression model provides insights into the potential drivers of visitor arrivals from mainland China, it is essential to interpret the results with caution and consider the limitations. For the “GDP” variable, its relationship with the dependent variable “visitor arrival from mainland China” might not remain constant over time. The time during COVID and the brief post-COVID period make up the majority of the data included in the regression. Therefore, rather than the influence of tourism, Hong Kong’s economic performances would be more reliant on external factors, policies, and other economic conditions as the time ticked down to the government’s announcement of the resumption of border availability. Causality between variables is another limitation of the regression. Other factors, such as exchange rate, could also be a determinant of the variable “visitor arrival”. Thus, the regression analysis only shows an association between GDP, border policy status, and visitor arrivals. Causality is not established.

### Regression Model 3

```
. reg gdp borderpolicystatus
```

Source	SS	df	MS	Number of obs	=	45
Model	5.0388e+09	1	5.0388e+09	F(1, 43)	=	4.18
Residual	5.1791e+10	43	1.2045e+09	Prob > F	=	0.0470
Total	5.6830e+10	44	1.2916e+09	R-squared	=	0.0887
				Adj R-squared	=	0.0675
				Root MSE	=	34705

	gdp	Coefficient	Std. err.	t	P> t	[95% conf. interval]
borderpolicystatus		-25452.73	12444.2	-2.05	0.047	-50548.85 -356.6111
_cons		723461.3	10974.75	65.92	0.000	701328.6 745594

Fig 7

Our goal in using this regression model is to investigate the connection between the GDP of Hong Kong and the execution of the border policies. According to the model, GDP is impacted by two primary factors: a constant term that represents the value of GDP when border

policies are not in place, and a coefficient that represents the change in GDP as a function of border policy presence or absence and is represented by a dummy variable. By using this model, we want to shed insight on the specific effects of border policies on economic indicators and the wider effects that policy decisions have on Hong Kong's economy.

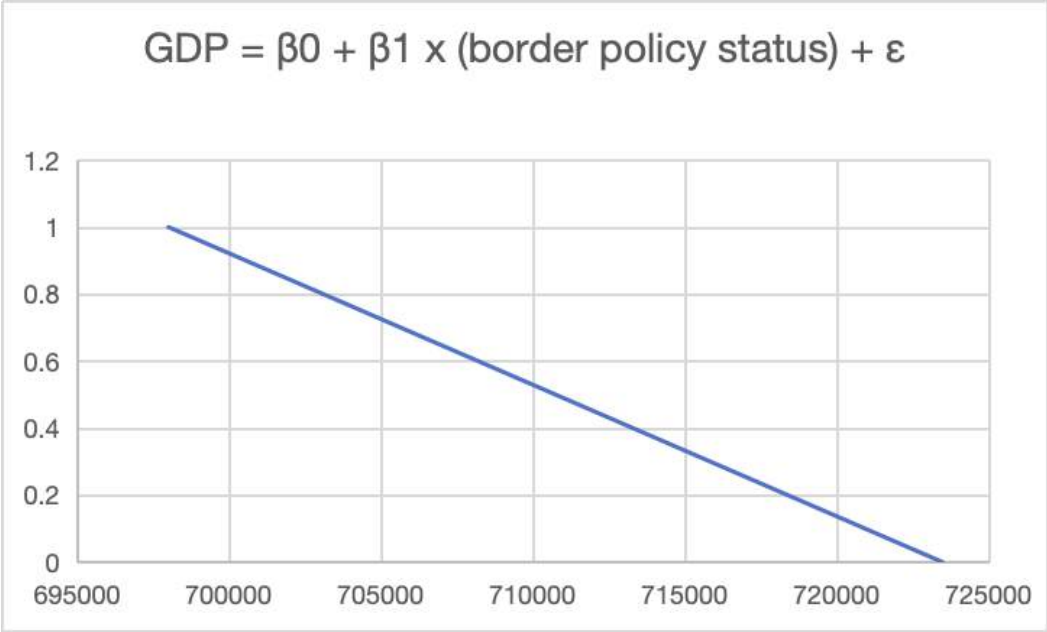


Fig 8: The data used in the model is the same with Figure 2.

In our regression model predicting GDP, the model resulted in a negative coefficient for the “border policy status” variable. This indicates that the presence of the border policy is associated with a decrease in GDP compared to when the policy is absent. This is consistent with our expectation that the implementation of border policies (i.e. limiting the number of visitors from mainland China) had always had an adverse impact on tourism, and Hong Kong’s economy as well. Since the term “border policy” in our study primarily refers to the border between Hong Kong and the mainland, this negative impact on GDP reflects the effect of a decrease in the number of tourists and visitors from China, revealing specifically the direct relationship between Hong Kong’s economy and pandemic-relevant measures. The drop of around 3.5% of the GDP is considered as significant, which thus enhances the conclusion that border policy has a significant impact on Hong Kong’s economy.

The negative coefficient implies that there may be economic implications associated with a change in border policy status, such as changes in trade, investment, or consumer behaviour. A plausible conjecture from researchers is that this portion of loss in GDP might be caused by the decrease in the total consumption or expenditure on the tourism sector. However, there are certain restrictions on the model. By limiting the explanatory variable in the model to border policy status alone, it may be oversimplified. Omitted variable bias may result from additional significant factors impacting GDP that are not taken into consideration by the model.

Furthermore, the model is unable to capture the immediate effects of border policy implementation due to GDP data gathered only quarterly.

## **Conclusion**

With a primary focus on the tourism industry, the study has demonstrated the connection between border policies and Hong Kong's economic performance. The impact of mainland Chinese tourists on Hong Kong's tourism sector has also been studied, as has how the implementation of pandemic-related rules changed travellers' choices. Our empirical approach leads to a number of important findings. First, there is a strong inverse correlation between the quantity of inbound tourists—especially those from mainland China—and the existence of border controls. The number of visitor arrivals, particularly from mainland China, has decreased significantly as a result of the abrupt and unprecedented application of such tight pandemic control regulations; in comparison to pre-COVID records, the majority of these arrivals are no longer there. Second, it is discovered that border policies and Hong Kong's economic performance are strongly correlated with the number of visitors from mainland China. There is a strong positive relationship between Hong Kong's GDP and visitor numbers in addition to the negative correlation between visitors and border regulations. This result highlighted the close association between two variables and highlighted the negligible impact of Hong Kong's tourism industry on the region's overall economic performance. Thirdly, there is a clear inverse relationship between Hong Kong's GDP and border policies. By contrasting the performance of the economy before and after COVID, this demonstrates the importance of the tourism sector to Hong Kong's economy and sheds light on the degree of the impact of the pandemic restrictions.

After combating the catastrophic worldwide pandemic for over two to three years, the globe has finally returned to normal, and the global economy is gradually recovering. The effects of the pandemic response on Hong Kong, a city with a highly liberal economic system and one that primarily depends on tourism for its development, are certainly something to research and debate in the context of the special governance of China and the unique relationship between mainland China and Hong Kong. Thus, the research question was raised. The COVID-19 dilemma has just ended, at least for the governments of China and Hong Kong. Hence, more research may examine the long-term socioeconomic effects of border controls enforced by the Chinese government on Hong Kong's tourism sector, as well as viable strategies for resilient and sustainable growth amidst evolving global conditions.

During the research process, certain limits are encountered. One of the largest challenges that we faced was the relatively short post-COVID period. The research is conducted in one year since the borders' policies were being adjusted. Therefore, this causes tremendous difficulty in collecting data. To add on, we could only do comparisons among small data sets which were collected in a short period of time. Moreover, as the effects brought by this once in a lifetime pandemic is significant and far-reaching, it is likely that more changes in regulations from the government or public trends would be committed. As a result, predictions and analysis on current trends could be transitory.

## Works Cited

- Hong Kong Tourism Board (HKTb). (2024). Visitor Arrival Statistics 2017/2018. Hktb.com. [https://securepartnernet.hktb.com/australia/en/research\\_statistics/research\\_publications/index.html?id=4109](https://securepartnernet.hktb.com/australia/en/research_statistics/research_publications/index.html?id=4109)
- Hong Kong: The Facts - Tourism. (2020). <https://www.gov.hk/en/about/abouthk/factsheets/docs/tourism.pdf>
- Tourism. (n.d.). [https://www.tourism.gov.hk/pdf/FactSheet/2020/tourism\\_fact\\_sheets\\_2020\\_en.pdf](https://www.tourism.gov.hk/pdf/FactSheet/2020/tourism_fact_sheets_2020_en.pdf)
- Healthcare, O. (n.d.). COVID-19 Timeline of Events - OT&P Healthcare. Www.otandp.com. <https://www.otandp.com/COVID-19-timeline>
- Zhang, H., Song, H., Wen, L., & Liu, C. (2021). Forecasting tourism recovery amid COVID-19. *Annals of Tourism Research*, 87(1), 103149. <https://doi.org/10.1016/j.annals.2021.103149>
- Yuki, K., Fujiogi, M., & Koutsogiannaki, S. (2020). COVID-19 pathophysiology: A review. *Clinical Immunology*, 215(1), 108427. <https://doi.org/10.1016/j.clim.2020.108427>
- World Health Organization. (2024). Coronavirus Disease (COVID-19) Pandemic. Wwww.who.int; World Health Organization. <https://www.who.int/europe/emergencies/situations/COVID-19>
- John Hopkins University. (2023, March 10). Johns Hopkins coronavirus resource center. Johns Hopkins Coronavirus Resource Center; Johns Hopkins University & Medicine. <https://coronavirus.jhu.edu/map.html>
- AlTakarli, Nourah S. (2020). China's Response to the COVID-19 Outbreak: A Model for Epidemic Preparedness and Management. *Dubai Medical Journal*, 3(2), 44–49. <https://doi.org/10.1159/000508448>
- Tourism Statistics | Hong Kong Tourism Board. (n.d.). Discover Hong Kong. <https://www.discoverhongkong.com/eng/hktb/newsroom/tourism-statistics.html>
- UNWTO World Tourism Barometer. (n.d.). <https://doi.org/10.18111/wtobarometereng>
- C. L. Tai, A., W. H. Wong, D., F. Lee, H., & W. Qiang. (2022). Tourism's long- and short-term influence on global cities' economic growth: The case of Hong Kong. *ProQuest*, e0275152. <https://doi.org/10.1371/journal.pone.0275152>
- The Government of the Hong Kong Special Administrative Region - Press Releases. (2022, August 24). Government increases quotas for entering Hong Kong through Shenzhen Bay Control Point under Come2hk Scheme from today. Info.gov.hk. <https://www.info.gov.hk/gia/general/202208/24/P2022082400522.htm?fontSize=1>
- ZHU, P., & TAN, X. (2022). Evaluating the effectiveness of Hong Kong's border restriction policy in reducing COVID-19 infections. *BMC Public Health*, 22(1). <https://doi.org/10.1186/s12889-022-13234-5>
- 入境事務處. (n.d.). Wwww.immd.gov.hk. <https://www.immd.gov.hk/>
- Statista. (2010). Global GDP 2010-2022 | Statista. Statista; Statista. <https://www.statista.com/statistics/268750/global-gross-domestic-product-gdp/>
- HKTDC Research. (n.d.). Research.hktdc.com. <https://research.hktdc.com/en/article/NjQzOTczODA5>



- Statista. (n.d.). Hong Kong: tourism industry GDP share. Statista.  
<https://www.statista.com/statistics/622227/hong-kong-tourism-industry-gdp-share/>
- Statista. (n.d.-a). Hong Kong: international tourism revenue 2019. Statista.  
<https://www.statista.com/statistics/918335/hong-kong-revenue-from-international-tourism/>
- Hong Kong Economy. (2019). Hong Kong Economy - Hong Kong Economic Situation - Latest Development. Hkeconomy.gov.hk.  
<https://www.hkeconomy.gov.hk/en/situation/development/index.htm>
- Wong, S.-C., Au, A. K.-W., Lo, J. Y.-C., Ho, P.-L., Hung, I. F.-N., To, K. K.-W., Yuen, K.-Y., & Cheng, V. C.-C. (2022). Evolution and Control of COVID-19 Epidemic in Hong Kong. *Viruses*, 14(11), 2519. <https://doi.org/10.3390/v14112519>
- Census and Statistics Department. (2024, February 28). Gross Domestic Product (GDP), implicit price deflator of GDP and per capita GDP. [Www.censtatd.gov.hk](http://www.censtatd.gov.hk).  
[https://www.censtatd.gov.hk/en/web\\_table.html?id=31](https://www.censtatd.gov.hk/en/web_table.html?id=31)

## **Consumer Purchase Intention When Buying Athletic Shoes By Vedant Kothari**

### **Abstract**

When buying a product, regardless of what category it falls under, individuals are usually faced with a plethora of decisions to make. Traditionally, research has been performed to determine side effects of this decision-making, including things such as overchoice or being overwhelmed. However, research focusing on what consumers really prioritize when making purchases is lacking. The present research delves into what factors make consumers buy a product, which in this case, is athletic shoes. This research aims to analyze the thought process of consumers when buying athletic shoes, specifically what prompts them to learn towards a particular product over another. It focuses on this topic to identify trends and preferences in the athletic shoe market, as well as what psychological factors in general people prioritize when buying athletic shoes. Data was collected through a questionnaire via Google Form from 47 participants classified into two groups: athletes and non-athletes residing in India and the United States. It was found that regardless of whether they were athletes or not, the participants rated the following categories (durability, comfort, odor, quality) in the same order of prioritization. Participants from both groups on average ranked comfort, quality, durability, and odor as the most important factors when purchasing athletic shoes, in descending order. It implies that both athletes and non-athletes have similar trends in terms of their thought processes when they buy athletic shoes.

**Keywords** Purchase Intention, Athletic Shoes, Psychology, Athletes, Non-Athletes

### **Introduction**

Over time, the factors which lead to consumers buying goods have evolved massively as products have become more complex. Factors such as price, size, and color have always been of interest, but more novel variables like odor, design, material, and brand loyalty are the focus of the market. For instance, Silva, Madhushani & Jayalath (2020) studied the impact of brand loyalty on consumers' purchase intention when buying basketball shoes [1]. It was found that over 50% of participants favored Nike shoes over the likes of Puma, Adidas, and Reebok. Furthermore, a study by Chu (2020) found that comfort and style were the two factors that had the most impact on influencing consumer purchase intention when buying athletic shoes [2].

Sports goods over time have been some of the most evolving types of products. As sports have advanced, having the top gear for sports has become an essential part of performance. The best of the best athletes make sure to have the best shoes, clothes and other important gear. For instance, Novak Djokovic, who signed a shoe deal with athletic shoe company Asics. Asics is one of the brands that produces high-end shoes for athletes [3]. Similarly, LeBron James and his signature shoe line with Nike is another example. LeBron has had a shoe deal with Nike for over 20 years, and over this time has released over 20 models [4]. However, there was more to the difference between each of these models than just looks. The comfort, structure, and fit of each

of these shoes were unique. In any sport, the most differentiating factor between athletes, whether in a team sport or individual sport, is the shoes.

Having the best shoes while playing in any sports is essential to ensuring the best performance in those sports. It is not in one's interest to have shoes that don't support the ankle, for example, and this can only lead to more injuries when playing the sport [5]. This leads to a variety of other factors which consumers tend to look for when purchasing athletic shoes, such as durability, comfort, and odor. In terms of durability, some consumers may like shoes which last a short amount of time, such as half a year, whereas other consumers may like a shoe that lasts them a long time, even more than a year in some cases. In addition, a big factor consumers look into when buying shoes is the odor that they produce, as no one wants a shoe which produces a bad odor after a very short time of usage [6]. Lastly, comfort is a big factor, as some consumers' definition of comfort correlates to a tight fit whereas other consumers' definition of comfort correlates to a loose fit. [7]

There is a big variance based on perspective as to how consumers prefer their athletic shoes to be, which is what this study aims to investigate. Athletes, such as track runners, may prefer shoes that have a tight fit and don't slip off their feet, whereas a non-athlete may prefer shoes that have a loose or perfect fit and give their foot more room to breathe. This study focuses on what factors consumers prioritize in relation to their age, location, and other demographic factors. On the other hand, when there is a case where there are too many factors to prioritize, it can lead to overchoice, which can be a cause of a phenomenon known as decision paralysis. According to Manolică, Gută, Roman, and Dragăn (2021) [8], providing more options at the time of purchase makes it more likely for a customer to have post-choice dissatisfaction. The influx of options actually made it harder for the customer to make a decision.

## **Methodology**

### **Aim of the Study**

The aim of this research paper is to dive deeper into the thought process and psychology of a consumer when purchasing athletic shoes to learn about what prompts them to buy a certain product over another. The design of this paper is a quantitative-based approach.

### **Sample**

The study consisted of 47 participants who were assured that the information provided by them would be kept confidential. Within these 47 participants, 39 identified as male, and 8 identified as female. Their ages ranged from 10 and on. They were separated into two groups: 24 athletes and 23 non-athletes. The inclusion criteria for the study was people who had some knowledge of the athletic world, and either played sports casually or more seriously on a team.

### **Ethical Considerations**

The survey was sent out to all participants via the same messaging application that was used to contact them. Their identity would not be revealed to any third party organization of any sorts, and there was an option to keep their identity anonymous within the form itself, if that was what they wished. In addition, they were allowed to eliminate their participation from the study at any time they wished. At the beginning of the study, participants were asked for their name, age, and gender. The research only started after receiving their consent through filling out the survey questionnaire.

### **Tools Used**

To gain a basic understanding of the structure of a study, Sagone, Commodari, Indiana, and La Rosa (2023) were used [9]. This study used a similar methodology where a survey was carried out and participants were asked to rank different factors using a scale similar to that of the Likert scale [10]. For example, the 10th question on the survey asked participants to rank durability, comfort, odor, quality, marketing, customer service, price, design, and brand on one of the following categories based on how much they prioritized them: least, low, medium, high, and most. The questionnaire consisted of 18 questions to gauge an understanding of what factors consumers prioritize when buying athletic shoes.

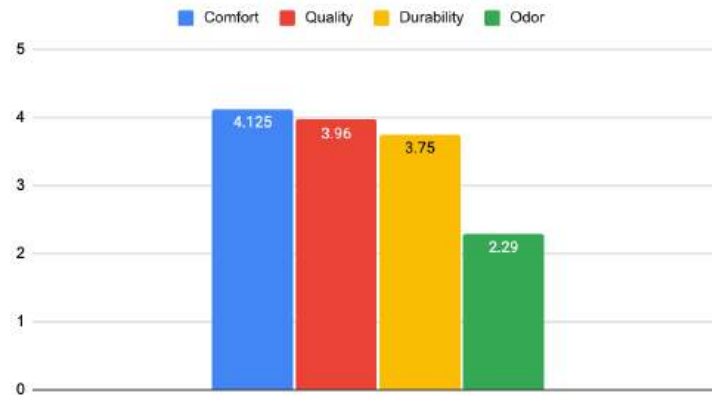
### **Data Collection Procedure**

The questionnaire via Google Forms was used to determine what different people prioritize when purchasing athletic shoes (durability, comfort, odor, quality, marketing, customer service, price, design, and brand). First, participants were informed about the purpose of the study, and were then asked whether they would like to be included in the research via questionnaire. Those who agreed to participate were sent the link to the questionnaire to fill out and return. It consisted of 18 questions from short responses to multiple choice to rankings. They were first asked basic demographic information. Then, they were asked questions to learn their preferences based on a multitude of factors when buying an athletic shoe, such as odor, durability, and comfort.

Finally, they were asked to rank each factor (durability, comfort, odor, quality, marketing, customer service, price, design, and brand) based on how important they were to them (least, low, medium, high, or most) and then justify why the variable they chose as the least important to them was the least important and why the variable they chose as the most important to them was the most important. Taking all these responses into account, it became possible to determine what factors different groups of people (male athletes, female athletes, male non-athletes, female non-athletes) prioritized when purchasing the same kind of product: athletic shoes. To extract data in an easier format, we converted all values to numbers. 1 replaced “least,” 2 replaced “low,” 3 replaced “medium,” 4 replaced “high,” and 5 replaced “most.” We then divided the participants into two groups, athletes and non-athletes, and averaged each group's rankings on the categories of comfort, quality, durability, and odor to see how they compared to each other.

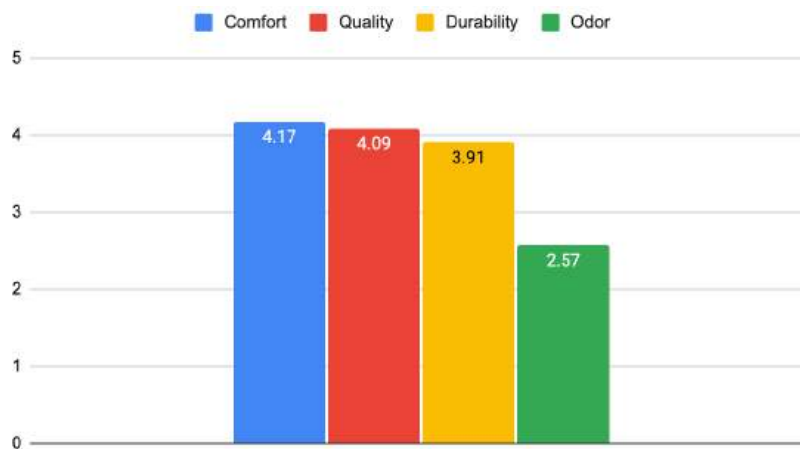
## Results and Discussion

After receiving all responses from the participants, descriptive statistics for the collected data was measured.



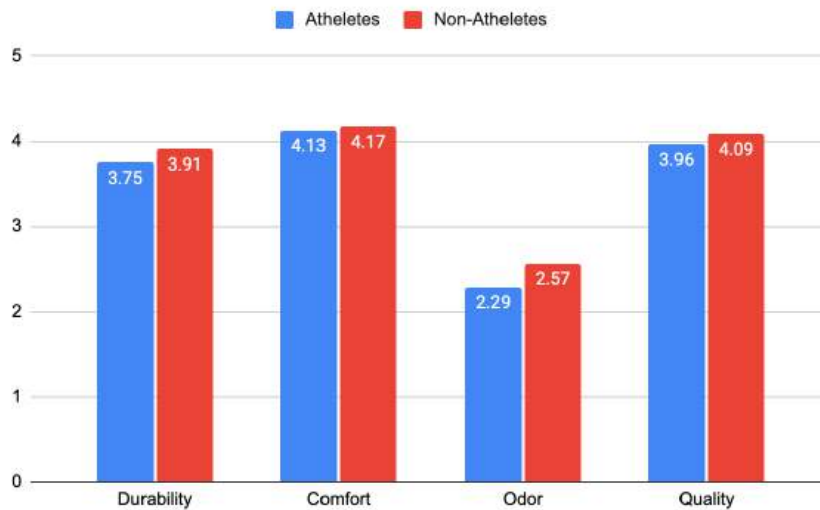
**Figure 1:** Graphical Representation Of Athletes’ Average Rankings of Factors When Buying Athletic Shoes (N=47)

In Figure 1, one can see how athletes who participated in the survey ranked the four factors (comfort, quality, durability, odor) in terms of importance. It can be seen that comfort was ranked the highest. With a rating of 4.125, this depicts that participants ranked comfort to be somewhere between “high” and “most” on the question from their form response. Next, with the second-highest ranking of 3.96, it can be observed that quality is somewhere between medium and high on the scale, but more closely towards high. Then, with the third-highest rating of 3.75, durability is observed to be somewhere between medium and high as well. Finally, with the lowest rating of 2.29 was odor, which was ranked somewhere between low and medium by athletes, on average.



**Figure 2:** Graphical Representation of Non-Athletes’ Average Rankings of Factors When Buying Athletic Shoes (N=47)

In Figure 2, one can see how non-athletes who participated in the survey ranked the four factors (comfort, quality, durability, odor) in terms of importance where comfort was ranked the highest. With a rating of 4.17, this depicts that participants believed comfort to be somewhere between high and most on the scale. Next, with the second-highest ranking of 4.09, it can be observed that quality is somewhere between high and most on the scale, but more closely towards high. Then, with the third-highest rating of 3.91, durability is observed to be somewhere between medium and high as well. Finally, with the lowest rating of 2.57 was odor, which was ranked somewhere between low and medium by athletes.



**Figure 3:** Graphical Representation of Athletes' Average Rankings For Each Category in Comparison With Non-Athletes' Average Rankings For Each Category (N=47)

Figure 3 depicts the contrast between the responses of athletes and non-athletes. In terms of raw rankings, the rankings remained the same for both groups in descending order: Comfort, Quality, Durability, Odor, which means that there is nothing too variable in terms of what athletes prioritize when purchasing athletic shoes versus what non-athletes prioritize when purchasing athletic shoes.

From the data collected, it can be seen that the comfort of the shoe is the most important variable to the participants. This could be for multiple reasons. For starters, wearing a comfortable shoe is crucial to preventing injuries with your foot, such as bunions and blisters, while providing support for your arch to reduce the effects of a flat foot. In the realm of athletes, specifically, preventing injury is crucial, and in order to do that, it's vital to have a shoe that fits our foot and is comfortable at the same time. *Effects of shoe inserts and heel heights on foot pressure, impact force, and perceived discomfort during walking* by Lee Yung-Hui observed that as participants wore heels with increased height, the impact force and perceived discomfort when walking increased, which corroborates the importance of comfort in a shoe. Hennig & Sterzing (2010) found that, amongst the soccer players they interviewed, the two factors on top of the

wish list for the players in a pair of cleats was comfort and shoe-fit [11]. It influenced their kicking speed, accuracy, and performance overall, which supports the significance of comfort in a shoe.

## **Conclusion**

As the sports world is growing increasingly, so is the significance of performance. The aim of this research paper was to assess what factors prompt a consumer to buy a certain product over another. It can be concluded that comfort was the most important factor. Out of the factors of durability, quality, comfort, and odor, both athletes and non-athletes ranked them in descending order in the following order: comfort, quality, durability, odor. This study can benefit those who struggle from overchoice, and may feel the need to buy multiple shoes in the event they can't choose one. As a result, this can lead to further implications, such as spikes in anxiety levels and a higher likelihood of purchases based on impulse in the future. Of course, there are limitations to this study. Things such as quality of thought and time taken to respond to the questionnaire were not considered. In addition, a majority of the participants are between the ages of 16-21, meaning that responses cannot be generalized to the broader sample of athletes and non-athletes. It's important to remember that the main priority is comfort, so the first thing one should look at in a shoe is how it fits on the foot.

## Works Cited

1. K.C. De Silva, AAL Madhushani, SDS Jayalath, “The Impact of Brand Loyalty on Customer Purchase Intention: An Empirical Study on Basketball Shoe Brands of Adidas, Nike and Puma in Asia,” ResearchGate.  
[https://www.researchgate.net/publication/345983988\\_The\\_Impact\\_of\\_Brand\\_Loyalty\\_on\\_Customer\\_Purchase\\_Intention\\_An\\_Empirical\\_Study\\_on\\_Basketball\\_Shoe\\_Brands\\_of\\_Adidas\\_Nike\\_and\\_Puma\\_in\\_Asia](https://www.researchgate.net/publication/345983988_The_Impact_of_Brand_Loyalty_on_Customer_Purchase_Intention_An_Empirical_Study_on_Basketball_Shoe_Brands_of_Adidas_Nike_and_Puma_in_Asia)
2. J. Chu, “Factors affecting the purchase decision of sneakers in post-90s consumer groups,” in *Advances in intelligent systems and computing*, 2020, pp. 480–487. doi: 10.1007/978-3-030-51828-8\_63.
3. L. Christou, “The lucrative sponsorship deals that made up 94% of Novak Djokovic’s earnings last year,” *Verdict*, Jun. 19, 2024. [Online]. Available: <https://www.verdict.co.uk/novak-djokovic-net-worth-sponsorship/#:~:text=Likewise%2C%20a%20deal%20with%20Asics,lucrative%20deal%20is%20with%20Lacoste.>
4. P. Benson, “LeBron James’ latest Nike sneaker is out now,” *Sports Illustrated FanNation Kicks News, Analysis and More*, Jan. 05, 2024. [Online]. Available: <https://www.si.com/fannation/sneakers/news/lebron-james-latest-nike-sneaker-is-out-now#:~:text=James'%20signature%20sneaker%20line%20has,officially%20hits%20shelves%20this%20week.>
5. “The best shoes for ankle support, recommended by experts — from sneakers to boots and more,” *TODAY.com*, Sep. 13, 2024.  
<https://www.today.com/shop/best-shoes-ankle-support-rcna170495>
6. K. Hui and K. Hui, “6 Best Odor-Fighting Shoes for stinky Feet,” *Well+Good*, Jun. 22, 2023.  
<https://www.wellandgood.com/best-odorless-shoes/>
7. U. S. S. Academy, “Factors that Influence African-American Millennials to Purchase Athletic Shoes,” *The Sport Journal*, Nov. 25, 2013.  
<https://thesportjournal.org/article/factors-that-influence-african-american-millennials-to-purchase-athletic-shoes/>
8. A. Manolică, A.-S. Guță, T. Roman, and L. M. Dragăn, “Is consumer overchoice a reason for decision paralysis?,” *Sustainability*, vol. 13, no. 11, p. 5920, May 2021, doi: 10.3390/su13115920.
9. E. Sagone, E. Commodari, M. L. Indiana, and V. L. La Rosa, “Exploring the Association between Attachment Style, Psychological Well-Being, and Relationship Status in Young Adults and Adults—A Cross-Sectional Study,” *European Journal of Investigation in Health Psychology and Education*, vol. 13, no. 3, pp. 525–539, Feb. 2023, doi: 10.3390/ejihpe13030040.
10. G. M. Sullivan and A. R. Artino, “Analyzing and Interpreting Data from Likert-Type Scales,” *Journal of Graduate Medical Education*, vol. 5, no. 4, pp. 541–542, Dec. 2013, doi: <https://doi.org/10.4300/JGME-5-4-18>.
11. E. M. Hennig and T. Sterzing, “The influence of soccer shoe design on playing performance: a series of biomechanical studies,” *Footwear Science*, vol. 2, no. 1, pp. 3–11, Mar. 2010, doi: <https://doi.org/10.1080/19424281003691999>.



**The Decline of Shui Script and Its Cultural Implications:  
A Study on the Changing Attitudes Towards Cultural Heritage Among the Shui Tribe  
By Rongyu Zhou**

**Abstract**

This study examines the reasons why the Shui script, a cultural heritage of the Shui tribe, is no longer as authoritative as it once was. Drawing on interviews with 17 Shui people ranging from 5 years old to 80 years old and observational data, the paper explores the characteristics of the development of Shui script culture and its changes in the modernization process from analyzing how different generations view Shui script, as well as the relationship between the inheritance of Shui script and Shui people's cultural identity. The study finds that people stop paying attention to and losing faith in Shui's character as a result of modernity and the impact of the Han culture. The essay presents the changes of Shui culture by discussing the understanding of Shui traditional morality and gender cognition by different generations.

**1. Introduction**

The cultural heritages of many ethnic minorities are facing the threat of extinction today as fewer individuals adopt their own historical cultural identity. The influences of culture shock, industrialization, and globalization have led people to gravitate towards more widely-accepted cultures. Among these communities, the Shui tribe has not received as much attention from scholars in comparison to the Miao and Mongolian ethnic groups.

The Shui ethnic group is a relatively small group of people among the 55 ethnic minorities. This ethnic group has a unique cultural heritage, which is the Shui script. Shui script contains Shui ancestors' experiences of life and primitive religious beliefs, and it is a tool that can be used by sorcerers to perform rituals. Shui people's life are helped by the Shui script. According to Shui script masters, Shui script is considered a sacred and symbolic cultural representation of the Shui tribe; yet, individuals in Shui, particularly young generations, are unfamiliar with it and some may even have never heard of Shui script.

The article investigates the decline of the cultural significance of the Shui script by focusing on the changing cultural identity of Shui people. This research will aid people to understand the impact of modernity on cultures such as the Shui by presenting them the huge difference of perspective between the elderly and young people in the Shui tribe. The research will also raise awareness of the rapidly disappearing civilizations and cultures. The study will employ cultural sociology and cultural anthropology approaches and collect ethnographic data.

**2. Literature review**

This section covers the literature on subjects such as culture shock and cultural identity and primitive religion. There are few studies on Shui ethnicity and Shui script. As a result, this study will examine some of the works that describe other ethnic minorities.

## **2.1 Cultural revival and cultural identity: the change and modern transformation of Miao culture in southeast Guizhou**

In this paper, the author argues that the cultural identity of the Miao people has altered as a result of modernism and globalization. Identity and belief shifted because of both their country and the market.

In China, Chinese governments are eager to accomplish modernization, and most actions are geared toward that end. They promoted huge economic growth in China. The rural areas where the Miao lived believed that by participating in the market economy and modernity, they may boost their own economic growth and wealth level.

This cooperative relationship between the government and Miao, however, is unequal and harmful to Miao culture.

Miao culture has been intentionally or unintentionally displaced by Han culture and modernity as a result of the strength and influence of Chinese governments. As a result of the interaction between the government and ethnic minorities, the Miao are obliged to evolve and introduce modern ways of living, such as administering their tribe, such as the construction of schools, banks, and police stations.

They are moving toward Han culture. Chinese governments, whether intentionally or unintentionally, employ authority to force ethnic minorities to behave similarly to the Han. That is the first phase of the Miao's society shift, in which most of their conventional social institutions were altered.

During the second phase from 1970-1980, the Miao experienced a cultural revival, with senior members of the tribe electing to head the restoration and assist the Miao in recalling all of their recollections of traditional clothes and awakening their cultural identity. During this time, practically all of the customs were restored, and the Miao used their vision to learn about themselves by enriching their traditional rituals.

The third phase, which was renovated in the 1990s, was motivated by both the government and the Miao's needs for tourism. The Miao actively promote and restore their traditions to create a better cultural background and impression to attract more visitors. The Chinese governments also contribute to the promotion of Miao and ethnic minority culture by legitimizing the presence of cultural customs to promote economic growth. As a result of "the other," (which refers to tourists) the Miao learned more about themselves and verified their identity.

Because the Miao dwell in the same region as the Shui, it is reasonable to believe that Shui culture has been influenced by globalization and modernization. The only difference is that cultural resurgence did not occur in Shui ethnicity as the author described in the second and third phases because the Shui tribe is not as well-known as Miao due to the small population. As a result, few people traveled to the Shui sites and did not bring them economic benefits and aid the Shui's cultural identity rebirth.

## **2.2 From "Hmong", "American Hmong" to "Hmong American" -- the change of Miao cultural identity in the United States**

The author argues that there are differences in ethnic identity and cultural identity between different generations.

The first generation, Hmong 1, is a generation of people born between the 1950s (or earlier) and the mid-1970s in their native Asian countries such as Laos and Vietnam, which corresponds to the English term "Hmong". They are all identified with traditional Hmong culture, they can fluently communicate in the Hmong language, and about 50% of them can read and write RPA Hmong.

As for Miao Generation 1.5, it refers to those born between the late 1960s and mid-1970s in their native Asian countries. The Miao 1.5 generation has typical characteristics of dual cultural identity, and recognize the traditional culture of Miao.

Miao Generation 2 is the young generation who were born in the United States and received American systematic school education. They emphasize their identity as American citizens and identify more with American culture.

Cultural identity declined among those three generations, according to the paper, showing how moving away from hometown would lead to people's identity shift. This pattern could also be applied to Shui nationality. In the study, there are three generations of Shui who have been affected to varying degrees by Han cultural shock after moving into Han civilizations.

### **2.3 The elementary forms of religious life- Emile Durkheim**

The book states that the most primitive religion can be found in the simplest organized societies. Shui ethnicity, a relatively underdeveloped society in China, is indeed a region that has a traditional form of religion-belief of Shui script divination and the ghost and gods. The most savage and outlandish rituals, as well as the most bizarre myths, carry some human need and some aspect of individual or social life. In Shui script, the myth of a butterfly helping the birth of the first Shui reflects their need toward nature. The reasons that believers use to justify these rituals and myths may (or usually do) be wrong; But real reasons do not exist, and it is science's job to discover them.

Durkheim believed that the reason why primitive people believe that elves are widespread is because they are different from modern people, they can not distinguish between living and inanimate matter. The elderly in Shui who were not exposed to scientific education, believed in elves and ghosts due to the same reason.

Durkheim believed that the system of nature worship is itself an "illusion system", which provides a realistic starting point for religion, confirming the feelings that natural phenomena inspire in us. His theories explain Shui people's belief of nature. According to him, the basic elements of religion are beliefs, ritual is only the external transmission, and it is the function of ritual to induce man to act and to stimulate consciousness and emotion. And the main purpose of religion is to exert influence on spiritual life. The belief of Shui script, a kind of religion, deeply represents the inner spirit of Shui people and their culture.

### **3. Research Methodology: Data collection and analysis of data**

Research was conducted by ethnographic observation and in-depth interviews to reveal how people in the Shui tribe interpret the Shui script and their attitudes. Qualitative research methods were implemented throughout the research.

### **3.1 Sampling**

This study was conducted in China, Guizhou, Sandu, where 63% of the Shui live.

#### **3.1.1 The geographic environment**

Sandu Shui Autonomous County, located in the southeast of China, is located in the hinterland of Moon Mountain and Leigong Mountain, between 107°40 'to 108°14' east longitude and 25°30 'to 26°10' north latitude. Due to the complex influence of topography, terrain, landform, and altitude, regional differences in the climate of Sandu are formed.

#### **3.1.2 The cultural environment**

People reside in wooden houses and the majority of the houses have two or three stories. People live on the second or third floors because of floods. Even though due to global warming, floods have decreased, the living pattern is still maintained.

The Shui ethnicity, as a whole, lives in a small and compact community.



*A photo of Sandu Layout Village*



*A photo Shui Autonomous County of Sandu*

Purposive sampling will be used to find interviewees for the study. The criteria for selecting in-depth issues are:

1. People of Shui ethnicity
2. People with extensive knowledge of the Shui script
3. Three generations of the Shui ( 1-18, 18-40, 40-80 )

Pedestrians of varied backgrounds and different stakeholders in Shui ethnicity will also be interviewed. The study implements triangulation, a strategy used to improve the credibility and validity of findings by cross-verifying information from several sources, data collection methods, or researchers' views to get a well-rounded grasp of the topic.

The following is the background and the key concepts of the research.

After the fall of the Shang Dynasty, the ancestors of the Shui nationality began to migrate south, which continues to today. The Shui people have many special customs and rituals. Shui script is a unique writing system and is the focus of this paper. The Shui script records a lot of

cultural information about astronomy, geography, religion, folklore, ethics, and philosophy. Shui ethnicity believes in the existence of ghosts and spirits, and the Shui script integrates many differences of ghosts and spirits, taboos, affecting all aspects of the lives of Shui people.

### 3.2 Data collection and analysis

In this study, follow-up questions were added based on the responses of the respondents. In-depth interviews were captured on both audio and video. Observation of the atmosphere and interviewees are included.

Main Interviewees:

*1. Photo of Mr. yang*



*2. Photo of Mr. wei*



3. Photo of Mr. wei (another one)



4. Photo of Mr. Lu



### 5. Photo of Mr.shi



Table 1. Information of Interviewees

Number	Gender	Schooling	Age	intro
1 Mr. Yang	Male	Vocational school	68	Shui script master provincial heir of Shui script Shui script researcher
2 Mr. Wei	Male	middle school	68	Shui script master
3 Mr. Wei	Male	Vocational school	59	Shui script master provincial heir of Shui script decipherer of the Shui calendar Shui script researcher
4 Mr. Lu	Male	middle school	57	Shui script master
5 Mr. Shi	Male	Middle school	46	Head of Shui cultural heritage Experience Center

### 3.3 Data collection method

The responses would be sorted into two categories: their grasp of the substance of Shui script and the significance of Shui script. The culture of the Shui script itself, the reason why the



Shui script began to be less authoritative, and the moralities disclosed by the Shui script became the three key parts of the investigation.

## 4. Findings

### 4.1 An overview of the culture of Shui script

#### 4.1.1 Shui script is a highly abstract reality of Shui culture, reflecting many cultural cores of Shui ethnicity.

There are many kinds of Shui scripts, which are divided into reading books, palace palm books, time image books, orientation books, star books, etc., and are divided into application volumes and reading volumes according to the content. "Shui script" is divided into "white book" (good) and "black book" (bad) because of the different occasions in which it is used.

##### 4.1.1 classification and types of Shui script

There are application volumes and divination volumes.

#### 4.1.2 The role of Shui script in divination

1. Calculating Lucky and Unlucky Days (Mr. Shi)

2. Manipulating Life Events and Destiny

Shui scripts will answer the questions of when to build new homes, when to marry, when to sacrifice, and so on. It can also be used to predict one's fate.

3. Interacting with Ghosts and Gods

Shui script can also help Shui script masters beseech to gods and ghosts for help, communicate with them during festivals and sacrifices, and change one's destinies.



*A photo of Mr. Wei communicating with gods and ghosts*

#### **4.1.3 The standard of choosing Shui script successors**

The standard of choosing Shui script successors is they must be respectful to their Shui ancestors and must be altruistic, be in a good temper and be smart and extremely interested in Shui script.

#### **4.1.4 Shui script and Jiagu script**

There are many similarities between Shui script and oracle bone script, as they are all hieroglyphic and based on pictures.

The Shui script, according to Mr. Moyouzhi, appeared before the oracle bone script.

#### **4.1.5. The basis of Shui script divination**

Mr. Ludegong, the developer of Shui script, based the Shui script on the "jixing script" Shui calendar. Shui masters would factor in people's lunar calendar birth dates when calculating Shui calendar birthdays. The 28 stars in the sky were used to form the Shui calendar, which is the basis of Shui script divination. Mr. Wei spent ten years interpreting the Shui calendar. According to him, when the 28 stars are face to face, it indicates fortune. It signifies ill luck when they are back to the right with one other.

#### **4.1.6 The inheritance manner of Shui script**

Shui script are copied by hand and conveyed between teachers and pupils, so there may be some differences when passing the knowledge. As a result, several meanings and versions of the same word exist in the Shui script. (Mr.shi).

Historically, Shui writing had to be passed down within the same family. However, even though the Shui script was taught in one family, a study discovered that most Shui scripts in different households follow the same fundamentals (Mr. Yang).

Furthermore, the Shui script was only allowed to be passed down to males rather than females because it was thought that females should do chores in the past.

#### **4.1.7 The rules in divination**

Mr. Wei claims that during divination, various laws are applied when solving different situations such as the direction and place of the divination.

#### **4.1.8 The limitations of Shui script divination**

Shui script cannot detect the best directions and locations to accomplish particular tasks. Furthermore, mistakes may occur and the condition might be worse after divination. (Mr.shi)

#### **4.2 The moral underpinnings embedded in Shui script**

The study discovered that, from the aspect of change of gender cognition, admiration toward nature and moral beliefs between younger generations and the elderly, the following shows the change of Shui script belief and Shui culture.

#### **4.2.1 the worship of Altruism**

To begin with, the Shui script demonstrates that Shui values people's selflessness. One of the most significant criteria for selecting successors to the old Shui script masters is benevolence. Shui script successors must be kind and selfless to prevent destructive or evil events with the help of Shui script. According to Mr. Yang, even when he must go through huge mountains and several towns to divine, he would travel if anyone asked. Also, he would not ask for money depending on the condition of other families. After all, the personality of altruism was highly praised in the Shui script. Besides, many Shui script masters are altruistic because the creator of the Shui script ludegong was altruistic. He once helps people of Shui ethnicity find good places to live while living in a cave of swallows himself. And that is the most important reason why Shui script masters are respected by Shui ethnicity.

Unfortunately, the Shui are no longer as charitable as they once were. They would return to their hometown to assist when members of their tribe were married or bereaved, but that was forced and enhanced by their parent's command and the rules of the village. Furthermore, members of the Shui tribe frequently assisted their village neighbors only due to the relationship their parents have with other families. Several young people reported a reduction in generosity after moving to cities. Because of the relatively isolated environment outside, they began to act more like city residents, managing their affairs rather than aiding and cooperating. They are also hesitant to assist because of the unfavorable news in the cities. Which shows the impact of modernity on people's moral beliefs.

#### **4.2.2 Belief in oneness between humans and the nature**

Second, Shui script shows the Shui tribe's desire to live in harmony with nature and their respect for nature. According to Mr. Yang and Mr. Wei, many ways of resolving disasters depend on trees and water, and things in nature. For example, one of the ways recorded in the Shui script is to worship huge trees. In the past, people only cut down a small number of old trees to protect the forest. They express their deep gratitude to nature because they grew up in villages surrounded by trees and water. In fact, the belief in the gods of nature is the way traditional religion began. (Durkheim, 1992). In his Introduction to the Science of Religion, Max Muller, founder of modern religious teaching, put forward the view that religion originated from natural worship. The biggest feature of nature worship is to ascribe certain divinity to natural objects. Nature worship is a very common form of belief in primitive society, and it is the original belief of primitive humans that natural objects and natural forces have life, will and great ability.

Unfortunately, as the population grows, this spirit declines. To survive, people must cut down more trees and use more water. Furthermore, two young individuals stated that they do not rely heavily on nature because they frequently travel to cities for school, and they have little experience defending the environment or worshipping nature gods during festivals. The ritual of worshipping tree gods is not as popular as it once was.

#### **4.2.3 the changing gender values**

Finally, the Shui script tells that the Shui tribe prefers males over females in the past. Shui successors had to be males because males married others and were thought to impart knowledge to other families. Furthermore, people in the Shui ethnicity felt that females should perform chores. Mr. Yang once found a female Shui script successor and taught her divination and other skills. And the older Shui script masters were against him and believed that he disobeyed old rules, but the younger ones support Mr. Yang as they believe the action could help inherit Shui script.

This worldview evolved throughout time as a result of the influence of Han Chinese governments, which called and promoted gender equality. Nowadays, everyone can learn the Shui script even if they are not of Shui ethnicity. (Mr. Shi ). Shui females are allowed to go to school and eat with males on a table, which is nearly impossible. Their consciousness greatly changed.



*A photo of gaozhai, longquan village*

*Translation: Boys and girls are equally strong and can grow up to support their nation.*

#### **4.3 The fading authority of Shui script**

##### **4.3.1 Decline in belief among youth**

The survey discovered that Shui scripts are no longer as authoritative among people of all ages, particularly younger people. Older people and Shui script masters still believe in Shui script. According to Mr. Yang, he thinks that the Shui script is reliable and Mr. Wei declared that

the Shui script is scientific. But it's different for the younger generations. Some individuals had never heard of the Shui script ( a shopkeeper in Sandu, 23 years old) .

And while some individuals have heard of the Shui script, they may not be aware of its potential for divination (a girl in Layou village) . Even though some people are aware of Shui characters, they do not believe that Shui scripts can predict the future because it's not scientific and he doesn't care about it because it can't bring him wealth. (A person who sells barbeque)

#### **4.3.2 shifting cultural identity and customs**

The survey also revealed that younger generations such as those aged 6-18 18-40 haven't preserved much of the culture and customs of Shui, instead, they live more like Han due to the fact that the Han educational system and working system is the preferred way of living in China. Some of them have already forgotten all of their customs and costumes, such as the Duan festival, one of the biggest festivals in Shui culture, and instead celebrate the typical Chinese New Year just like Han. As for clothing, according to Mr. Shi: "In the past, young women wore our traditional customs, but now, they wear skirts and modern clothes.

There is a special group of people who knows quite much about the Shui script and show full respect for their customs and the divination of the Shui script even if they cast doubt on the reliability of the Shui script. Those people are often one of the most educated people in the tribe. Due to the fact of education in sciences, those people do not see any evidence of the Shui script, therefore, often won't ask Shui script masters to divine their future or use the Shui script to tackle illness. A 20-30years old man, he mentioned that his only experience of using the Shui script was when he was very young, living with his grandparents. He said that only people aged like his grandparents would believe in divination. Younger ones will go to hospitals and take medicines.

#### **4.3.3 the natural and man-made factors affecting Shui script preservation**

The Shui script is in grave danger of dying out owing to both human and natural factors.

Non-human catastrophes include things like fire accidents that destroyed recordings of Shui writing and made it difficult to pass them on to future generations.

In terms of human factors, military conflicts in China lead to Shui people leaving their hometowns to fight. That group of Shui consequently loses connections with their tribe, and had no chance to learn the Shui script. (an old soldier who fought in the War to Resist US Aggression and Aid Korea shared). He left his hometown around 10, so he doesn't have the opportunity to learn his own culture. Second, because of advances in technology and urban growth, members of the Shui tribe who live in the mountains desire to leave their homes and seek out higher-paying jobs and chances outside of the mountains to increase their income. In that way, they are farther and farther away from their own culture and cultural heritage such as the Shui script.

One of the most important effects that hurt the inheritance of the Shui script is the Han cultural shock. Mr. Shi once said that "people in Shui are starting to be more and more affected by Han culture." To begin with, younger generations would learn Han characters because Shui characters will not be useful. When they are young, due to the influence and advancement of Han

education, most Shui dreams are to score higher on the gaokao( Chinese university entrance exam), which is written in Han characters, and get into better universities in China, which are all Han Universities. To succeed, they must study Han characters. According to Mr. Shi, most Shui script masters and successors nowadays are 40 and above. That is because younger ones began to study Han culture and prepare for gaokao. They have no time to learn the Shui script in their childhood. And when they grow up to work, they will have to learn how to communicate with people in the city.

#### **4.3.4 diminishing audience and trust of Shui script**

Additionally, the advancement of medical knowledge aids in the discovery of new, accurate, and science-based treatments for illness. So the audience group for the Shui script was lost. According to Mr. Shi, though he studies the Shui script a lot, his daughter does not believe in the Shui script and learns the Shui script because she thinks it's not scientific. And without that audience, since fewer individuals seek their advice when utilizing the Shui script(According to Mr. Lu, only 3-4 people come and ask him to divine every week, the number is extremely low compared to the time when his grandfather was divining), so fewer people learn Shui script. According to Mr. Shi, it's like water and fish. Water represents people who seek Shui script divination, fishes represent Shui script masters. And when the area of water is smaller, the fishes start to die out.

Furthermore, the advancement of science education encourages people to trust science rather than the superstitious Shui script. Less individuals are interested in learning the Shui script if there is no longer a belief in it.

#### **4.3.5 obstacles to learning Shui script**

-Lack of financial incentive;

Younger generations did not see many benefits to learning the Shui script. They couldn't make money by learning Shui script because Shui master is not a job, but rather a position that hands out help. According to locals, Shui script masters would not set specific prices for a single divination or so. As a result, studying the Shui script will not help people become wealthy.

-lengthy and tedious process;

It also takes too long to become proficient in Shui script because one needs to understand geography, algebra, Shui characters, and a variety of other subjects to use Shui script for divination.

-Erosion of the teaching ability of Shui masters

Aside from these external factors, some Shui script masters have refused to teach their successors all of the skills of Shui script for fear that their successors will use those skills to do bad things that will harm society or take over their employment. After numerous generations, less and less knowledge was passed down. According to Mr. Wei, Shui script masters used to be able to control natural forces, but now they can only divine some obvious good fortune and treat minor maladies. Based on the above reasons, the Shui script faces the challenge of disappearing.

#### **4.3.6 The existing actions of protecting the Shui script**

The lucky thing is that Chinese governments and organizations are working hard to ensure the inheritance of the Shui script. With the help of governments, Shui scripts were taught in schools, printed in little books that could be handed out, and held Shui script activities. But most of these are about learning how to read Shui script instead of using Shui script to divine. And even if Shui script divination is taught to younger generations, most of them are afraid to use it to divine by themselves, they rely on older Shui script masters to help. That is because mastering divination requires various knowledge such as math, astronomy and so on. Luckily, the Shui script is already a cultural heritage in Asia, and it is going to apply to worldwide cultural heritage for the next step.

### **5. Conclusions and discussions**

Overall, the Shui script is no longer as authentic as it once was due to the transformation in the Shui's cultural identity. Modernization, globalization, and the Han cultural shock have all contributed to a transformation in cultural identity, reflected by their change in their traditional moral beliefs. These moralities include altruism, the desire to keep humans and nature in harmony, and a preference for males.

In fact, many ethnic cultures are on the verge of extinction as traditional beliefs among generations decline. Furthermore, cultural preservation should be prioritized across distinct ethnic cultures. Because experiencing modernization can offer growth and economic benefits to certain areas, it is best to figure out how to make modernization and ethnic cultures coexist. Some possible solutions include the government continuing to provide advanced technical skills to rural areas while also providing employment positions that study cultural inheritance such as Shui script for living and salaries. Schools in those communities should respect their own cultural identity and teach their own characters, such as Shui characters, while still educating popular culture figures. Individuals should engage in cultural exchange with others, learning more about their own culture and the current world while engaging in cultural communication.

In conclusion, the study employed the Shui script as a case study to investigate and reveal the decline of Shui culture. Further research could dive deeper into the calculations of Shui script divination using fields of study such as astronomy, arithmetic, and so on. To obtain a more comprehensive analysis, they might potentially go deeper into a larger number of samples and interview individuals about their backgrounds and attitudes toward the Shui script.

## Works Cited

- Wen, Yi, Lin, Bosheng, and Meng, Yaoyuan. "Interpretation of Shui Script: Yin-Yang and Five Elements Volume." 凯里学院学报, vol. 30, no. 4, 2012, pp. 46-49.  
DOI:10.3969/j.issn.1673-9329.2012.04.15.
- Bai, Xiaoli. "The Character Variants of the Earthly Branches in Shui Script: The Correct Seven Volumes." 贵州民族学院学报(哲学社会科学版), no. 5, 2010, pp. 14-17.  
DOI:10.3969/j.issn.1003-6644.2010.05.004.
- "A Demonstration of the Advantages of Qualitative Methodologies in Stuttering Research." ScienceDirect, 7 May 2001, [https://doi.org/10.1016/S0094-730X\(01\)00094-8](https://doi.org/10.1016/S0094-730X(01)00094-8).
- Flowerdew, R., and D. M. Martin. *Methods in Human Geography: A Guide for Students Doing a Research Project*. Prentice Hall, 2008.
- Tang, M., Xie, S., and Liu, X. "Ancient Character Recognition: A Novel Image Dataset of Shui Manuscript Characters and Classification Model." *Chinese Journal of Electronics*, vol. 32, no. 1, Jan. 2023, pp. 64–75. <https://doi.org/10.23919/cje.2022.00.077>.
- "Our Story - The Shui Nationality." CCTV English, [Video].  
YouTube, <https://www.youtube.com/watch?v=kKysSF0jwD0>. Accessed 15 Dec. 2018.
- "Harmonious Life of Shui and Miao Ethnic Groups in South China's Ancient Village." [Video].  
YouTube, <https://www.youtube.com/watch?v=uEdrCz2mAlE>. Accessed 6 Sept. 2021.
- "The Mysterious Shui Ethnicity." [Video]. Bilibili, <https://b23.tv/xSqWv8t>. Accessed 26 June 2019.
- "Shui Ethnic People's New Year: The Longest Festival in China." [Video].  
YouTube, <https://www.youtube.com/watch?v=uHYqG8fO1XI>. Accessed 30 Apr. 2018.
- "Shui Ethnic Group in China Marks Duan Festival." [Video].  
YouTube, <https://www.youtube.com/watch?v=k3IlbC0Od-k>. Accessed 8 Nov. 2020.
- "From 'Miao' and 'American Miao' to 'Miao Americans' — Changes in Cultural Identity of the American Miao Community." *世界民族*, no. 1, 2017, pp. 83–93. <https://www.nssd.cn/html/1/156/159/index.html?lngId=671406730>.
- Wang. "Cultural Revival and Cultural Identity — Changes and Modern Transformation of Miao Culture in Qiandongnan." *贵州工业大学学报(社会科学版)*, vol. 7, no. 1, Feb. 2005, [https://xueshu.baidu.com/usercenter/paper/show?paperid=1r5y04q00g4b04e06k5d0pn0hd118380&site=xueshu\\_se](https://xueshu.baidu.com/usercenter/paper/show?paperid=1r5y04q00g4b04e06k5d0pn0hd118380&site=xueshu_se).
- Durkheim, Emile. *The Elementary Forms of Religious Life*. Oxford Paperbacks, 2008, [https://monoskop.org/images/a/a2/Durkheim\\_Emile\\_The\\_Elementary\\_Forms\\_of\\_Religious\\_life\\_1995.pdf](https://monoskop.org/images/a/a2/Durkheim_Emile_The_Elementary_Forms_of_Religious_life_1995.pdf).



# **The Roles and Mechanisms of Burst Firing Neurons in the Lateral Habenula Circuit in Depression, and Implications of Targeted Inhibition By Antidepressants as Potential Treatment By Emily Chen**

## **1. Introduction**

### ***1.1 What is depression?***

Depressive disorder, or depression, is a debilitating mental disorder with clinical manifestations of anhedonia and anxiety (Krishnan and Nestler), and is often accompanied by other neurological disorders, increased risk of suicidal behavior, and psychiatric and psychosocial morbidity (Birmaher). As the leading cause of disability worldwide (WHO), the alarming growth in the prevalence of depression – reaching over 1 in 6 adults in the U.S. diagnosed with depression in 2023 (McPhillips) – calls for a better understanding of its pathology.

### ***1.2 Related studies and systems involved***

Previous literature has been consistent in attributing a role in the neurological manifestation of clinical depression to the pyramidal neurons of the motor cortex microcircuit (Zhang et al.) (Younts, Chevalyere and Castillo). However, recent breakthroughs have suggested clinical depression may also be associated with the neurons of the Lateral Habenula (LHb) Circuitry as well.

Synaptic and electrophysical abnormalities of the cerebral cortical pyramidal cells (L5PCs) have been posited to exhibit a causal relationship with neurological disorders such as depression (Zhang et al.). L5PCs are principal excitatory neurons that transform synaptic input into a patterned output of action potentials that stimulate movement (Bekkers), by sending axons subcortically or contralaterally. In mice, depression is characterised by poorer mobility and decreased activity in behavioral tests (Gonzalez and Aston-Jones), and so in a study conducted by Zhang et al., the Layer V (L5) of the motor cortex microcircuit, which contains the highest number of cerebral cortical pyramidal cells (PCs), was investigated as it could act as a convenient region for neural microcircuit investigation. Zhang et al. studied the properties of the L5PCs by multi-neuron whole-cell patch clamping using light deficiency as a trigger to simulate depression-like symptoms in mouse models (Zhang et al.).

More recently, excessive activity in the LHb, specifically increased burst firing of the LHb neurons, has also been correlated with depressive-like symptoms in human and mouse models (Liu et al.) (Huang et al.). This pattern has been suggested to be associated with a suppression of the dopaminergic, serotonergic, or glutamatergic neuron activity in brain regions that LHb<sup>DRN</sup>, LHb<sup>VTA</sup> and LHb<sup>MnR</sup> project to – dorsal raphe nucleus (DRN), ventral tegmental area (VTA), and median raphe nucleus (MnR), thereby mediating the symptoms of depression.

### ***1.3 Proposed Drugs for Treatment***

Selective serotonin reuptake inhibitors (SSRIs) are the first choice when it comes to depression treatment, and are the most commonly prescribed antidepressants (Vroomen Durning). Fluoxetine (sold under the brand name Prozac), the most popular SSRI (Gilmore), acts by preventing the reuptake of serotonin in presynaptic serotonin neurons, thus increasing serotonin levels in the brain.

As observed in fluoxetine, the most widely used form of treatment operates by increasing dopamine or serotonin levels in the brain, to combat the lower levels of said neurotransmitters exhibited in depression. However, few drugs have been researched to target the neurons, such as those in the LHb, that regulate the release of these neurotransmitters. One such drug is ketamine, which is often recommended as an alternative medication for treatment-resistant depression. Ketamine elicits a much more rapid antidepressant effect than SSRIs, often showing effects as soon as half an hour after a single administration (Yang et al.). However, the long-term effects of ketamine exposure have yet to be researched and remain unclear.

#### ***1.4 Bursting mechanisms in general***

Bursting neurons are found in a vast array of networks and play many roles in coordinating rhythmic movements such as swimming in *Clione* (Pirtle). Besides controlling locomotion, a series of bursting neurons also work together in other systems such as the cardiac neuromuscular system in lobsters (Williams et al.), or the stomatogastric system in crustaceans (Selverston and Moulins). During bursting, a neuron, or a group of neurons, repeatedly fires a group of closely separated spikes (Fox, Rotstein and Nadim), and their functions, as mentioned above, vary greatly – to create rhythmic movement, to further regulate other neurological processes, etc.

#### ***1.5 Bursting Mechanisms in the Lateral Habenula***

The lateral habenula (LHb), along with the medial habenula, form the mammalian habenula, which regulates serotonergic and dopaminergic activities. The LHb hosts mainly glutamatergic neurons that project to regions involved in the release of dopamine (the ventral tegmental area (VTA)), and serotonin (the median and dorsal raphe nucleus) (Hikosaka) in the midbrain and the hindbrain. Recent studies have revealed that LHb neurons acted oppositely to dopaminergic neurons in monkey models (Matsumoto and Hikosaka), implicating the possibility of LHb being a negative reward signal. It is now hypothesized that the LHb is the brain's "anti-reward centre", responsible for encoding negative rewards that are exhibited in excess in depression.

It has been widely established that an increase in burst-firing neurons in the LHb is associated with depressive symptoms in various models (Bray). The dopaminergic neurons in the LHb are intrinsically active and can be subclassified into three categories – silent, tonic-firing, and burst-firing (Yang et al.). This study targets the rhythmic burst firing of neurons in the Lateral habenula (LHb), which has been discovered to play a major role in depression manifestation and has become a large topic of interest in pathophysiological research of

depression in recent years. Hyperactivity in the Lhb neurons, predicted to promote neurotransmitter release to respective regions in the brain (Shao et al.) has been exhibited in numerous models with depression. More specifically, a known alteration to the individual burst of an Lhb neuron (excluding the quantity, duration and other more holistic factors of bursts), is the hyperpolarized resting membrane potential (RMP) following each burst (Young, Lyons and Piggins).

### 1.6 Hyperpolarization in burst-firing neurons

Prior to a study conducted in 2018, the consensus was that increased Lhb burst firing was associated with depressive symptoms. A Kir4.1 channel was also hypothesized to cause hyperpolarization of Lhb neurons, and so correlations between the channel and hyperpolarization, and between burst-firing and depression were established. However, it wasn't yet addressed whether hyperpolarization was what was causing the bursting, which caused depression. In the 2018 study conducted by Cui et al., through using a whole-cell patch clamp to force hyperpolarization, it was revealed that hyperpolarization of the Lhb neurons was a direct cause of increased Lhb burst firing.

Another study investigating the effect of blocking Ketamine on Lhb bursting activity discovered that the RMP of Lhb neurons ranged from  $\sim -70\text{mV}$  to  $\sim -30\text{mV}$ , depending on the type mode of bursting – silent, tonic or burst (Figure 1.1) (Yang et al.).

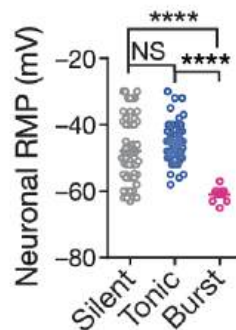


Figure 1 - There are three modes under which Lhb neurons can be sorted – silent, tonic-firing, and burst-firing. Out of the three, burst-firing Lhb neurons exhibit the most hyperpolarised resting membrane potential (RMP) (Yang et al.)

The Kir4.1 channel in the astrocyte membrane that surrounds an Lhb neuron is a part of the glial Kir channel and is primarily responsible for the mediation of  $\text{K}^+$  conductance and regulation of the resting membrane potential (RMPs) in astrocytes.  $\text{K}^+$  released extracellularly is removed by astrocytes which depend on a Kir4.1-related mechanism, meaning that downregulation of the said channel would result in an increased  $[\text{K}^+]_{\text{out}}$  and a depolarized neuronal RMP, upon further calculation (according to the Nernst Equation). Therefore, the study hypothesizes that upregulation of Kir4.1 should lead to a decreased  $[\text{K}^+]_{\text{out}}$  and hyperpolarize neuronal RMP, thereby increasing neuronal bursting. Upon testing this hypothesis, 8 out of 15 originally tonic-firing neurons transitioned into bursting mode, indicating that overexpression of Kir4.1 that causes enhanced extracellular  $\text{K}^+$  clearance may underlie the neuronal

hyperpolarization required before burst initiation (Cui et al.). This finding is corroborated by further studies that discovered significantly more hyperpolarized RMPs in bursting neurons, compared to silent and tonic-firing ones, measured using a whole-cell patch clamp (Yang et al.). A study that investigated the properties of the low-voltage-activated transient  $\text{Ca}^{2+}$  current that underlies burst firing of LHB neurons also discovered that it had an activation threshold of near  $-65\text{mV}$  and that the LHB neurons exhibit a biphasic inactivation process (Huguenard, Gutnick and Prince). Consequently, the 2018 study used a ramp-like mechanism (Figure 1.2) to inject a hyperpolarizing current into LHB brain slices. The induced hyperpolarization elicited high-frequency burst firing of LHB neurons, a clear indicator of its relationship to burst-firing.

Two other components that have been identified to be involved in LHB neuronal burst mechanisms are the *N*-methyl-D-aspartate receptors (NMDAR) and the low-voltage-sensitive T-type calcium channels (T-VSCCs). The importance of NMDAR-mediated calcium influx in burst generation has been highlighted in brain regions of other organisms, including in lampreys and rats (Zhu et al.) (Schiller et al.) (Grillner et al.). The study conducted by Yang et al. discovered the complete stopping of burst firing upon application of specific NMDAR antagonist AP5, supporting the idea that NMDARs are required for burst generation in the LHB, and therefore opening further paths for ketamine, an NMDAR-blocker, to act as a potential treatment for depression.

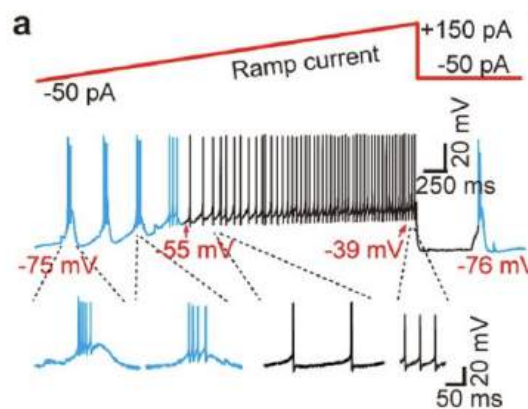


Figure 2 - Indicates that a ramp-like current can stimulate burst-firing in LHB neurons. The blue traces indicate the region that was hyperpolarised. The diagram with an upwards-sloping red line represents the ramp-like hyperpolarizing current with parameters ( $-50\text{pA}$  to  $+150\text{pA}$ ), that was injected into LHB brain slices via an electrode for a duration of  $100\text{ms}$ . The action potential diagrams below indicate the stimulation of LHB neurons burst firing (blue), compared to the non-bursting neurons (reference: Cui et al.)

Upon injection of tetrodotoxin, a toxin that blocks sodium channels (CDC), a plateau potential was left in bursting neurons, which suggested that the ionic component of the plateau potential is unlikely to be sodium, but rather calcium. When the membrane potential is hyperpolarized for longer than  $100\text{ms}$ , the pacemaker channel T-VSCC can be de-inactivated to initiate burst firing, which was detected in both rat and mouse LHB neurons (Yang et al.). Hence,

it is possible that depression directly or indirectly hyperpolarises the RMP of LHb neurons to a level that de-inactivates T-VSCCs that stimulate more bursting.

### ***1.7 Aspects of Depression Manifestation***

Common behavioral indices of clinical depression that have been used previously as markers in mouse models most notably include anhedonia, despair, impaired locomotor activity and decreased activity in behavioral tests (Zhang et al.) (Liu et al.).

## **2. Research Question**

This study aims to investigate the association between burst firing of LHb neurons and depression, as well as the implications of targeted inhibition by antidepressants as a potential form of treatment.

To do this, this study aims to design a hypothetical experiment to investigate the specific neural networks associated with clinical depression and how their burst firing patterns are impacted by its manifestation. The study also aims to propose various chemicals that may inhibit specific researched networks, to provide a framework for potential pharmacotherapies to treat clinical depression.

## **3. Specific Aims**

The genetics of the mice are so well-researched that detailed procedures, such as the isolation of a single neural pathway, can be carried out. There also exists multiple established models for inducing depression-like behaviours in mice. Hence, mice are the most appropriate model for investigating complex neural mechanisms underlying depression.

In this study, the outputs of burst-firing and tonic-firing LHb neurons will be investigated, as well as the role of ketamine as a potential drug for treatment.

### ***AIM 1: What impacts do different bursting modes of LHb neurons have on the VTA?***

Current research confirms that the projection of LHb neurons to GABAergic interneurons within the VTA, DRN and MnR is direct, and that burst firing enhances synaptic transmission and facilitates synaptic plasticity. Another discovery is that depressive symptoms involve a proportion of tonic-firing neurons transitioning into burst-firing mode. (Liu et al.) However, the neurological implications of the different modes of firing on the inhibitory interneurons within the respective regions to which the LHb neurons project remain unknown. In this study, the differences between the output effect of tonic-firing and burst-firing neurons are investigated in one such pathway, specifically from the LHb to the VTA, indicated by the number 4 in Figure 3.1. The study hypothesizes that burst-firing neurons would exhibit greater output on the inhibitory interneurons within the VTA, potentially indicated by a greater number of action potentials in the post-synaptic region, or a greater voltage input in electrophysiological recordings.

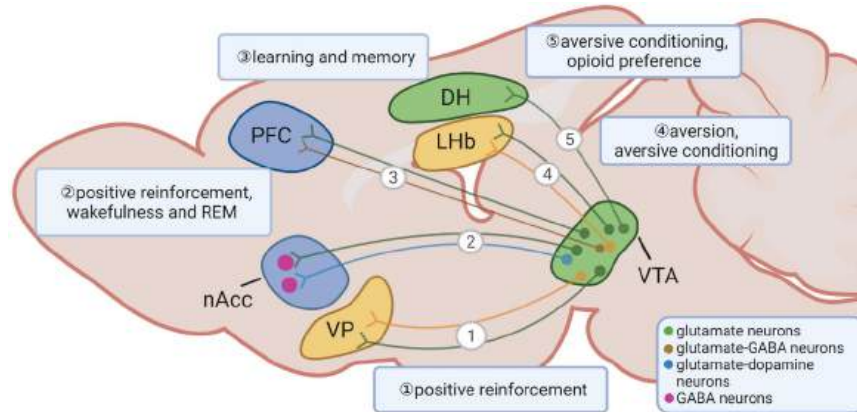


Figure 3 - Multiple regions of the brain project to the VTA. The specific pathway investigated in this study is from the LHb to the VTA, labelled with the number 4 (reference: Cai and Tong)

**AIM 2: The mechanisms of ketamine as an anti-depressant in depression treatment**

A major class of antidepressants are SSRIs, which are monoaminergic drugs, meaning they act directly on monoamine neurotransmitters such as serotonin, dopamine etc. by blocking the reuptake of them at synapses, thus increasing their extracellular concentrations (Mayo Clinic). However, they have a major limitation of a lag period, sometimes lasting up to three weeks, before their therapeutic effects take place (Maeng et al.). Studies also show that although the classical SSRI-type antidepressant fluoxetine reduces bursts after chronic treatment for depression, it does not instantly block LHb bursts. These findings suggest that reduced burst firing of LHb neurons may be a common endpoint for antidepressant drugs, but that bursting is not directly impacted by the application of these drugs. Lastly, another point worth noting is that SSRIs require repeated administration (usually once daily, but up to 3 times per day depending on the severity of the depression) (Chu and Wadhwa).

To address these limitations, more attention has been turned to glutamatergic-based treatments such as ketamine. Ketamine has been discovered to elicit a much more rapid antidepressant effect (Yang et al.), and unlike SSRIs, it targets the NMDAR receptors in the LHb directly, with studies showing that it reduces LHb burst-firing and mediates depressive symptoms in behavioral paradigms (Figure 3.2) (Yang et al.). Additional studies show that ketamine has a half-life of 13 minutes in mice (Maxwell et al.), and 186 minutes in humans (Clements et al.). Surprisingly, however, the antidepressant and antisuicidal responses last much longer, up to 24 hours in mice, and 3-14 days in humans (Berman et al.) (Autry et al.), even after the elimination of ketamine. One study discovered that the long-term antidepressant effects of ketamine are decreased upon stimulation of LHb neurons at low ketamine concentrations (Ma et al.).

This study aims to build on its basis (Ma et al.), to investigate the following three questions:

*Specific aim 2.1:* Whether the stimulation of the LHb at high ketamine levels is a possible mechanism that enables more ketamine “trapping”, i.e. a larger amount of ketamine retained within NMDAR channels, enabling more blockage.

*Specific aim 2.2:* Whether negative stimuli post-ketamine injection shortens the long-term antidepressant effects of ketamine.

*Specific aim 2.3:* Whether varying concentrations of ketamine per injection would amplify its long-term antidepressant effects.

All three aims would elucidate mechanisms that could prolong the long-term antidepressant effects of ketamine.

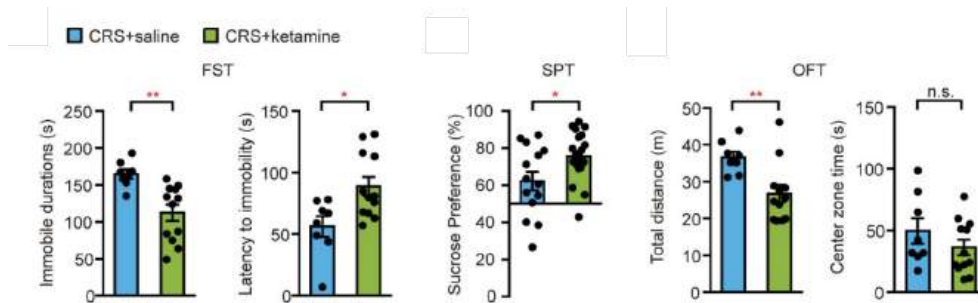


Figure 4 - Ketamine efficacy was shown by comparing the results of two groups of CRS-induced mice tested in a series of behavioral paradigms. The group of mice without ketamine injection displayed higher immobile duration, lower latency to immobility and lower sucrose preference, all classified as depressive-like behavior. (reference: Yang et al) (Behavioral paradigms: Forced Swim Test (FST), Sucrose Preference Test (SPT), Open Field Test (OFT))

## 4. Testing the Aims

### 4.1 AIM 1: Determining the effect and specific outputs of bursts on the VTA

LHb neurons project to three main regions – the dorsal raphe nucleus (DRN), ventral tegmental area (VTA), and median raphe nucleus (MnR). All three regions are involved in reward processing, including regulating the release of hormones such as dopamine and serotonin. The VTA, more specifically, is well known for mediating dopamine (DA) release in downstream regions and is also known to be heterogeneous, containing other neurons such as glutamate neurons. Compared to the extensive research on DA neurons, the role of VTA glutamate neurons is understudied (Cai and Tong). However, recent studies have hinted at the importance of glutamate neurons in regulating behaviors related to reward processing, stress modulation, addiction and more, all of which are associated with depression. Given that retrograde viral tracing conducted in a previous study revealed that projections from LHb neurons to the three regions are largely separate (Liu et al.), the study aims to investigate the neural pathway of LHb<sup>VTA</sup> neurons to the VTA specifically and to determine the output of different modes of bursting on the region.

#### *4.1.1 General Overview*

In a simplified outline of the methodology, the study first begins by mapping out the projection pathways of LHB neurons to the VTA through Retrobead analysis and retrograde tracing. To do so, the neurons must be stimulated for the pathway to light up and be detected. This stimulation is also essential for the burst-firing modes to be manipulated and controlled, allowing the effects of LHB neuronal activity on the VTA to be investigated in the main part of the aim. In order to control the stimulation of LHB neurons, retrograde injection of an adeno-associated virus coupled to channelrhodopsin is carried out in the LHB and the VTA, then by flashing light according to a fixed stimulation protocol, the LHB neurons can be manipulated to begin burst-firing or tonic-firing. The effect of the LHB neurons on the VTA can be recorded in two ways – electrophysiologically, by implantations of tetrodes coupled to an optical fibre; or optogenetically, through voltage-sensitive dye injections from which voltage can be measured using a computerized system.

#### *4.1.2 Specific methodology*

The study hypothesizes that the output to inhibitory interneurons in the VTA is strengthened by higher frequencies of burst-firing mode compared to tonic-firing and silent modes, thereby suppressing dopamine release by the interneurons in the VTA. To investigate this, the pathways through which LHB neurons signal to the VT must first be identified. Current literature suggests VTA projections originate from two different LHB subpopulations, one direct and another indirect (Figure 4.1) (Groos and Fritjof Helmchen). The direct pathway involves the medial LHB division projecting immediately onto the DA interneurons mainly located in medial part of the VTA, which then project to the medial prefrontal cortex (mPFC); the indirect pathway involves neurons in the lateral part of the LHB projecting to the Rostromedial tegmental nucleus (RMTg), predominantly composed of GABAergic neurons that then project to the VTA (Sun et al.). The ratio of the type of neurons LHB<sup>VTA</sup> neurons target remains controversial, with some studies suggesting they target dendrites of DA neurons and GABAergic ones to a similar extent (Omelchenko, Bell and Sesack), while another discovering LHB neurons target a higher percentage of GABAergic VTA neurons with much less direct contact on DA neurons (Brinschwitz et al.).



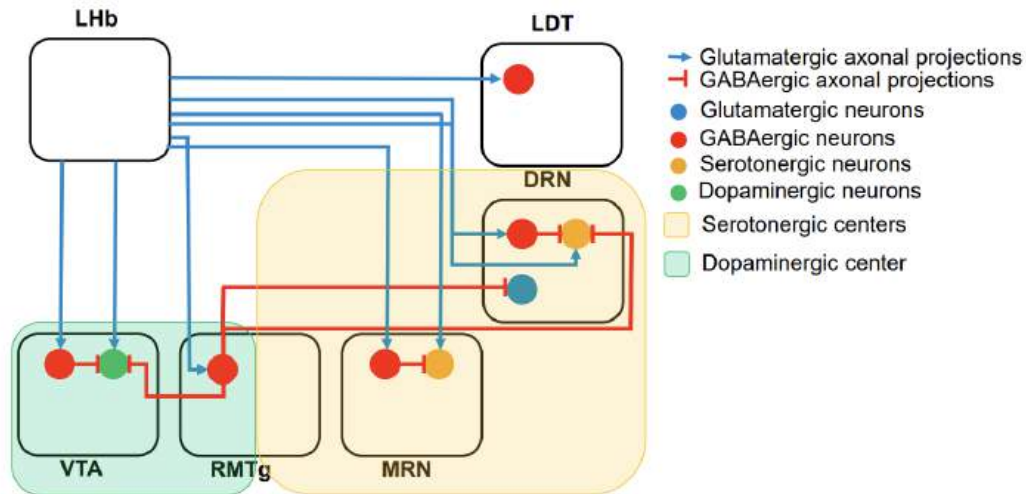


Figure 5 - There are two pathways, direct and indirect, through which LHB neurons project to the VTA. The direct pathway is indicated by the vertical, downward arrow to the very left of the diagram, directly from the LHB to the DA neurons (green). The indirect pathway is highlighted by the connector arrow (blue) from the LHB to the RMTg (middle bottom box), then the red arrow from the RMTg to the DA neurons (green) in the VTA (reference: Groos and Fritjof Helmchen)

In order to test the pathway of projection by LHB neurons, and subsequently to control and shift between burst-firing and tonic-firing modes of the neurons, a method of stimulation must be established. Optogenetics involving opsins can be used to manipulate LHB neurons to fire at a controlled rate, mimicking tonic-firing and burst-firing modes. Opsins are light-activated membrane-bound proteins that can cause cell activation (i.e. depolarization), cell inhibition (i.e. hyperpolarization), or activate other signalling cascades (Guru et al.).

One inhibitory opsin is eNpHR3.0, which is a fast-light-activated electrogenic chloride pump that inhibits neural activity (Gradinaru, Thompson and Deisseroth). eNpHR3.0 is determined to have a maximum excitation at 590nm, and can be stimulated by wavelengths of green, yellow or red light (Guru et al.).

As depression has been found to be associated with increased LHB neuronal activity involving higher spike frequency, a study showed that it was not merely the increase in spikes that was responsible for inducing depressive-like symptoms, but rather the specific pattern of burst-firing (Yang et al.). This was proven by using a stimulation protocol that resulted in same overall firing rate as eNpHR3.0 – 5 Hz on adeno-associated virus (AAV)-oChIEF, but which did not display similar phenotypes of depressive symptoms (Yang et al.).

This investigation therefore would use the inhibitory opsin eNpHR3.0 at 1Hz, 100ms, and 589nm to stimulate rebound bursting in LHB<sup>VTA</sup> neurons, which has triggered bursting in over 90% of trials in previous studies (Figure 4.2) (Liu et al.) (Yang et al.), and hence can be used to mimic neurons in burst firing mode.

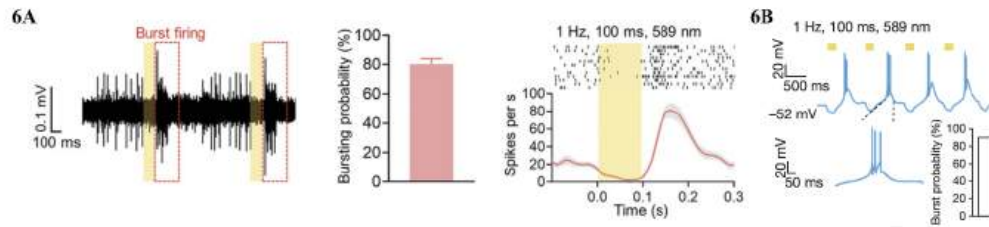


Figure 6 - Burst stimulation using eNpHR3.0 was proven to be successful, as shown by the boxed sections that indicate burst firing in panel 6A. Bursting probability also increased up to 80%, and spike frequency increased to ~80 spikes per second using stimulation protocol 1Hz, 100ms, 589nm. In panel 6b, the same stimulation protocol was used and yielded similar results, with a success rate of over 90%. (Liu et al.) (Yang et al.)

The LHb has been previously discovered to receive glutamatergic signals from the lateral hypothalamic area (LHA), therefore the VGlut2 LHA-LHb terminals would be the main target for optogenetic stimulation in this study. The specific method of adenovirus injection follows that used in a previous study investigating the convergence of aversion and reward signals in individual neurons of the mouse lateral habenula (Li et al.). Following anesthesia by 1% pentobarbital (80 mg kg<sup>-1</sup>) of VGlut2-ires-cre mice, and ensuring a stable body temperature of 37°C (Li et al.), 300 nl of pGM-AAV-hsyn-eNpHR3.0-EYFP (Adeno Associated Virus (AAV) eNpHR3.0) would be slowly injected with a micro syringe pump into the bilateral LHb region, via a glass pipette at a flow rate of 100 μlmin<sup>-1</sup>. After this, the glass pipette would be left in the region for a minimum of 10 minutes before withdrawing. The LHb neurons in the mice can then be stimulated using a light stimuli procedure of 1Hz, 100ms, 589nm, and subsequently used to record the activity of VTA neurons following rebound burst excitation.

After establishing a method of LHb neuronal stimulation, the aforementioned mapping of the LHb<sup>VTA</sup> projection pathway begins by labelling the LHb<sup>VTA</sup> neurons with fluorescent Retrobeads (Figure 4.3), integrating the procedures of two studies that investigated the neural activity patterns in response to chronic mild stress (CMS) in mice (Cerniauskas et al.), and the effects of an inhibitory circuit within the VTA on reward promotion (Stamatakis et al.). 200nL of red/green fluorescent retrobead labelling would be injected using a 1μL Hamilton syringe into the VTA. The LHb neurons are also tagged with channelrhodopsin-2 (CHR2) by retrograde injection of an adeno-associated virus (AAV). The specific virus used would be AAV5-EF1a-DIOhChR2(H134R)-eYFP; AAV5-EF1a-DIO-eYFP; AAV5-CaMKII-ChR2-eYFP, with an injection volume of 300-500nL of concentration virus solution into the LHb. The injection needle is withdrawn at least 10 minutes before extraction after the end of infusion. Then, a minimum of 3 weeks is given for the mice to recover and for retrograde transport to take place. Light illumination of the LHb<sup>VTA</sup> neurons may accelerate this process, allowing the output of the VTA to be recorded optically sooner. An advantage of the Retrobead analysis is that the pathway of LHb<sup>VTA</sup> neuron projection can be mapped, locating the neurons in the LHb that are burst-firing and causing the output in the VTA.

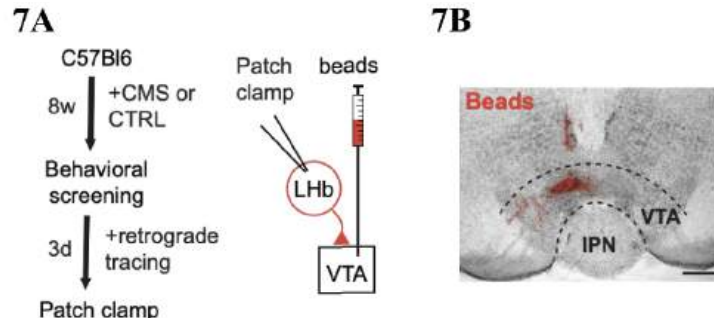


Figure 7 - panel 7A is a simplified pictorial form of experiment setup – patch clamp recording in the LHb neurons, which projects to the VTA, retrograde injection of Retrobeads into the VTA. Panel 7B then shows the location of Retrobead injection (reference: Cerniauskas et al.)

### ***Tonic firing versus burst firing***

As of right now, there are no distinct stimulation protocols that detail the specific frequency and duration at which light must be flashed to differentiate between modes of firing – i.e. tonic firing versus burst firing. However, it is known that to stimulate tonic firing, a depolarising current is required. This can be achieved by employing ChR2, a channelrhodopsin that allows fast depolarisation through stimulating ion channels (addgene). ChR2 has a peak response spectra of 450nm, and modifications can be made to the stimulation protocol for burst-firing to elicit tonic-firing.

For the burst-firing mode, the methodology undertaken in previous studies that involved stimulating rebound bursting through a hyperpolarizing current injection was able to stimulate a robust burst-firing pattern that resembled the burst-firing patterns exhibited in mouse models of depression (Li et al.). Hence, by injecting both the inhibitory and the excitatory channelrhopsins, either tonic firing or burst firing can be elicited using a dual laser (one at 450nm, one at 589nm).

Two methods can then be used to measure the output in the VTA region:

### ***Electrophysiological recording***

Within three to six weeks after virus injection, electrophysiological recording in vivo can be applied downstream in the VTA region, where the effects of the LHb neuronal activity, including RMP and extent of hyperpolarization can be recorded and analyzed. In order to do so, the specific instruments used would be involve plating four tetrodes made of four nichrome wires (diameter: 17µm) that extend around an optical fibre (diameter: 200µm), as was done in a previous study (Li et al.). That way, single-cell recordings can be obtained and analyzed more in-depth. At the end of recording, a DC current of 30s, 0.1mA is injected to electrically damage the VTA, and through stimulating spike firing in ChR2-expressing neurons with low-intensity light stimuli, the specific cell type within the bilateral LHb region that was stimulated can be identified. The type of LHb neuron most likely stimulated is glutamatergic, as indicated by the results of previous studies (Li et al.).

Once the tetrode is fixed in place and the mice recover from anaesthesia, data collection can take place. Following the procedure from a study investigating chronic stress and its neurological impacts on the LHB, the “Rank Surprise” burst detection algorithm can be used to detect bursts in spike trains and hence analyze the effects of a given pattern of bursting on the excited neurons in the VTA. An advantage of this compared to the simpler ISI-threshold burst detection method, as noted by the study, is that tonic spikes identified as bursts can be eliminated, and the number of false positives greatly reduced (Cerniauskas et al.).

### ***Voltage-sensitive dye (VSD)***

Another method that could enable the output of the VTA to be measured in vivo without the use of a tetrode is by injecting a voltage-sensitive dye (VSD). One advantage of this compared to electrophysiological methods is that the field output of large populations of neurons can be measured and recorded at once, versus tracing an individual LHB neuron, which would be much more difficult as the chances of finding the specific cell activated in the VTA by the chosen cell stimulated in the LHB using an electrophysiological method is very low (electrophysiologically, unrealistically large electrode arrays would be needed to achieve the same effect). Unlike extracellular electrophysiology techniques that can only detect characteristics such as spiking activity, there’s the opportunity to detect other aspects such as subthreshold synaptic potentials.

As of now, there is limited research that has used voltage-sensitive dye in the VTA region of mice, however, the procedure has been used in a study that conducted voltage-sensitive dye imaging of primary motor cortex activity by VTA stimulation (Kunori, Kajiwara and Takashima) can be referenced.

To introduce the dye into the brain, past studies (Kunori, Kajiwara and Takashima) first had to expose the motor cortex of the models used for in vivo imaging. In the case of this study, the simplest way to obtain even staining is also by applying the dye directing onto its surface, which may be slightly more challenging in the VTA. By referencing a study that investigated neural connections between the sensorimotor cortex and the VTA (Ferezou et al. ), a craniotomy can theoretically be performed over the region of interest. However, as the VTA is situated deep in the midbrain, adjacent to the substantia nigra (Kalivas), it may be better to use methods such as microinjection pipettes or viral transduction to introduce the dye into the VTA.

The specific dye used in the study can be VSD RH-795 (Invitrogen) dissolved at 0.8 mg/ml in ACSF containing 125 NaCl, 5KCl, 2 CaCl<sub>2</sub>, 1.25 MgSO<sub>4</sub>, 1.25 NaHPO<sub>4</sub>, 22 NaHCO<sub>3</sub> and 10 glucose, referencing a study investigating primary motor cortex activity by VTA stimulation (Kunori, Kajiwara and Takashima). Following dye incubation, the excess dye is washed away and neural activity (i.e. the output of the VTA) can be measured by changes in fluorescence. This is recorded by a Micam01 system using a tandem-type epifluorescence microscope, and optical signals can be expressed as fractional changes in fluorescence (Kunori, Kajiwara and Takashima).

#### 4.1.3 Expected outcomes and interpretations:

Assuming that the burst-firing mode of neurons increases the output to the VTA, we would expect to see an increased voltage output to neurons in the VTA detected by the voltage-sensitive dye. From the electrophysiological recordings, burst-firing neurons would be expected to cause greater VTA stimulation, which would manifest in subthreshold characteristics such as higher burst frequencies within bursts. Conversely, tonic-firing neurons would stimulate a lower frequency of bursts, as well as a lesser extent of hyperpolarisation between each series of bursts.

#### 4.2 AIM 2: Investigating potential methods of extending Ketamine efficacy

The study aims to confirm whether stimulation of the LHb at high ketamine concentrations encourages further trapping, whether minimising negative emotional episodes after ketamine injection slows down the process of ketamine untrapping, and if increasing ketamine concentrations per injection would boost blockage of NMDAR receptors, all theoretically extending ketamine efficacy.

In order to investigate the short and long-term physiological impacts of ketamine as an anti-depressant, mice must first be stimulated to express depressive symptoms. There are several methods for depression induction in mice. The first one is a well-established mouse model of depression, in which the mice would be exposed to a series of aversive stimuli for 28 days, including foot shock, 20 times per day; air puff, 20 times per day; fox urine, 30 minutes per day; physical restraint, 1 hour per day (Liu et al.).

Another method is the lipopolysaccharide (LPS)-induced depressive-like rat model. According to a study that investigated the effects of Astroglial Kir4.1 in the LHb, one week of LPS injection at 0.5mg kg<sup>-1</sup>/day, intraperitoneal) can induce a strong depressive-like phenotype in three-month-old Wistar rats (Cui et al.).

Mice subjected to chronic restraint stress (CRS) is also another established animal model of depression. It has been shown to instigate reliable depression-like phenotypes (Figure 4.4), as well as increased burst firing in the LHb (Yang et al.). The CRS model involves placement in 50mL conical tubes with holes for airflow for 2-3 hours per day for 14 consecutive days (Kim and Han).

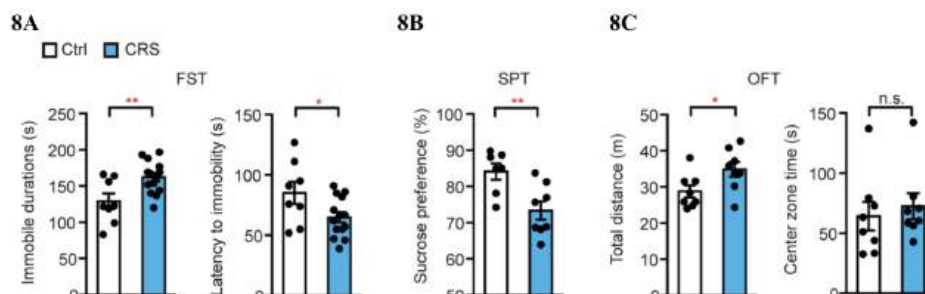


Figure 8 - The chronic restraint stress model is shown to be effective in mediating depressive symptoms as the group of mice subjected to CRS demonstrated higher immobile durations, lower latency to immobility, lower sucrose preference, and greater centre zone time, all indicators of depressive behavior. (reference: Yang et al.) (Behavioral paradigms: Forced Swim Test (FST), Sucrose Preference Test (SPT), Open Field Test (OFT))

Out of the three models, the mouse model involving exposure to a series of aversive stimuli is the most suitable to begin with, as it is the most comprehensive and well-established. The experiments would then ideally be repeated with the other models to ensure generality.

As aim 2 involves investigating ketamine efficacy, the exhibition and extent of depressive-like behavior must be tested and measured. Under the expected outcomes and interpretations of each aim, the specific behavioral paradigms that would be used are introduced.

*4.2.1 Specific aim 2.1: Whether the stimulation of the Lhb right after ketamine injection is a possible mechanism that enables more ketamine trapping.*

Lhb bursts critically depend on NMDARs (Figure 4.5), and ketamine is an effective NMDAR receptor blocker. A study that investigated sustained antidepressant effects of ketamine (Ma et al.) theorised that if more Lhb NMDARs could be blocked by ketamine treatment, the recovery of those channels would be delayed and hence the therapeutic effects of ketamine could be extended. This study wishes to build upon this theory and following the stimulation procedure of the Lhb neurons when investigating previous aims, test whether different intensities of Lhb stimulation would affect the extension of ketamine's antidepressant effects.

Following the study's (Ma et al.) methodology, the mice expressing depressive symptoms would be injected with ketamine at  $5\text{mg kg}^{-1}$  when the ambient ketamine level is high ( $>6\mu\text{M}$ ) and is greater than its dissociation constant ( $K_d$ ), i.e. immediately after ketamine injection (Figure 4.6, Figure 4.7). Then, using the inhibitory opsin eNpHR3.0 and stimulation protocol (1Hz, 100ms, 589nm) previously used when investigating aim 1, rebound bursting would be stimulated in Lhb<sup>VTA</sup> neurons. In the following trials the stimulation protocol would be ramped up in increments by increasing amplitude and shortening the duration, thus causing higher frequency bursts in a shorter amount of time, theoretically increasing the number of NMDAR channels opened to allow greater ketamine binding and hence blocking of the NMDAR channels, termed 'ketamine trapping'. Then, the mice from each stimulation intensity would be tested using the following behavioral paradigms: the sucrose preference test (SPT) and forced swim test (FST) 10 hours post-stimulation.

### **Sucrose preference test (SPT)**

The SPT is carried out in the dark phase, and prior to testing, food would be provided ad libitum. To initiate the SPT, the mice would first be isolated and individually housed with 2 bottles of water for 2 days, then given 2 bottles of 2% sucrose for another 2 days. Then, the mice would be deprived of any water for 24 hours and subsequently exposed to a bottle of 2% sucrose and water for 2 hours in the dark phase. The positions of the two bottles are switched after an hour and the volume of sucrose and water consumed is recorded. Sucrose preference is

determined by the percentage of sucrose solution consumed in comparison to total volume intake (methodology adapted from (Yang et al.)). In the Sucrose Preference Test, mice that exhibit depressive symptoms, specifically anhedonia, have a significantly reduced preference for water containing sucrose. (Shieh and Carter)

### Forced swim test (FST)

The FST is used to assess and measure the immobility of mice. The FST is performed in the light phase, and in the FST, mice would be placed in a cylinder (diameter 12 cm, height 25 cm) of water 23-25°C and allowed to swim for 6 minutes under normal light conditions (Powell, Fernandes and Schalkwyk). Video cameras would be positioned to the side of the cylinder to record animal behaviors. In the last 4 minutes of the test, immobile time would be recorded. Immobility is quantified through immobile time, in which immobile is defined as mice remaining floating or motionless, displaying movements only necessary for keeping their balance in the water (methodology adapted from Yang et al.). In the Forced Swim Test, depressive-like behavior is indicated by lower latency to immobility and greater time spent immobile. Although FST has been pharmacologically validated for depression (Knackstedt and Schwendt), it is worth noting that there is controversy surrounding whether decreased immobility is a true indicator of depression, as some suggest rodents may remain immobile to conserve energy, which cannot be classified as an expression of learned helplessness, but rather a demonstration of adaptation (Knackstedt and Schwendt).

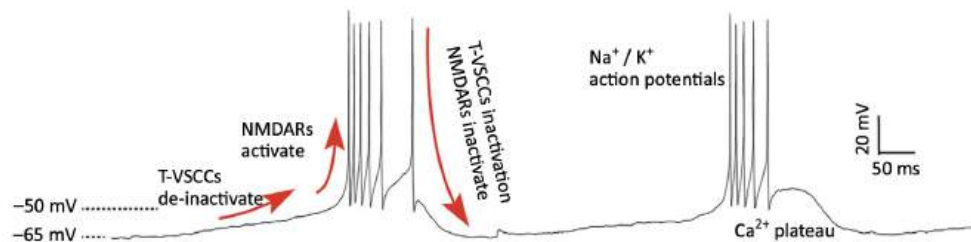


Figure 9 - At an RMP of -65mV, T-VSCCs de-inactivate and NMDARs activate, stimulating the burst-firing of neurons, as shown by the subsequent appearance of burst spikes. Following T-VSSCs and NMDARs inactivation, burst firing is terminated. Burst firing of an LHB neuron shows that is dependent on NMDARs for burst initiation and activation (reference: Cui, Hu and Hu)

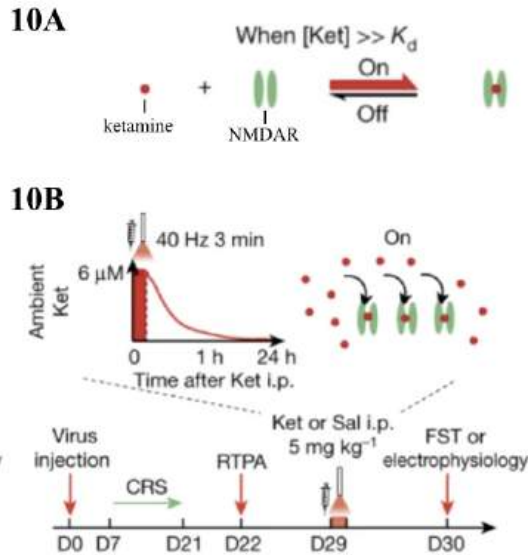


Figure 10 - Demonstrates how Lhb stimulation post-ketamine injection affects the extent of ketamine binding. This finding is summarized by panel 10A, which suggests when ambient ketamine concentration  $[Ket]$  is greater than the dissociation constant ( $K_d$ ), the interaction between ketamine and NMDAR leads to greater binding (ketamine trapping). Panel 10B shows the theoretical ambient ketamine levels following ketamine injection and Lhb stimulation, with protocol 40 Hz for 3 minutes. Immediately after the injection of ketamine, ambient ketamine levels rise to  $6\mu\text{m}$ . They then drop to  $\sim 0.23\mu\text{m}$  one hour after injection. The study discovers that Lhb stimulation right after ketamine injection (time after Ket i.p.: 0 h) with protocol 40 Hz for 3 minutes leads to greater ketamine trapping. (edited, reference: Ma et al.)

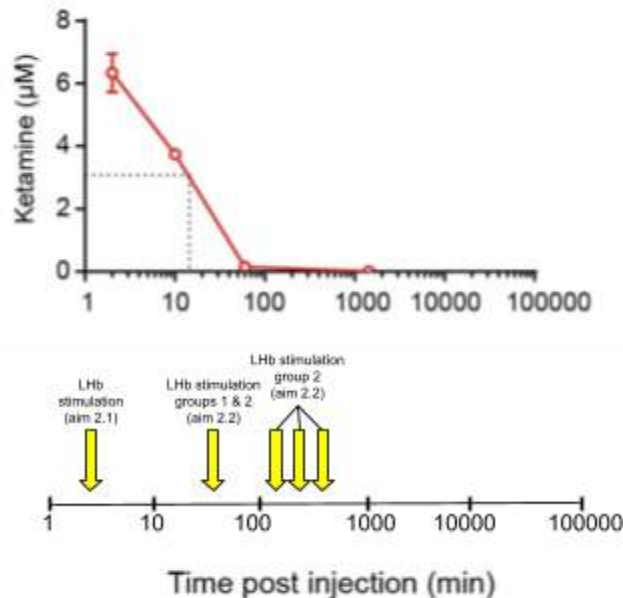


Figure 11 - above is a more detailed depiction of the theoretical concentration of ketamine in the brain, post-ketamine injection. Right after injection, ketamine levels are at  $6\mu\text{m}$  and subsequently complete one half-life in around 15 minutes post-injection. Below shows the differences in time at which Lhb would be stimulated when investigating aims 2.1 and 2.2 (edited, reference: Ma et al.)



### **Expected results and interpretations:**

Assuming the hypothesis that stimulation of the LHb right after ketamine injection, i.e. when ambient ketamine levels are still larger than  $K_d$ , enables increased binding between ketamine and NMDAR channels, the group of mice that received the greatest amount of LHb stimulation post-ketamine injection would be expected to exhibit the least depressive-like symptoms. As mentioned previously, this would be evident in the highest preference for sucrose in the SPT among the groups tested and the higher latency to immobility in the FST.

#### *4.2.2 Specific aim 2.2: Whether negative stimuli beginning 1-hour post-ketamine injection shortens the long-term antidepressant effects of ketamine.*

A previous study also discovered shortened antidepressive effects upon stimulation of the LHb at low ketamine levels (Ma et al.). Stimulation of the LHb post-ketamine injection, when  $[Ket] \ll K_d$  (Figure 4.9) is hypothesized to open NMDAR channels, releasing the ‘trapped’ ketamine that was postulated to be causing the long-term antidepressant effects. To further elaborate on this, the study also aims to investigate whether increasing the frequency of exposure to negative stimuli post-ketamine injection would shorten its long-term antidepressant effects. To do so, after the initial ketamine injection, the mice would be split into two groups and their LHb would be stimulated with a low and high frequency 1h post-injection (Figure 4.7); these mice would subsequently be tested for depressive symptoms using behavioral paradigms. The LHb would be stimulated following the method used to stimulate bursting when addressing aim 1, and both groups of mice would be exposed to stimulation once, 1h post-injection. Only one group of mice would go on to be stimulated three more times, 2 hours post-injection, 4 hours post-injection, and so on. Both groups of mice would then be tested with the following behavioral paradigms: Open Field Test (OPT), Tail Suspension Test (TST), 10 hours post-ketamine injection.

### **Open field test (OPT)**

The OFT is used to assess levels of anxiety and motor activity in mice (Liu et al.). It is carried out in the light phase and in the OPT, mice would be placed in a plastic box, size 40 cm x 40 cm x 40.5 cm, in a room with dim light for 10 minutes (methodology adapted from (Yang et al.)). Video cameras would be positioned on opposite sides to capture animal movement, and the software Ethovision XT would be used to measure the locomotion and activity of the mice in the box. (methodology adapted from (Liu et al.))

In the Open Field Test, most rodents remain around the borders of the walls and avoid the centre. After adapting for some time, normal mice would venture into the centre, compared to the anxious mice, which would spend significantly less time in the open area and more time near the edges (Figure 4.8) (Shieh and Carter).

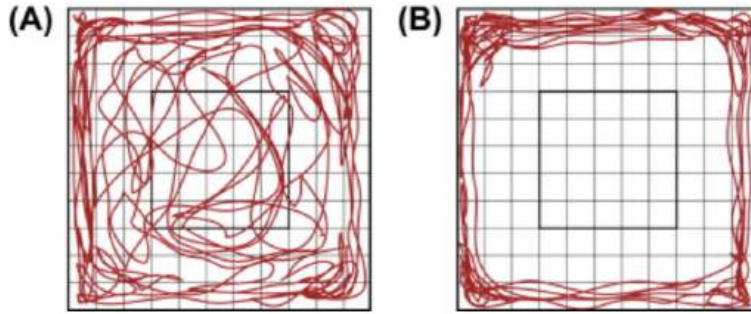


Figure 12 - Comparison between the path tracing of mice in the Open Field Test. (A) indicates the locomotion of normal mice, whereas (B) refers to those who exhibit anxiety. As shown from the tracings, mice that exhibit anxiety are much more confined to the walls of the box and completely avoid the centre, compared to the normal mice, which have path tracings all over the grid (reference: (Shieh and Carter))

### Tail suspension test (TST)

TST is used to assess the level of despair-like behavior in mice (Liu et al.). The TST assay is performed during the light phase, and in this procedure, the tails of the mice would be fixed to a horizontal bar using adhesive tape for 6 minutes, and immobility would be recorded in the last 4 minutes (methodology adapted from (Liu et al.)).

In the Tail Suspension Test, which is a test for despair, mice displaying depressive-like behavior exhibit longer durations of immobility. (Proenca et al.)

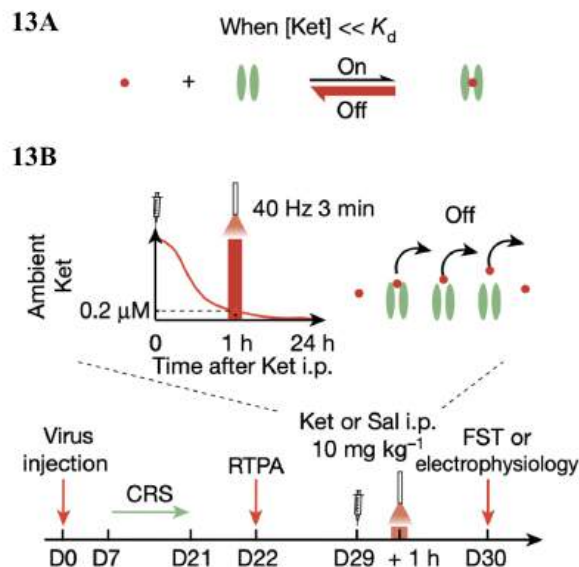


Figure 13 - The study also discovered that LHB stimulation at time: 1h after Ket i.p. with protocol 40 Hz for 3 minutes led to greater ketamine release from NMDAR channels – ‘ketamine untrapping’. This finding is summarized by panel 13A, which suggests when ambient ketamine concentration [Ket] is less than the dissociation constant ( $K_d$ ), the interaction between ketamine and NMDAR leads to dissociation (reference: Ma et al.)

### **Expected results and interpretations:**

Assuming that stimulation of the LHb when ambient ketamine levels are lower than  $K_d$  (i.e. 1 hour after ketamine injection) encourages ketamine unbinding (ketamine untrapping), the group of mice that was exposed to stimulation four times post-ketamine injection would be expected to exhibit much more depressive-like symptoms in the behavioral paradigms. To assess the significance of these results, as well as for all the other quantitative results from other aims, a T-test can be performed, using a significance level of  $p=0.05$ .

#### *4.2.3 Specific aim 2.3: Whether varying concentrations of ketamine per injection would amplify its long-term antidepressant effects.*

To address this aim, different levels of ketamine would be administered to separate groups of mice to produce varying ketamine concentrations in the brain. A local ketamine concentration of  $5.2 \pm 1.3 \mu\text{M}$  can be achieved both through cannulae infusion of  $25 \mu\text{g}$  ketamine, and intraperitoneal injection of  $255 \text{mg kg}^{-1}$  (Yang et al.).

Drug delivery of ketamine, following the methodology adopted by a study that investigated the pathophysiological effects of ketamine in the LHb, involves intraperitoneally injecting ketamine dissolved in 0.9% saline, at  $10 \text{mg kg}^{-1}$  for mice (Yang et al.). In this investigation, the concentration would be altered, with two groups of mice administered at concentrations of  $8 \text{mg kg}^{-1}$  and  $16 \text{mg kg}^{-1}$ . One hour after injection, the mice can be used for behavioral studies. The mice would be tested with behavioral paradigms once again after 10 hours post-ketamine injection, as was the case for the groups of mice investigated in specific aims 2.1 and 2.2.

### **Expected results and interpretations:**

Under the prediction that higher ketamine concentrations per injection would amplify its long-term antidepressant effects, the group of mice that received the higher concentration ( $16 \text{mg kg}^{-1}$ ) of ketamine would theoretically exhibit less depressive-like symptoms in both behavioral paradigms tested post-injection.

## **5. Future Implications**

Given that the study only investigates the pathway from the LHb to the VTA, future studies investigating the various different projection pathways of LHb neurons would be necessary to gain a full picture of the heterogeneity of LHb neurons. It is plausible that LHb neurons that undertake other projection pathways operate and are stimulated by different factors, as high-throughput single-cell transcriptional profiling used in recent studies (Hashikawa et al.) discovered neuronal cell types in the LHb that are transcriptionally different, which had distinct downstream targets as well; further research is required to clarify this. Furthermore, the study only examines the antidepressant effects of ketamine, as an alternative to traditional SSRIs. The therapeutic effects of other NMDAR antagonists remain largely unknown, which warrants further investigation in future studies, where their antidepressant effects and mechanisms can be compared to those of ketamine.

## **6. Conclusion**

Through the utilisation of optogenetic and electrophysiological techniques, the study would demonstrate the specific effects of different LHb neuron firing modes on the VTA, clarifying the differences between tonic-firing and burst-firing LHb neurons. The results of the study may have significant implications for future research on the aetiology of dementia, as well as potential pharmacological applications in investigating whether the ratio between silent, tonic-firing and burst-firing neurons can be altered as a form of treatment.

This study is also one of the first to propose an in-depth investigation into the mechanisms underlying the long-term therapeutic effects of ketamine, and if the hypotheses are correct, these studies would improve our understanding and provide deeper insight into the antidepressant effects exhibited by ketamine.

## Works Cited

- addgene. "Addgene: Optogenetics Guide." *W*ww.addgene.org, 17 July 2017, [www.addgene.org/guides/optogenetics/](http://www.addgene.org/guides/optogenetics/).
- Autry, Anita E., et al. "NMDA Receptor Blockade at Rest Triggers Rapid Behavioural Antidepressant Responses." *Nature*, vol. 475, no. 7354, June 2011, pp. 91–95, <https://doi.org/10.1038/nature10130>.
- Bekkers, John M. "Pyramidal Neurons." *Current Biology*, vol. 21, no. 24, Dec. 2011, p. R975, <https://doi.org/10.1016/j.cub.2011.10.037>.
- Berman, Robert M., et al. "Antidepressant Effects of Ketamine in Depressed Patients." *Biological Psychiatry*, vol. 47, no. 4, Feb. 2000, pp. 351–54, [https://doi.org/10.1016/s0006-3223\(99\)00230-9](https://doi.org/10.1016/s0006-3223(99)00230-9).
- BIRMAHER, B. "Course and Outcome of Child and Adolescent Major Depressive Disorder." *Child and Adolescent Psychiatric Clinics of North America*, vol. 11, no. 3, July 2002, pp. 619–37, [https://doi.org/10.1016/s1056-4993\(02\)00011-1](https://doi.org/10.1016/s1056-4993(02)00011-1).
- Bray, Natasha. "Bursting with Depression." *Nature Reviews Neuroscience*, vol. 19, no. 4, 2018, pp. 181–81, <https://doi.org/10.1038/nrn.2018.28>.
- Brinschwitz, K., et al. "Glutamatergic Axons from the Lateral Habenula Mainly Terminate on GABAergic Neurons of the Ventral Midbrain." *Neuroscience*, vol. 168, no. 2, June 2010, pp. 463–76, <https://doi.org/10.1016/j.neuroscience.2010.03.050>. Accessed 20 Aug. 2020.
- Cai, Jing, and Qingchun Tong. "Anatomy and Function of Ventral Tegmental Area Glutamate Neurons." *Frontiers in Neural Circuits*, vol. 16, no. 867053, May 2022, <https://doi.org/10.3389/fncir.2022.867053>.
- CDC. "Tetrodotoxin: Biotoxin | NIOSH | CDC." *W*ww.cdc.gov, 25 May 2023, [www.cdc.gov/niosh/ersbdb/emergencyresponsecard\\_29750019.html#:~:text=8%2DHOUSING%20EXPOSURE-](http://www.cdc.gov/niosh/ersbdb/emergencyresponsecard_29750019.html#:~:text=8%2DHOUSING%20EXPOSURE-).
- Cerniauskas, Ignas, et al. "Chronic Stress Induces Activity, Synaptic, and Transcriptional Remodeling of the Lateral Habenula Associated with Deficits in Motivated Behaviors." *Neuron*, vol. 104, no. 5, Dec. 2019, pp. 899-915.e8, <https://doi.org/10.1016/j.neuron.2019.09.005>. Accessed 9 Nov. 2020.
- Chu, Andrew, and Roopma Wadhwa. "Selective Serotonin Reuptake Inhibitors." *PubMed*, StatPearls Publishing, 1 May 2023, [www.ncbi.nlm.nih.gov/books/NBK554406/](http://www.ncbi.nlm.nih.gov/books/NBK554406/). Accessed 11 Aug. 2024.
- Clements, J. A., et al. "Bioavailability, Pharmacokinetics, and Analgesic Activity of Ketamine in Humans." *Journal of Pharmaceutical Sciences*, vol. 71, no. 5, May 1982, pp. 539–42, <https://doi.org/10.1002/jps.2600710516>.
- Cui, Yihui, et al. "Astroglial Kir4.1 in the Lateral Habenula Drives Neuronal Bursts in Depression." *Nature*, vol. 554, no. 7692, Feb. 2018, pp. 323–27, <https://doi.org/10.1038/nature25752>. Accessed 5 Aug. 2024.

- . “Lateral Habenular Burst Firing as a Target of the Rapid Antidepressant Effects of Ketamine.” *Trends in Neurosciences*, vol. 42, no. 3, Mar. 2019, pp. 179–91, <https://doi.org/10.1016/j.tins.2018.12.002>.
- Ferezou, Isabelle, et al. “Spatiotemporal Dynamics of Cortical Sensorimotor Integration in Behaving Mice.” *Neuron*, vol. 56, no. 5, Dec. 2007, pp. 907–23, <https://doi.org/10.1016/j.neuron.2007.10.007>. Accessed 16 Jan. 2020.
- Fox, David M., et al. “Bursting in Neurons and Small Networks.” *Encyclopedia of Computational Neuroscience*, edited by Dieter Jaeger and Ranu Jung, Springer New York, 2013, pp. 1–17, [https://doi.org/10.1007/9781461473206\\_4541](https://doi.org/10.1007/9781461473206_4541).
- Gilmore, Joe. “Most Common Antidepressants.” *Renaissance Recovery*, 26 Dec. 2023, [www.renaissancerecovery.com/most-common-antidepressants/#:~:text=Common%20SSRIs%20include%20Prozac%20\(fluoxetine](http://www.renaissancerecovery.com/most-common-antidepressants/#:~:text=Common%20SSRIs%20include%20Prozac%20(fluoxetine). Accessed 24 Aug. 2024.
- Gonzalez, M. M. C., and G. Aston-Jones. “Light Deprivation Damages Monoamine Neurons and Produces a Depressive Behavioral Phenotype in Rats.” *Proceedings of the National Academy of Sciences*, vol. 105, no. 12, Mar. 2008, pp. 4898–903, <https://doi.org/10.1073/pnas.0703615105>. Accessed 21 Mar. 2022.
- Gradinaru, Viviana, et al. “ENpHR: A Natronomonas Halorhodopsin Enhanced for Optogenetic Applications.” *Brain Cell Biology*, vol. 36, no. 1-4, Aug. 2008, pp. 129–39, <https://doi.org/10.1007/s11068-008-9027-6>. Accessed 28 June 2019.
- GRILLNER, S., et al. “Activation of NMDA-Receptors Elicits ‘Fictive Locomotion’ in Lamprey Spinal Cord in Vitro.” *Acta Physiologica Scandinavica*, vol. 113, no. 4, Dec. 1981, pp. 549–51, <https://doi.org/10.1111/j.1748-1716.1981.tb06937.x>. Accessed 22 Mar. 2022.
- Groos, Dominik, and Fritjof Helmchen. “The Lateral Habenula: A Hub for Value-Guided Behavior.” *Cell Reports*, vol. 43, no. 4, Cell Press, Apr. 2024, pp. 113968–68, <https://doi.org/10.1016/j.celrep.2024.113968>. Accessed 22 Aug. 2024.
- Guru, Akash, et al. “Making Sense of Optogenetics.” *International Journal of Neuropsychopharmacology*, vol. 18, no. 11, July 2015, p. pyv079, <https://doi.org/10.1093/ijnp/pyv079>.
- Hashikawa, Yoshiko, et al. “Transcriptional and Spatial Resolution of Cell Types in the Mammalian Habenula.” *Neuroresource*, vol. 106, no. 5, June 2020, pp. 743-758.e5, <https://doi.org/10.1016/j.neuron.2020.03.011>. Accessed 22 Aug. 2024.
- Hikosaka, Okihide. “The Habenula: From Stress Evasion to Value-Based Decision-Making.” *Nature Reviews Neuroscience*, vol. 11, no. 7, July 2010, pp. 503–13, <https://doi.org/10.1038/nrn2866>.
- Huang, Lu, et al. “A Visual Circuit Related to Habenula Underlies the Antidepressive Effects of Light Therapy.” *Neuron*, vol. 102, no. 1, Feb. 2019, <https://doi.org/10.1016/j.neuron.2019.01.037>.
- Huguenard, J. R., et al. “Transient Ca<sup>2+</sup> Currents in Neurons Isolated from Rat Lateral Habenula.” *Journal of Neurophysiology*, vol. 70, no. 1, American Physiological Society, July 1993, pp. 158–66, <https://doi.org/10.1152/jn.1993.70.1.158>. Accessed 9 Aug. 2024.

- Kalivas, Peter W. “Neurotransmitter Regulation of Dopamine Neurons in the Ventral Tegmental Area.” *Brain Research Reviews*, vol. 18, no. 1, Jan. 1993, pp. 75–113, [https://doi.org/10.1016/0165-0173\(93\)90008-n](https://doi.org/10.1016/0165-0173(93)90008-n).
- Kim, Kyoung-Shim, and Pyung-Lim Han. “Optimization of Chronic Stress Paradigms Using Anxiety- and Depression-like Behavioral Parameters.” *Journal of Neuroscience Research*, vol. 83, no. 3, Feb. 2006, pp. 497–507, <https://doi.org/10.1002/jnr.20754>.
- Knackstedt, Lori A., and Marek Schwendt. *Metabotropic Glutamate Receptors in Psychiatric and Neurological Disorders*. Elsevier, 2023, pp. 221–64.
- Krishnan, Vaishnav, and Eric J. Nestler. “Linking Molecules to Mood: New Insight into the Biology of Depression.” *American Journal of Psychiatry*, vol. 167, no. 11, Nov. 2010, pp. 1305–20, <https://doi.org/10.1176/appi.ajp.2009.10030434>.
- Kunori, N., et al. “Voltage-Sensitive Dye Imaging of Primary Motor Cortex Activity Produced by Ventral Tegmental Area Stimulation.” *Journal of Neuroscience*, vol. 34, no. 26, June 2014, pp. 8894–903, <https://doi.org/10.1523/jneurosci.5286-13.2014>. Accessed 17 May 2020.
- Li, Jicheng, et al. “The Convergence of Aversion and Reward Signals in Individual Neurons of the Mice Lateral Habenula.” *Experimental Neurology*, vol. 339, May 2021, p. 113637, <https://doi.org/10.1016/j.expneurol.2021.113637>. Accessed 14 Apr. 2022.
- Liu, Xianwei, et al. “Burst Firing in Output-Defined Parallel Habenula Circuit Underlies the Antidepressant Effects of Bright Light Treatment.” *Advanced Science*, Wiley, June 2024, <https://doi.org/10.1002/advs.202401059>. Accessed 30 July 2024.
- Ma, Shuangshuang, et al. “Sustained Antidepressant Effect of Ketamine through NMDAR Trapping in the LHb.” *Nature*, vol. 622, no. 7984, Nature Portfolio, Oct. 2023, <https://doi.org/10.1038/s41586-023-06624-1>.
- Maeng, Sungho, et al. “Cellular Mechanisms Underlying the Antidepressant Effects of Ketamine: Role of  $\alpha$ -Amino-3-Hydroxy-5-Methylisoxazole-4-Propionic Acid Receptors.” *Biological Psychiatry*, vol. 63, no. 4, Feb. 2008, pp. 349–52, <https://doi.org/10.1016/j.biopsych.2007.05.028>.
- Matsumoto, Masayuki, and Okihide Hikosaka. “Lateral Habenula as a Source of Negative Reward Signals in Dopamine Neurons.” *Nature*, vol. 447, no. 7148, May 2007, pp. 1111–15, <https://doi.org/10.1038/nature05860>.
- Maxwell, Christina R., et al. “Ketamine Produces Lasting Disruptions in Encoding of Sensory Stimuli.” *Journal of Pharmacology and Experimental Therapeutics*, vol. 316, no. 1, Sept. 2005, pp. 315–24, <https://doi.org/10.1124/jpet.105.091199>.
- Mayo Clinic. “Selective Serotonin Reuptake Inhibitors (SSRIs).” *Mayo Clinic*, 2019, [www.mayoclinic.org/diseases-conditions/depression/in-depth/ssris/art-20044825](http://www.mayoclinic.org/diseases-conditions/depression/in-depth/ssris/art-20044825).
- McPhillips, Deidre. “More than 1 in 6 Adults Have Depression as Rates Rise to Record Levels in the US, Survey Finds.” *CNN*, 17 May 2023, [edition.cnn.com/2023/05/17/health/depression-rates-gallup/index.html](http://edition.cnn.com/2023/05/17/health/depression-rates-gallup/index.html).

- Nikolay Aseyev, et al. “Biolytic Delivery of Voltage-Sensitive Dyes for Fast Recording of Membrane Potential Changes in Individual Neurons in Rat Brain Slices.” *Journal of Neuroscience Methods*, vol. 212, no. 1, Elsevier BV, Jan. 2013, pp. 17–27, <https://doi.org/10.1016/j.jneumeth.2012.09.008>.
- Omelchenko, Natalia, et al. “Lateral Habenula Projections to Dopamine and GABA Neurons in the Rat Ventral Tegmental Area.” *European Journal of Neuroscience*, vol. 30, no. 7, Oct. 2009, pp. 1239–50, <https://doi.org/10.1111/j.1460-9568.2009.06924.x>. Accessed 8 Nov. 2019.
- Pirtle, Thomas J. “A Review of the Circuit-Level and Cellular Mechanisms Contributing to Locomotor Acceleration in the Marine Mollusk *Clione limacina*.” *Frontiers in Neuroscience*, vol. 16, Dec. 2022, <https://doi.org/10.3389/fnins.2022.1072974>.
- Powell, Timothy R., et al. “Depression-Related Behavioral Tests.” *Current Protocols in Mouse Biology*, vol. 2, no. 2, June 2012, pp. 119–27, <https://doi.org/10.1002/9780470942390.mo110176>. Accessed 12 Aug. 2024.
- Proenca, Catia C., et al. “Slitrks as Emerging Candidate Genes Involved in Neuropsychiatric Disorders.” *Trends in Neurosciences*, vol. 34, 3, 2011, pp. 143–53, <https://doi.org/10.1016/j.tins.2011.01.001>.
- Sawchuk, Craig. “Depression (Major Depressive Disorder).” *Mayo Clinic*, Mayo Foundation for Medical Education and Research, 14 Oct. 2022, [www.mayoclinic.org/diseases-conditions/depression/symptoms-causes/syc-20356007](http://www.mayoclinic.org/diseases-conditions/depression/symptoms-causes/syc-20356007).
- Schiller, Jackie, et al. “NMDA Spikes in Basal Dendrites of Cortical Pyramidal Neurons.” *Nature*, vol. 404, no. 6775, Mar. 2000, pp. 285–89, <https://doi.org/10.1038/35005094>. Accessed 20 Mar. 2020.
- Selverston, Allen I., and Maurice Moulins. *The Crustacean Stomatogastric System*. Springer Science & Business Media, 2012.
- Shao, Jie, et al. “Neural Burst Firing and Its Roles in Mental and Neurological Disorders.” *Frontiers in Cellular Neuroscience*, vol. 15, Sept. 2021, <https://doi.org/10.3389/fncel.2021.741292>.
- Shieh, Jennifer C., and Matt Carter. *Guide to Research Techniques in Neuroscience*. Elsevier Science Publishing Co, 2015, pp. 39–71.
- Stamatakis, Alice M., et al. “A Unique Population of Ventral Tegmental Area Neurons Inhibits the Lateral Habenula to Promote Reward.” *Neuron*, vol. 80, no. 4, Nov. 2013, pp. 1039–53, <https://doi.org/10.1016/j.neuron.2013.08.023>.
- Sun, Yanfei, et al. “Rostromedial Tegmental Nucleus-Substantia Nigra Pars Compacta Circuit Mediates Aversive and Despair Behavior in Mice.” *Experimental Neurology*, vol. 333, Nov. 2020, p. 113433, <https://doi.org/10.1016/j.expneurol.2020.113433>. Accessed 18 Dec. 2023.
- Vroomen Durning, Marijke. “Drug Treatment for Depression.” *WebMD*, 7 June 2024, [www.webmd.com/depression/drugs-treat-depression](http://www.webmd.com/depression/drugs-treat-depression). Accessed 5 Aug. 2024.



- WHO. “‘Depression: Let’s Talk’ Says WHO, as Depression Tops List of Causes of Ill Health.” *Www.who.int*, 30 Mar. 2017, [www.who.int/news/item/30-03-2017--depression-let-s-talk-says-who-as-depression-tops-list-of-causes-of-ill-health](http://www.who.int/news/item/30-03-2017--depression-let-s-talk-says-who-as-depression-tops-list-of-causes-of-ill-health).
- Williams, Alex H., et al. “The Neuromuscular Transform of the Lobster Cardiac System Explains the Opposing Effects of a Neuromodulator on Muscle Output.” *Journal of Neuroscience*, vol. 33, no. 42, Society for Neuroscience, Oct. 2013, pp. 16565–75, <https://doi.org/10.1523/jneurosci.2903-13.2013>. Accessed 5 Aug. 2024.
- Yang, Yan, et al. “Ketamine Blocks Bursting in the Lateral Habenula to Rapidly Relieve Depression.” *Nature*, vol. 554, no. 7692, Feb. 2018, pp. 317–22, <https://doi.org/10.1038/nature25509>.
- Young, Callum J., et al. “Circadian Influences on the Habenula and Their Potential Contribution to Neuropsychiatric Disorders.” *Frontiers in Behavioral Neuroscience*, vol. 15, Frontiers Media, Jan. 2022, <https://doi.org/10.3389/fnbeh.2021.815700>. Accessed 21 Aug. 2024.
- Younts, T. J., et al. “CA1 Pyramidal Cell Theta-Burst Firing Triggers Endocannabinoid-Mediated Long-Term Depression at Both Somatic and Dendritic Inhibitory Synapses.” *Journal of Neuroscience*, vol. 33, no. 34, Aug. 2013, pp. 13743–57, <https://doi.org/10.1523/jneurosci.0817-13.2013>. Accessed 26 Nov. 2021.
- Zhang, Yun-Feng, et al. “Alterations of Motor Cortical Microcircuit in a Depressive-like Mouse Model Produced by Light Deprivation.” *Neuroscience*, vol. 341, Elsevier BV, Jan. 2017, pp. 79–94, <https://doi.org/10.1016/j.neuroscience.2016.11.026>. Accessed 15 Apr. 2024.
- ZHU, Z., et al. “NMDA Enhances a Depolarization-Activated Inward Current in Subthalamic Neurons.” *Neuropharmacology*, vol. 49, no. 3, Sept. 2005, pp. 317–27, <https://doi.org/10.1016/j.neuropharm.2005.03.018>. Accessed 10 Jan. 2023.

# **The Influence of Digital Platforms on Hindu Spiritual Guidance: A Comparative Study of Online and Real-Life Gurus By Karan Shroff**

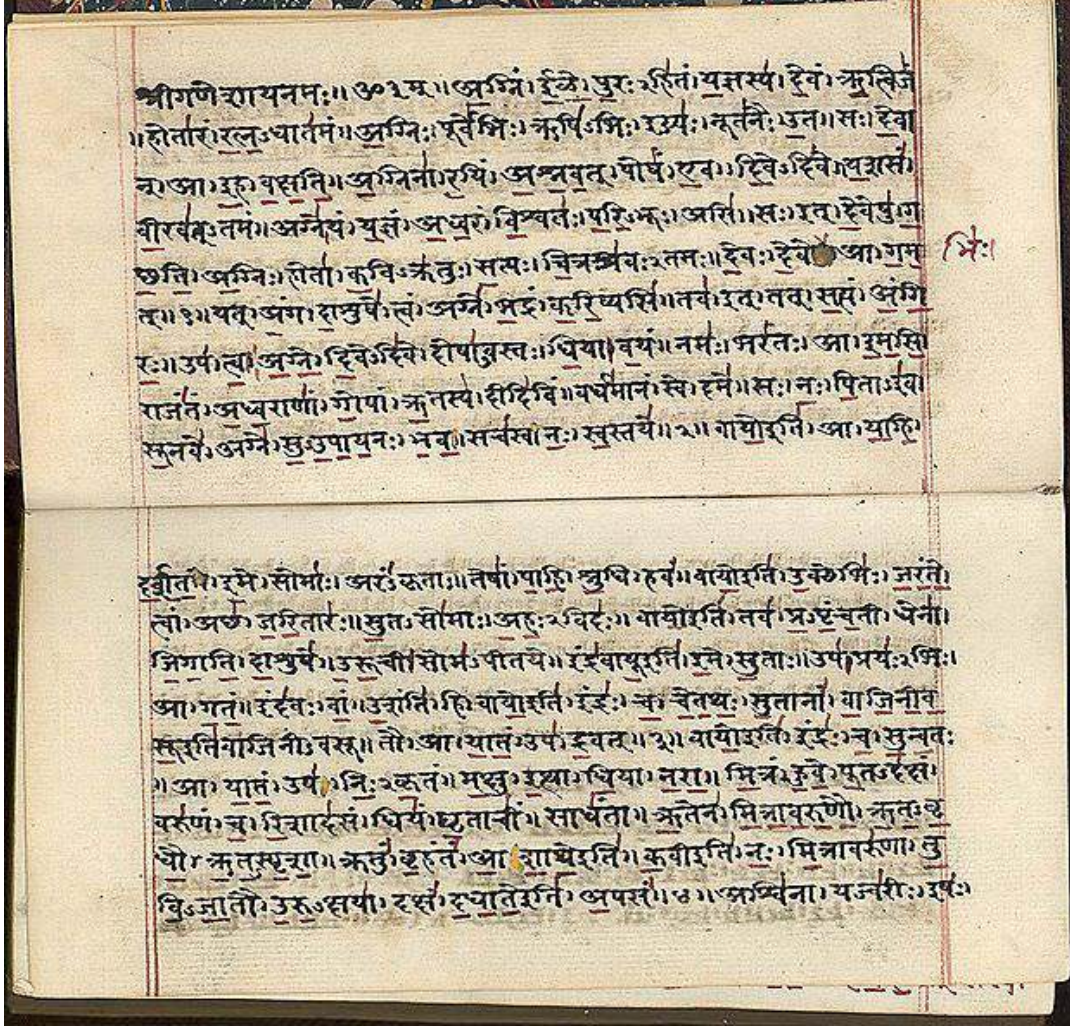
## **0. Introduction**

In the Hindu religion, the Guru serves as a spiritual guide for followers. This research paper explores the influence of the digital age on Hindu religious counseling. Specifically, the current difference between modern (online) and traditional Gurus.

## **1.0 Background**

Hinduism, considered by many the oldest religion in the world, dates back over 4,000 years. Many scholars argue that the polytheistic religion started in the Indus River Valley, one of the five major cultural hearths of the world, somewhere between 2300 BC and 1500 BC. However, many Hindus believe that their religion is even older, as argued by Mrs. Anupam, a local Hinduism and Hindi teacher in the UAE.

The original Hindu teachings were encapsulated by the Vedas, a collection of four scriptures (the Rigveda, Samaveda, Yajurveda, and Atharvaveda) dated by most scholars to be written around 1500 - 1200 BC. These sacred texts were passed down orally through multiple generations before finally being put to paper. Originally meaning “knowledge” in Sanskrit, the language it was written in, the Vedas contained hymns, philosophical discussions, and praise for a pantheon of Gods. In addition to that, the texts emphasized the importance of Gurus, people who are seen as spiritual guides and teachers. Following the Vedas, many other important religious scriptures and epics such as the Bhagavad Gita, Ramayana, and Mahabharata were put in black and white, expanding the breadth of the Hindu faith.



An extract from the Rig-Veda, written in Sanskrit (early 19th century CE)

For Hindus, more than most other religions, one's own religious beliefs and version of the truth is heavily influenced by where one was born, what scriptures one identifies with and what Gods they predominantly pray to. These differences are by nature of Hinduism's polytheistic system which contains thousands of deities and demi-Gods.

Nevertheless, the core beliefs of the religion stay relatively the same among all its followers: That is, the cycle of reincarnation and the eventual return to God. To explain more thoroughly, central to Hindu religious philosophy is the Sanskrit term Dharma. Essentially, Dharma refers to a set of religious duties that people can follow to live better and more meaningful lives; this ethical code attaches great importance to values such as truth, right conduct, love, peace and non-violence. Moreover, each person's Dharmic duties are individualized, with the path for every living soul being different.

Your accordance to your Dharma is then used to help decide your next life in the reincarnation cycle. This process represents the important Hindu principle of Karma. Essentially, if you follow your duties in this life, God will reward correspondingly in your next life and vice

versa. This would come in the form of you being born into a higher or lower class family, or as a higher or lower species. This cycle of rebirth, known as Samsara in Sanskrit, finally ends with the eventual re-joining with God, known as Moksha.

The return to God is a return to Brahman, the supreme being in the Hindu faith. Interestingly, Hindu deities are different manifestations of the Brahman. This is considering the fact that in its raw form Brahman has no shape or form, making it harder for humans to understand. Therefore, through distinct forms of Hindu idols, Hindus can learn about pieces of Brahman in a way that they can understand. For example, the God Shiva is depicted as holding a trident, which symbolizes destruction in the Hindu faith and thereby in Brahman. For that reason, the view that Hinduism is a polytheistic religion is not a complete one, owing to the fact that in reality it is based in one supreme being.

Now the basics of Hindu thought are clear, it begs the question of the Guru's role in the grand scheme of the religion.

"Guru is Shiva sans his three eyes,  
Vishnu sans his four arms  
Brahma sans his four heads.  
He is parama Shiva himself in human form"  
~ Brahmanda Purana

The Puranas are a number of Hindu texts that primarily focus on Hindu lore and legends. Literally translating to old/ancient, these scriptures date back to the 3rd century AD. The excerpt above, from one of the oldest Puranas, perfectly illustrates how the Guru is seen in Hindu religious philosophy. It compares the Guru to a God, saying that they are God in human form. In other words, Hinduism sees the Guru being as essential as the many Gods. The significance of the Guru is in view of the fact that they are a mentor that comes to you in the flesh. Just as God helps you from the heavens, a Guru imparts knowledge on you to live a better life. To put it simply, a Guru is a spiritual teacher, a guide.

The word Guru, along with other terms derived from Hindu tradition, has since expanded and is now used more loosely in everyday language. For example, someone would say that their life coach is their Guru, and a Tony Robinson quote is their daily mantra.

As technology continues to grow through the centuries, human perspectives have changed, the flow of information has increased, and the number of Hindu followers worldwide has grown. Consequently, the role of Gurus has expanded and changed. Now more than ever, people now have the ability to tap into an almost infinite pool of information. As a result, a Hindu is now able to compare the advice of multiple Gurus from around the world on a singular issue. This is compared to mere hundreds of years ago, where there would only be one spiritual guide you would be able to access. This change prompts inquiry into how the role of the Guru has shifted in the digital age, and what the actual differences are between traditional and modern spiritual guides.

## 2.0 Traditional Spiritual Guidance

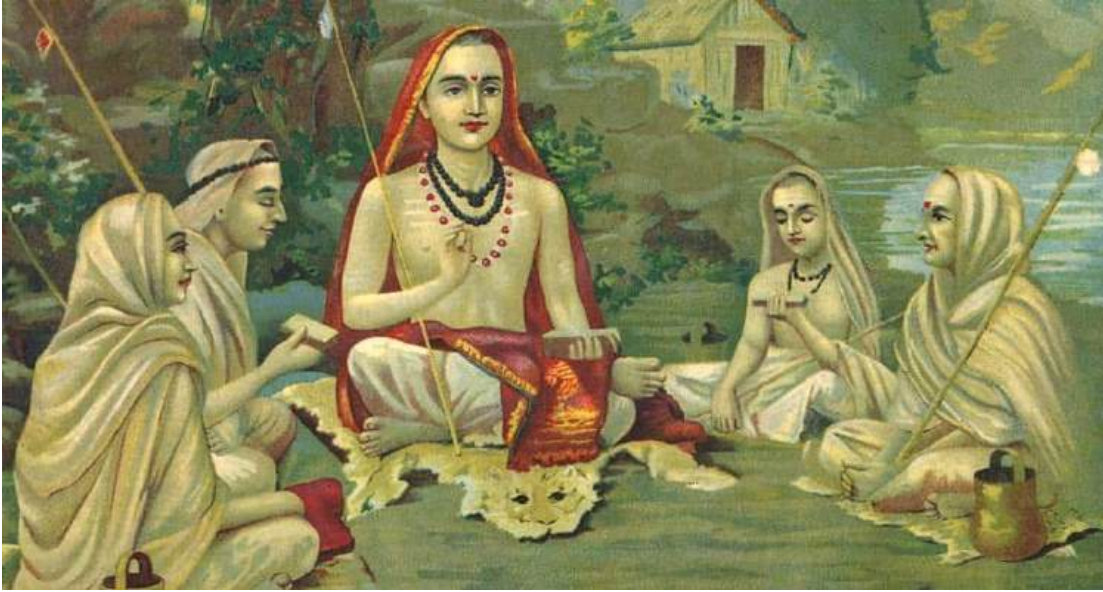
Guruship is as old as the Hindu religion itself. Traditionally, a Guru would live a simple married life and accept students to teach. Although there is scriptural guidance for the qualifications for good students, the ultimate decision lies with the Guru himself.

A student's journey would begin by the new disciple offering a gift to his new teacher to show their gratitude for years of education to come. The student would then move into the household of the Guru, called a Gurukula in Sanskrit. This could either be the personal residence of the Guru, the temple where he resides, but most often a hut in the woods. At the Gurukula, the student would be taught basic Vedic principles and study sacred scriptures. Once a student's teaching is complete, a Guru may ask his student for a Dakshina, a gift for teaching them. However, this does not repay the enormous debt the student has for his Guru, as he is obligated to spread his newfound knowledge to younger generations.

“When the teacher finds from signs that knowledge has not been grasped or has been wrongly grasped by the student, he should remove the causes of non-comprehension in the student. This includes the student's past and present knowledge, want of previous knowledge of what constitutes subjects of discrimination and rules of reasoning, behavior such as unrestrained conduct and speech, courting popularity, vanity of his parentage, [and] ethical flaws that are means contrary to those causes. The teacher must enjoin means in the student that are enjoined by the Śruti and Smṛti, such as avoidance of anger, Yamas consisting of Ahimsa and others, also the rules of conduct that are not inconsistent with knowledge. [The teacher] should also thoroughly impress upon the student qualities like humility, which are the means to knowledge.”

—Adi Shankara (788-820),  
Upadesha Sahasri 1.4-1.5

The life of Adi Shankara, an Indian Vedic Scholar and Teacher that championed the philosophy of Advaita Vedanta (a viewpoint which emphasizes the identity shared between Atman and Brahman), can help us better understand the life of a Guru. Shankara was born into a poor Brahman family and lost his father at an early age. With his mother as his teacher, he developed a brilliant mind, speaking fluent Sanskrit by the age of two and memorizing all four Vedas by age of four. By his twelfth birthday, he had left home to live an ascetic lifestyle. He then began searching for a Guru and landed up becoming a student of Govinda Bhagavatpada. There, Shankara further studied ancient Hindu scriptures such as the Vedas, Upanishads, Brahma Sutras, and more. At the age of thirty-two, he achieved Maha Samadhi, an experience in death where an enlightened one consciously releases themselves from the cycle of reincarnation. Before his death, he walked the Indian subcontinent spreading his knowledge as well as writing several books and commentaries.



Adi Shankara (788-820) with Students

Now that the role of a traditional Guru is clear, it raises the question of why one would even seek out a Guru when ample scriptures are available. To answer this question, we go back to the notion previously mentioned that a Guru is comparable to a God, but rather teaches you in the flesh. In other words, a Guru can help a student focus on and work towards their own individualized Dharma, whereas a text gives all round advice. With that importance in mind, a Guru has the respect and full devotion of his students.

An extreme example of a student's devotion for his Guru is exemplified in the story of Eklavya and Dronacharya present in the Mahabharata, one of the most important Hindu epics. The story begins long ago in Hastinapur (a city in northern India), where there lived a son of a local tribal chief named Eklavya. Wanting to learn archery from the renowned royal guru Dronacharya, Eklavya journeyed to his Gurukula in the nearby forest. When he arrived, Eklavya was disappointed to hear that Dronacharya could not accept him as a pupil because, according to state law, he could only teach royals. Eklavya, his desire to master archery still steadfast, took to the surrounding forest and sculpted a mud statue of Guru Drona, taking its blessing and practicing archery everyday. After a few years, Eklavya had become a master. One day while practicing, Eklavya was disturbed by a dog barking. In order to stay focused, he managed to shoot arrows into the dog's mouth so the dog could not bark but was still unharmed. Coming across the dog later in the day, Guru Drona knew it was the work of a master archer and set out to find him. After finding Eklavya, Guru Drona became perplexed by his skill and asked the young man who taught him. Eklavya touched Dronacharya's feet and said, "You, Sir. You are my guru. ... I have been practicing archery religiously in front of your statue [after you said you could not teach me]." Dronacharya was shocked yet proud at the same time and knew this would put him in a difficult position as a royal teacher. As his Guru then, Drona asked Eklavya for his Dakshina, Eklavya's right thumb, knowing that it would make it almost impossible for him to

pursue archery in the future. Without hesitation, Eklavya cut it off and offered his thumb with gratitude. Touched by Eklavya's dedication, Dronacharya blessed him to become the most gifted archer and to be a model of devotion. All in all, this story shows the faith that disciples have in their Gurus as well as a Guru's ability to bless their students, even if a student is not taught face to face.

Traditional Gurus still exist in modern day. However, these spiritual guides often exist in rural settings where access to the internet is limited. Now, most often, Gurus as well as Hindu religious institutions have integrated modern online networks and technology to increase their audiences.

### **3.0 Modern Spiritual Guidance**

As modern technology has expanded, so has the scale of Hindu mentorship. Because of the internet and electronic devices, hundreds of thousands of people can learn from a Guru's lecture instead of a handful in his Gurukula. However, despite what may look different on the outside, at their core, a Guru's primary goal has remained the same.

The notion that Gurus are still teachers, from the first centuries of the Hindu religion and continuing into the modern day, is essential in understanding the reason for how the way they operate has evolved. This is the view of the fact that a Guru's purpose has remained the same: to impart knowledge that they deem necessary for their students. Moreover, for that reason, Gurus have to speak in a language that their disciples understand and appreciate.

This idea of shaping your language to your audience is significant because the audience of many Gurus has expanded and changed remarkably in the last century alone. If the same rhetoric used to teach back in the first century AD was used now, the idea of a Guru would be dead completely. Therefore, Gurus have to change their language to fit a modern audience, speaking with modern terminology and about issues that affect the current day populace. Moreover, as Hinduism spreads globally and expands throughout the world, the largest Gurus have to use a language and impart teachings that are universally applicable.

In addition to broadening their teachings, Gurus have also broadened their ventures. Instead of being a one man show in the woods, Gurus are now the leads of multi national charitable organizations with hundreds of staff members that promote Hindu religious teachings. In other words, from a macro scale, the biggest Gurus look more like brand marks of transnational organizations than spiritual guides.

In addition to resembling companies, Gurus have to sell themselves as a Fortune 500 Company has to sell their products. In ancient India, the journey to find another Guru would last weeks or even months, as most would have to walk or ride horses to their destination. In other words, local Gurus had limited competition, and were high in demand. This is starkly contrasted to the modern day, where a student can choose between thousands of Gurus who have lectured online. Therefore, in order to become successful and pull audiences, modern Gurus have to sell themselves to their audience and market themselves as unique and useful products: this could be through a philosophy, product, charisma or certain method. As Amanda Lucia would put it,

“Contemporary global gurus are some of the most vibrant innovators in the field of Hindu religiosity” (Innovative Gurus: Tradition and Change in Contemporary Hinduism). In order to keep their followers, like an influencer, Gurus need to continuously innovate and bring new lessons to their followers.

The Mohan Ji Foundation and its founder Mohan Ji serves as a prime example of a contemporary Guru’s teachings and organization. Mohan initially led a happy life. However, after his daughter's death in 2000, he separated from his wife, his investments went sour, and he even lost his job. This prompted him to spend time in the remote spaces meditating the Himalayas, and in the wake of his return in 2003 he founded Ammucare Charitable Trust without any investments or compulsory payment. Ten years after his return from the mountains, he was finally able to find silence in the noise of life. Currently, he is head of Mohan Ji Foundation among other ventures teaching the philosophy that “spirituality is a lifestyle” (Mohan Ji Foundation). As part of the foundation, Mohan Ji charitable ventures include building places of worship across the globe, feeding those who cannot afford it, and planting fruit trees (that help the planet and feed people). Moreover, he travels around the world imparting knowledge on his followers across the globe. In 2024 alone, he traveled to Europe, Africa & North America.

1. Accept people (Don’t Judge)
2. Be positive
3. No agenda (Detachment)
4. Be Purposeful
5. Living insignificance (Selflessness)
6. Accept situations
7. Attitude of Service
8. Expect Nothing
9. No Binding [Agreements]
10. Feed the Hungry

#### The Ten Commandments of Being Mohanji

Unlike the deep scriptural study of the past, Guru Mohan Ji preaches implementing basic Hindu principles (as illustrated above) into daily life. His morning prayer includes gratefulness for another day and questioning what he can do for the world that day.

Mohan Ji lectures in English and teaches people across the globe. However, most of his followers hail from the western world. Therefore, through the focus on spirituality (a growing trend in the countries he is targeting) and using a more familiar language, Mohan Ji is able to capture his target audience. In addition to that, through unique teachings, charitable foundations, and a long history, he is able to advertise himself as a unique spiritual guide and trusted source of information. These ideas exemplify the previously mentioned ways of garnering an audience, because without them a Guru gets lost in the surrounding voices.





Mohan Ji at a New Fruit Tree Plantation

Although they might look like multinational conglomerates from an outside viewpoint, the foundations built by contemporary Gurus are far from that. They are charitable institutions deeply rooted in Hindu philosophy lead by Gurus speak on more generalized teachings of the religion. Essentially, as technology has grown and the world has become a smaller place, Gurus have adapted to take advantage of their new available resources and brought their teachings to a larger audience.

#### **4.0 Comparison and Conclusion**

Despite the change in how our societies operate and function, the role of the Guru in the Hindu religion has remained the same through the centuries: to teach. However, much has changed in the way Gurus guide their disciples, as the common language has changed and audiences have expanded tenfold.

Traditionally, Gurus teach their disciples in their homes, known as Gurukulas in Sanskrit. Moreover, they only teach a handful of students at a time, with full control on who to accept as disciples and how much to teach them. From there, students dive deep into Hindu philosophical teaching by reading scriptures, meditation, and other forms of lessons. Once their journey is complete, students have the obligation to spread the knowledge they've learned to others and create a good name for the Guru who taught them.

Conversely, because of the expansion of technology and decrease of time information takes to travel, Gurus no longer are constrained to teaching those in their homes. Through the utilization of the internet, Hindu spiritual guides have truly become global, and have started charitable organizations that have come to represent multinational corporations. Moreover, just like companies listed on the NASDAQ, Gurus have competition and have to advertise their spiritual services to students. This is in view of the fact that with the internet, students now can

choose to follow an almost innumerable amount of Gurus from the comfort of their home. Therefore, Gurus have become the largest innovators in Hindu religiosity by trying to demonstrate their uniqueness to potential disciples. Now, they have to adapt their language and teachings to best fit their audience, as to grasp and keep their attention, especially in a world full of other distractions.

Nevertheless, contemporary Gurus are far from the companies they resemble. They run charitable organizations that encourage positive change instead of companies that only strive to extract more profit from their clients.

In essence, contemporary Gurus are just a continuation of the Gurus of the past. As technology has evolved, these spiritual guides have learned to take advantage of their new resources to help cultivate a larger audience and promote the religion globally.

5.0 Bibliography (APA)

## Works Cited

- Bikram. (2022, November 18). Adi Shankaracharya – The greatest teacher of Advaita Vedanta. VedicFeed. <https://vedicfeed.com/adi-shankaracharya/>
- Britannica, T. Editors of Encyclopaedia (2024, June 12). guru. Encyclopedia Britannica. <https://www.britannica.com/topic/guru-Hinduism>
- Das, S. (2019, March 8). The significance of the guru. Learn Religions.
- Doniger, W. (2024, June 21). Veda. Encyclopedia Britannica. <https://www.britannica.com/topic/Veda>
- Doniger, W. (2024, June 7). Purana. Encyclopedia Britannica. <https://www.britannica.com/topic/Purana>
- Gold, A. G. , Basham, . Arthur Llewellyn , Buitenen, . J.A.B. van , Dimock, . Edward C. , Doniger, . Wendy , Smith, . Brian K. and Narayanan, . Vasudha (2024, June 19). Hinduism. Encyclopedia Britannica. <https://www.britannica.com/topic/Hinduism>
- Hinduism - origins, facts & beliefs. (2023, November 16). History. <https://www.history.com/topics/religion/hinduism>
- Kamath, R. (n.d.). The ten commandments of being Mohanji. Mohanji. <https://mohanjichronicles.com/mohanji-blog/the-ten-commandments-of-being-mohanji/#:~:text=Mohanji%20has%20zero%20expectations%20from,doership>
- Lucia, A. (2014). Innovative Gurus: Tradition and Change in Contemporary Hinduism. *International Journal of Hindu Studies*, 18(2), 221–263. <http://www.jstor.org/stable/24713672>
- M, B. (2013, January 16). The Vedas (Rig-veda). *World History Encyclopedia*. <https://www.worldhistory.org/image/1027/the-vedas-rig-veda/>
- Mahasamadhi. (2023, December 21). Yogapedia. <https://www.yogapedia.com/definition/5825/mahasamadhi>
- Menon, S. (n.d.). Advaita Vedanta. *Internet Encyclopedia of Philosophy*. <https://iep.utm.edu/advaita-vedanta/>
- The nature of God and existence in Hinduism. (n.d.). BBC. <https://www.bbc.co.uk/bitesize/guides/zmtj2nb/revision/1#:~:text=The%20nature%20of%20God%20and%20existence%20in%20HinduismBrahman,deities%20%2D%20presentations%20of%20the%20divine>
- Ruchika. (2022, September 12). The story of Eklavya and Dronacharya for kids. First Cry. <https://www.firstcry.com/intelli/articles/the-story-of-eklavya-and-dronacharya-for-kids/#:~:text=I%20give%20you%3F-,%27,teaching%20me%20all%20this%20while>
- Temple Purohit. (2015, July 27). The guru shishya parampara. Temple Purohit.
- Who is Mohanji. (n.d.). Mohanji. <https://mohanji.org/who-is-mohanji/>

## Study of Statistics of Solar Flares By Rania Shekhar

### 1. Introduction

Solar flares are intense and dynamic eruptions that occur on the Sun's chromosphere, characterised by sudden and rapid developments that span from minutes to several hours. These eruptions typically manifest as localised brightening on the solar surface, covering relatively small regions [1].

Classifying solar flares is essential for understanding their varying intensities and potential impacts on space weather. They are categorised based on their X-ray output, with classifications ranging from A to X, indicating escalating intensity levels. This classification system is pivotal for assessing the potential disturbances solar flares might cause to Earth's magnetic environment and space weather conditions [2].

The influence of solar flares on space weather is profound and far-reaching. These energetic eruptions emit high-energy particles and electromagnetic radiation that can disrupt satellite operations, communication networks, and navigation systems. Additionally, solar flares play a crucial role in modulating solar wind and amplifying geomagnetic storms, highlighting the importance of comprehensive studies on these phenomena [3].

Earth, situated within the heliosphere, experiences the effects of solar flares directly. The interaction between the energetic particles emitted during solar flares and Earth's magnetic field leads to the mesmerising auroras seen at high latitudes. However, these interactions also pose significant risks to terrestrial infrastructures, including power grids, aviation systems, and satellite communications [4].

Furthermore, solar flares' emission properties provide valuable insights into the underlying physical processes driving these phenomena. Radio emissions from solar flares offer diagnostic tools that help address fundamental questions about energy release, plasma heating, particle acceleration, and transport mechanisms in magnetised plasmas. Electrons with energies ranging from tens of kilo electron volts to several mega electron volts play a pivotal role in carrying a substantial portion of the energy released during the impulsive phase of solar flares [5]. Coherent plasma radiation at decimetre and metre wavelengths also contributes significantly to the emission processes during these events. Flare-like processes may be responsible for large-scale restructuring of the magnetic field in the corona as well as for its heating. Large flares influence interplanetary space and substantially affect the Earth's ionosphere. Flare scenarios have slowly converged over the past decades, but every new observation still reveals major unexpected results, demonstrating that solar flares, after 150 years since their discovery, remain a complex problem of astrophysics including major unsolved questions. [6]

The periodic nature of solar activity has been studied using parameters such as the sunspot Wolf numbers, calcium plage areas and flare indices. The magnitude of the solar activity based on these parameters reveals periodicities other than the most pronounced 11-yr one. Any absolute detection of periodicity in active phenomena would have fundamental significance for our understanding of solar activity. [7] The Earth's magnetosphere is the outermost layer of the

Geospace system deflecting energetic charged particles from the Sun and solar wind. The solar wind has major impacts on the Earth's magnetosphere, but it is unclear whether the same holds for solar flares—a sudden eruption of electromagnetic radiation on the Sun [8]

Sixty-three solar particle events with proton fluence ( $E > 10$  MeV)  $> 107$  cm<sup>-2</sup> were identified from October 1972 to March 1987. The average omnidirectional flux of protons with kinetic energy  $> 10$  MeV for cycle 21, 64 cm<sup>-2</sup>s<sup>-1</sup>, is lower than the corresponding number for cycle 20 (92 cm<sup>-2</sup>s<sup>-1</sup>) based on satellite data and for the cycle 19 (378 cm<sup>-2</sup>s<sup>-1</sup>) based on lunar sample data. Six large events contributed 70% of the total proton fluence during solar cycle 21. [9] According to the electrodynamical model of the solar flare [10], the flare energy is released in a current sheet above an active region. The protons can gain relativistic energy during fast plasma inflow into a current sheet. The Lorenz electric field  $-V \times B/c + j/\sigma$  is applied along a magnetic X-type singular line. Here  $V$  is the inflow velocity at magnetic reconnection, and  $B$  is the magnetic field of the current sheet.

"Type I" confined flares are characterised by slipping reconnection, strong shear, and a stable filament. "Type II" flares have almost no slipping reconnection and have a configuration in potential state after the flare. [11]

One of the largest solar hard X-ray (HXR) flares and solar energetic particle (SEP) events recorded by the *Mars Odyssey* mission while orbiting Mars occurred on 2002 October 27 and is related to a very fast ( $\sim 2300$  km s<sup>-1</sup>) coronal mass ejection (CME). [12] From the Earth, the flare site is  $40.4^\circ \pm 3.5^\circ$  behind the solar limb and only emissions from the high corona at least  $1.5 \times 10^5$  km radially above the main flare site can be seen. Nevertheless, the Earth-orbiting *Reuven Ramaty High Energy Solar Spectroscopic Imager (RHESSI)* observed HXR emission up to 60 keV with a relatively flat, nonthermal spectrum ( $\gamma$  between 3 and 3.5) that has an onset simultaneous with the main HXR emission observed above 60 keV by the Gamma-Ray Spectrometer (GRS) orbiting Mars. [13] While GRS records several smaller enhancements after the main peak, the high coronal source observed by *RHESSI* shows a long exponential decay ( $\tau = 135 \pm 5$  s) with progressive spectral hardening. The emissions from the high corona originate from an extended source ( $\sim 1.5 \times 10^5$  km in diameter) that expands ( $390 \pm 70$  km s<sup>-1</sup>) and moves upwards ( $750 \pm 80$  km s<sup>-1</sup>) in the same direction as the CME. [14] These observations reveal the existence of energetic electrons in the high corona in closed magnetic structures related to the CME that are accelerated at the same time as the main energy release in the flare. Although the number of energetic electrons in the high corona is only a small fraction of the total accelerated electrons, about 10% of all electrons in the high coronal source are nonthermal ( $>10$  keV). [15] Studies of solar flare energetics have shown that a significant portion of the flare energy is often released in the form of accelerated particles, which can then drive secondary phenomena such as coronal mass ejections (CMEs). [16]

The influence of solar flares extends beyond the immediate vicinity of the Sun. When Earth-directed, the high-energy particles and radiation from flares can disrupt the ionosphere, leading to radio signal degradation and navigation system errors [17]. Additionally, the interaction of flare-associated CMEs with the Earth's magnetosphere can trigger geomagnetic

storms, posing risks to power grids and increasing radiation exposure for airline passengers and astronauts [18].

Despite significant progress, many questions about solar flares remain unanswered. The precise conditions that trigger magnetic reconnection, the detailed particle acceleration mechanisms, and the interaction between flares and CMEs are active areas of research [19]. Advancing our understanding of these processes is essential for developing predictive models that can forecast solar flare activity and its potential impacts on Earth [20]. Given the intricate nature of solar flares and their potential impacts on both space weather and terrestrial systems, this project aims to investigate the patterns and time intervals between solar flare occurrences. By examining these phenomena in depth, we seek to enhance our understanding of solar dynamics and their implications for Earth and space-based technologies.

## 2. Methodology

In our solar flare research project, we initiated by accessing data from the Solar Dynamics Observatory (SDO) website, focusing on various time intervals and intensity levels. This data was systematically organised and downloaded into a CSV format, facilitating subsequent analysis through Python programming.

Following data extraction, we constructed an Intensity-Time bar graph to visualise the temporal distribution of solar flares. This graph provided a comprehensive overview of flare occurrences over time, highlighting periods of heightened activity and variations in intensity.

To further refine our analysis, we applied a threshold frequency criterion, considering only flares with intensities exceeding 0.5. This enabled us to focus on significant flare events and discern patterns of increased intensity. By mapping these higher intensity solar flares, we gained valuable insights into the dynamics of solar activity and its potential implications for space weather.

Overall, this systematic approach facilitated the exploration and analysis of solar flare data, laying the groundwork for deeper insights into the behaviour and characteristics of these dynamic solar phenomena.

The following is the explanation of the code used:

- This part of the code imports the necessary libraries: `matplotlib.pyplot` for plotting and `pandas` for data manipulation. The code then reads a CSV file named 'Solar flare things copy.numbers' located at the specified path into a `DataFrame` (`df`). The `print(df)` statement is used to display the contents of the `DataFrame`.
- Here, two lists `x` and `y` are defined, which represent the x and y-axis values for the plot.
- This code segment plots the data points from the lists `x` and `y` on a graph. The plot is styled with a green dashed line (`linestyle='dashed'`) and circular markers (`marker='o'`)

with blue marker faces (`markerfacecolor='blue'`). The line width is set to 3, and the marker size is set to 12.

- These lines set the range for the y-axis (`ylim`) and x-axis (`xlim`) to display values between 1 to 8.
- These statements label the x-axis as 'x - axis' and the y-axis as 'y - axis'.
- This line adds a title 'Solar flares' to the plot.
- 

In summary, this code reads data from a CSV file, plots the data points on a graph with specific styling and labels, and then displays the plot. Afterwards I added restrictions on the graph so it would display specific regions of the graph to be analysed.

### 3. Result and Analysis

The analysis of data of the solar flare graphs show that solar flares are sporadic with an unpredictable pattern. This was clearly represented in the graphs that mapped the frequency and intensity of the solar flares over time.

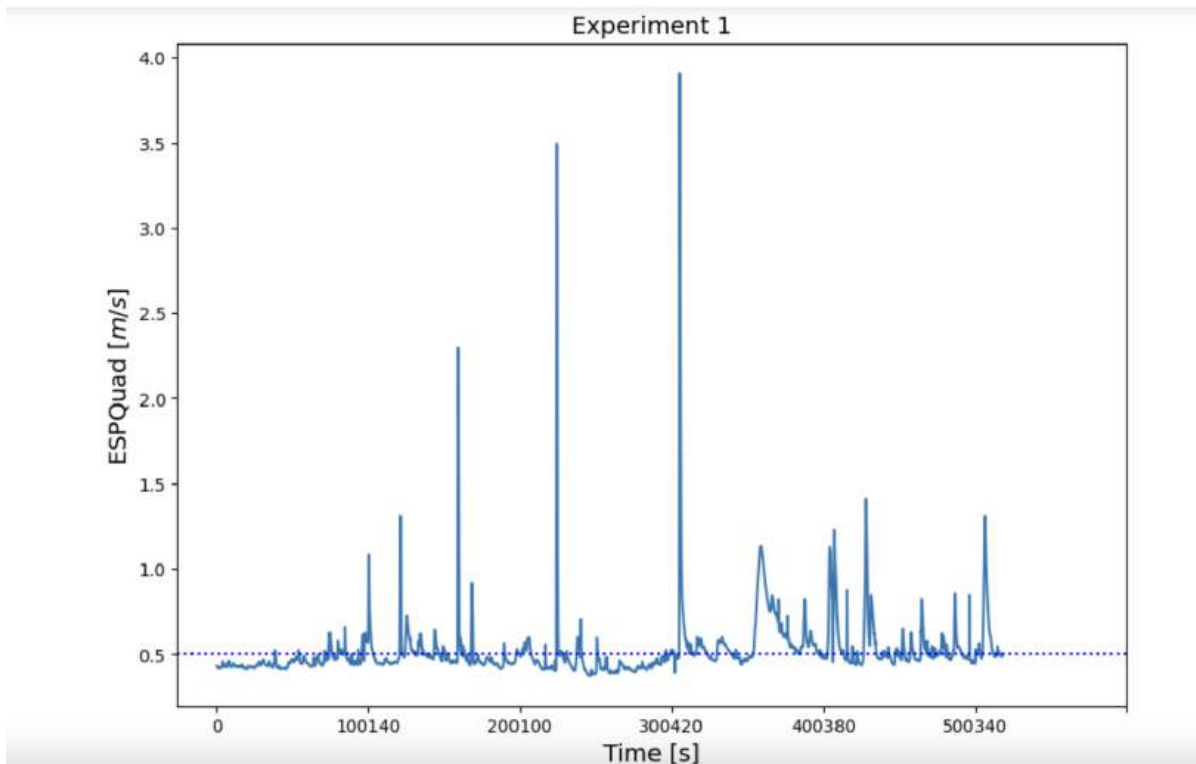


Fig. 1. Graph plot mapping the intensity of solar flares over a period of time with threshold frequency 0.5 (September 2023-December 2023)

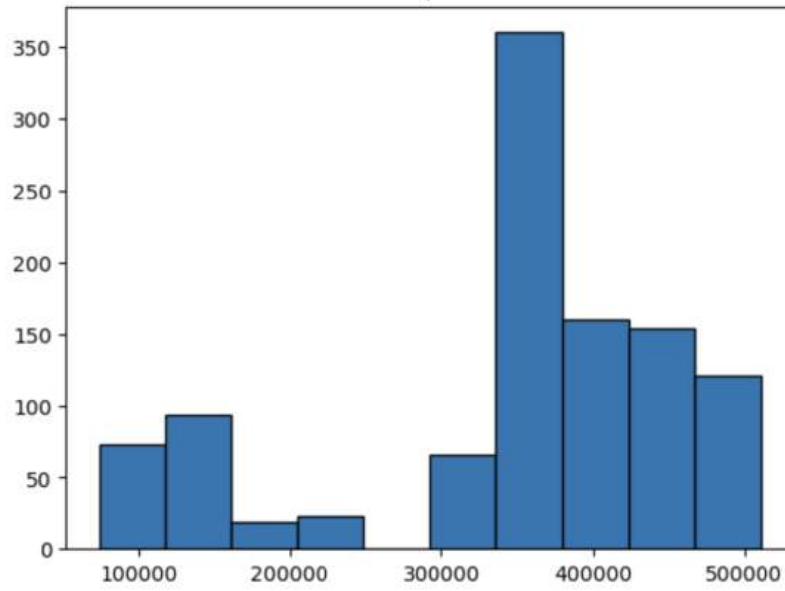


Fig. 2. Graph plot of the solar flare over 0.6 threshold frequency

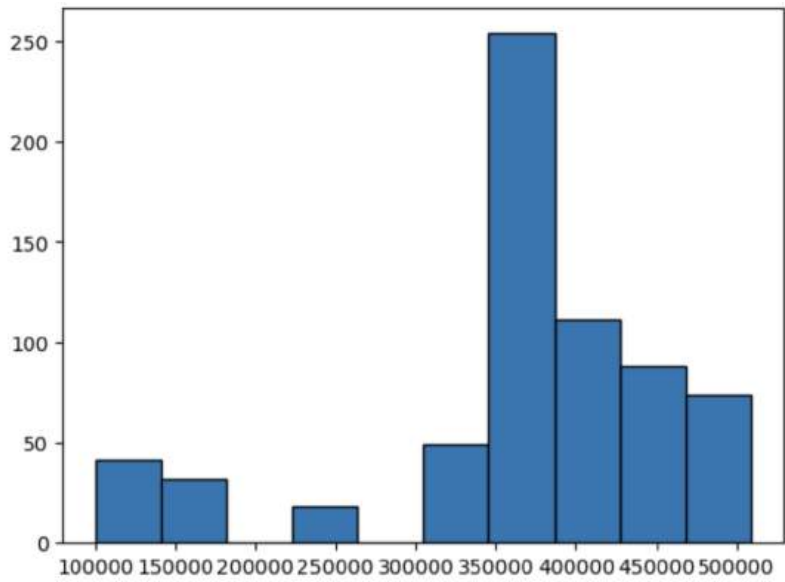


Fig. 3. Graph plot of the solar flare over 0.7 threshold frequency



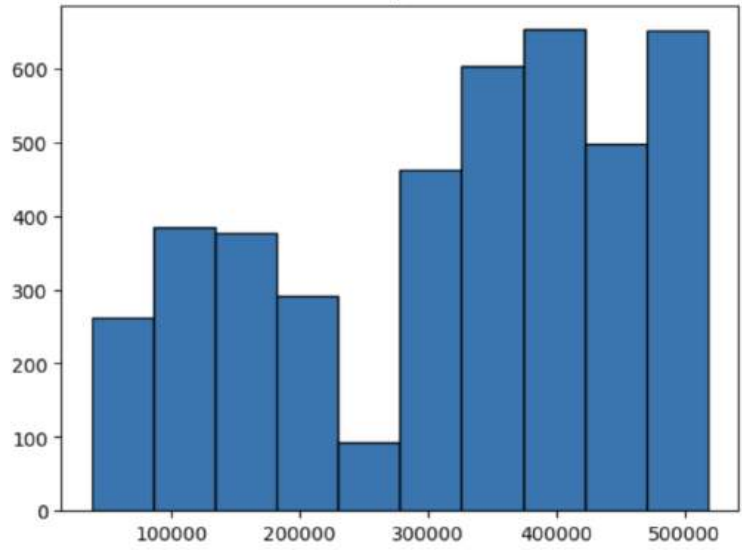


Fig. 4. Graph plot of the solar flare over 0.48 threshold frequency

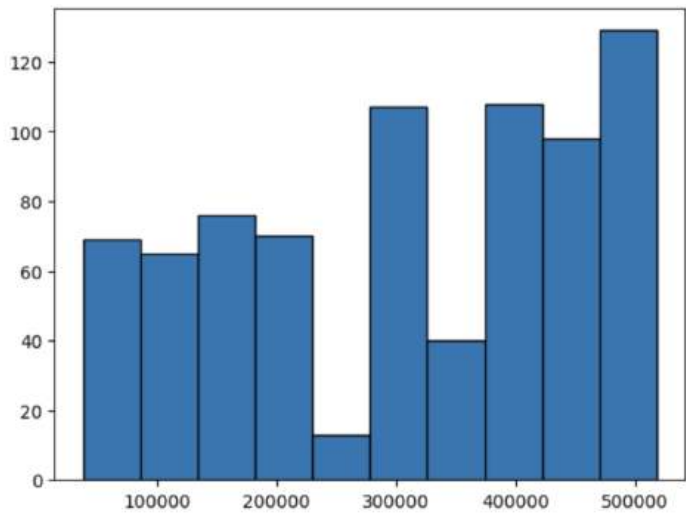


Fig. 5. Graph plot of the solar flare over 0.5 threshold frequency

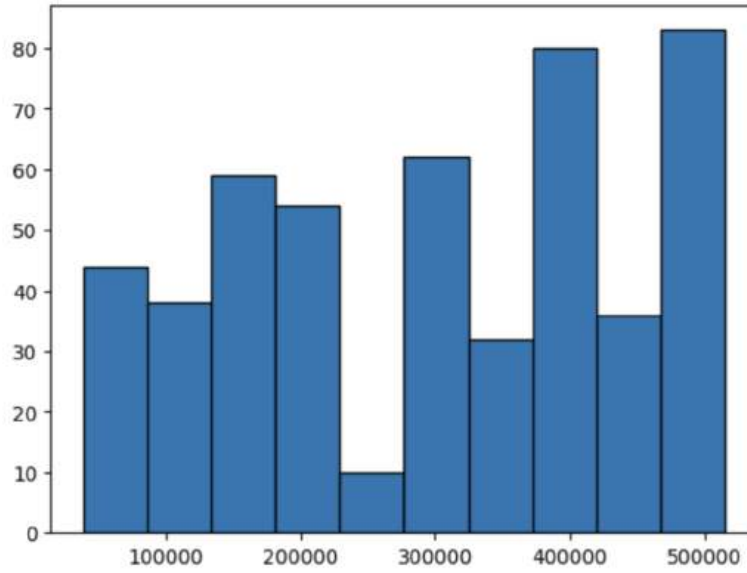


Fig. 6. Graph plot of the solar flare over 0.52 threshold frequency

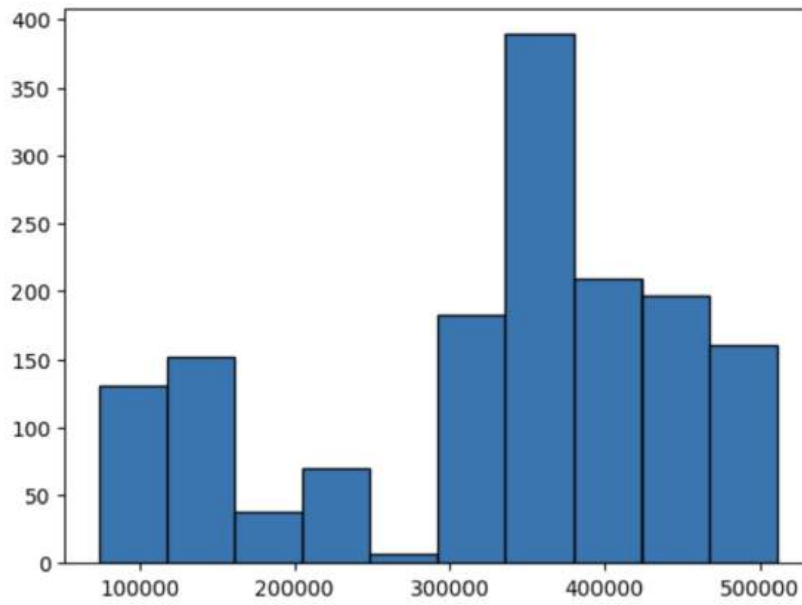


Fig. 7. Graph plot of the solar flare over 0.56 threshold frequency

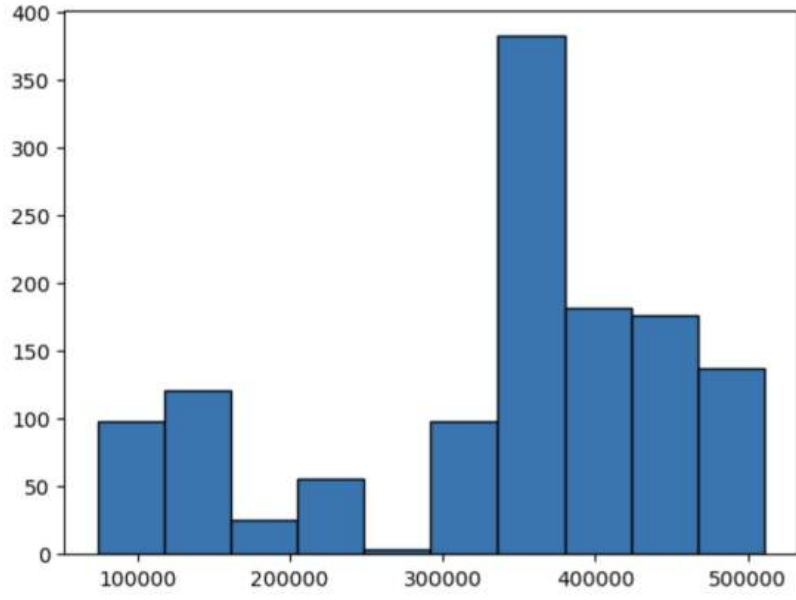


Fig. 8. Graph plot of the solar flare over 0.58 threshold frequency

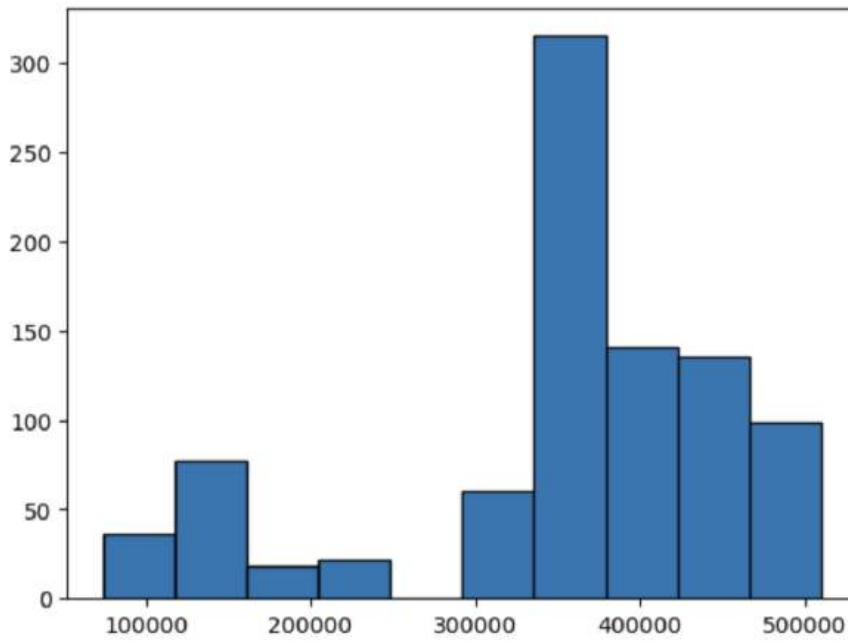


Fig. 9. Graph plot of the solar flare over 0.62 threshold frequency

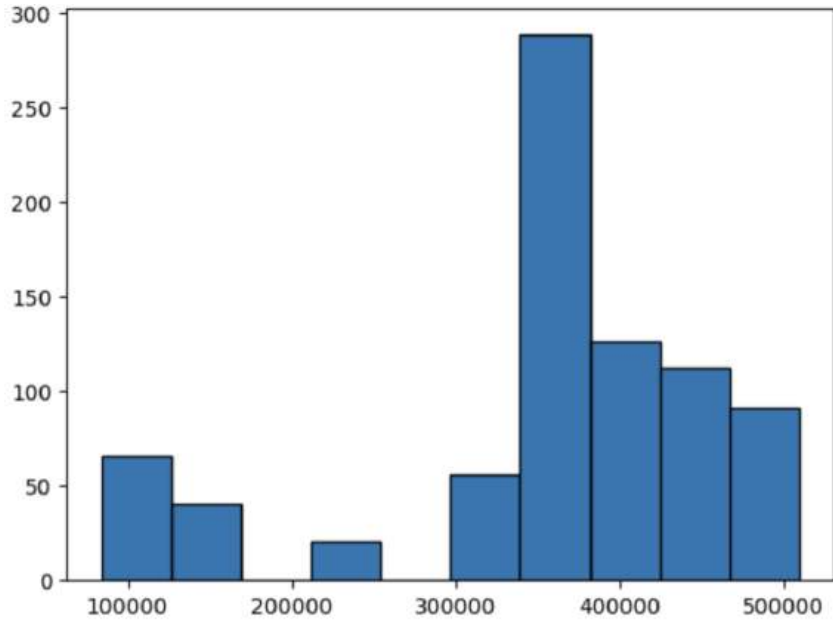


Fig. 10. Graph plot of the solar flare over 0.64 threshold frequency

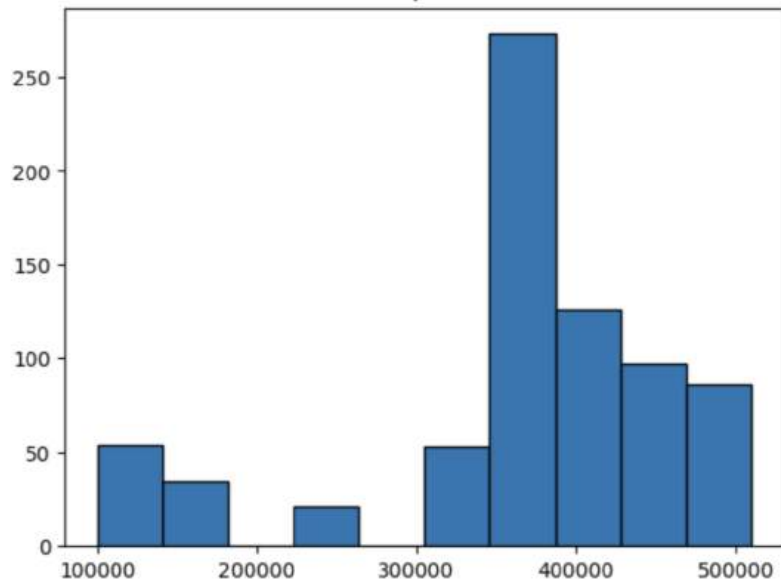


Fig. 11. Graph plot of the solar flare over 0.66 threshold frequency

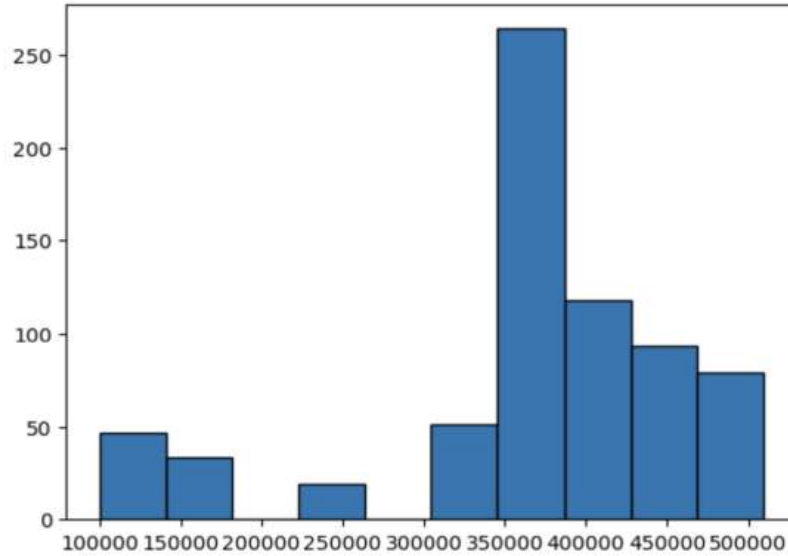


Fig. 12. Graph plot of the solar flare over 0.68 threshold frequency

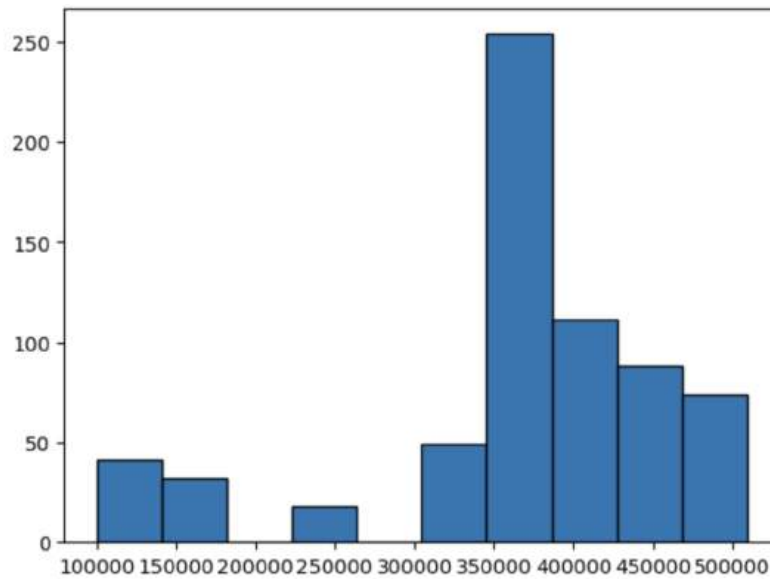
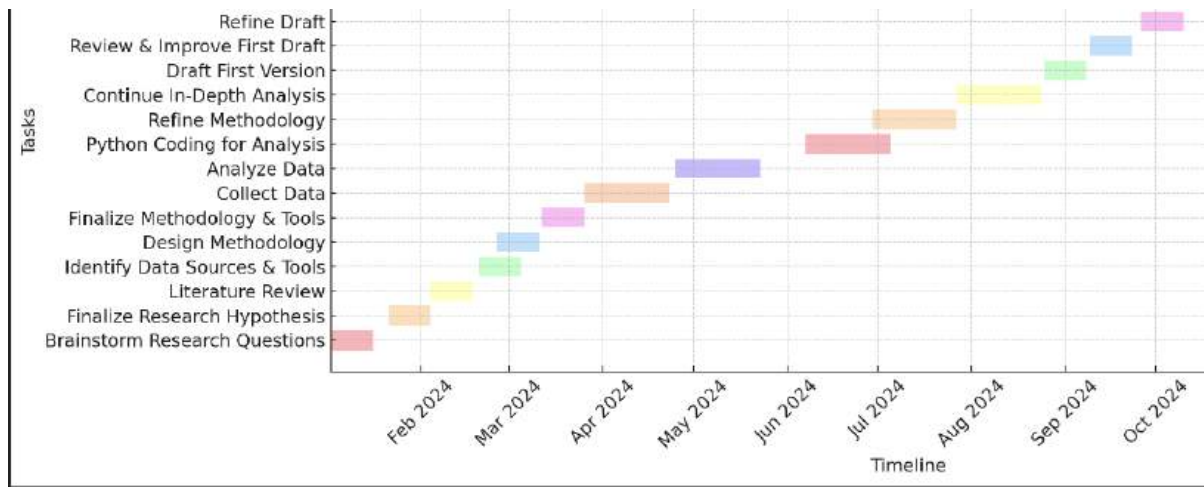


Fig. 13. Graph plot of the solar flare over 0.70 threshold frequency

In these graphs, we can clearly see that when we start from the lower threshold, we see a more uniform distribution. The more threshold we put, we see that the middle values of the intensity decrease, and we get more of lower and higher intensities, with a greater number of higher intensities. This gives a double peak structure to the histogram, hinting at a complex statistical pattern in the distribution of the solar flare intensities.

#### 4. Gantt Chart

This shows the entire duration of the project's timeline in a visual format. The y-axis represents the tasks completed and the x-axis is the duration of the project.



#### 5. Conclusion

Through our research project, we've gained valuable insights into the patterns of solar flares and their temporal distribution. By analysing data from the Solar Dynamics Observatory and employing Python-based techniques, we've mapped out the intensity-time relationship of solar flare events. Our findings underscore the sporadic and localised nature of solar flares, highlighting their potential impacts on space weather and Earth's magnetic environment.

## Works Cited

- [1] Svestka, Z., 2012. *Solar flares* (Vol. 8). Springer Science & Business Media.
- [2] Hudson, H. S. (2011). Solar flares, hard and soft. *Nature*, 474(7353), 616-617.
- [3] Schrijver, C. J., Title, A. M., Yeates, A. R., DeRosa, M. L., & Schrijver, K. (2015). The helioseismic and magnetic imager (HMI) vector magnetic field pipeline: Overview and performance. *Solar Physics*, 289(9), 3483-3530.
- [4] Bastian, T.S., Benz, A.O. and Gary, D.E., 1998. Radio emission from solar flares. *Annual Review of Astronomy and Astrophysics*, 36(1), pp.131-188.
- [5] Shin, S., Lee, J.Y., Moon, Y.J., Chu, H. and Park, J., 2016. Development of daily maximum flare-flux forecast models for strong solar flares. *Solar Physics*, 291, pp.897-909.
- [6] Benz, A.O., 2017. Flare observations. *Living reviews in solar physics*, 14, pp.1-59.
- [7] Ichimoto, K., Kubota, J., Suzuki, M., Tohmura, I. and Kurokawa, H., 1985. Periodic behaviour of solar flare activity. *Nature*, 316(6027), pp.422-424.
- [8] Shin, S., Lee, J.Y., Moon, Y.J., Chu, H. and Park, J., 2016. Development of daily maximum flare-flux forecast models for strong solar flares. *Solar Physics*, 291, pp.897-909.
- [9] Goswami, J.N., McGuire, R.E., Reedy, R.C., Lal, D. and Jha, R., 1988. Solar flare protons and alpha particles during the last three solar cycles. *Journal of Geophysical Research: Space Physics*, 93(A7), pp.7195-7205.
- [10] Podgorny and Podgorny, 1992, Podgorny and Podgorny, 2002, Podgorny et al., 2008
- [11] Li, T., Liu, L., Hou, Y. and Zhang, J., 2019. Two types of confined solar flares. *The Astrophysical Journal*, 881(2), p.151
- [12] Krucker, S.Ä.M., White, S.M. and Lin, R.P., 2007. Solar flare hard X-ray emission from the high corona. *The Astrophysical Journal*, 669(1), p.L49.
- [13] Jain, R., Aggarwal, M. and Sharma, R., 2008. X-ray emission from solar flares. *Journal of Astrophysics and Astronomy*, 29, pp.125-145.
- [14] Svestka, Z., 2012. *Solar flares* (Vol. 8). Springer Science & Business Media.
- [15] Heyvaerts, J., Priest, E.R. and Rust, D.M., 1977. An emerging flux model for the solar flare phenomenon. *Astrophysical Journal, Part 1, vol. 216, Aug. 15, 1977, p. 123-137.*, 216, pp.123-137.
- [16] Emslie, A. G., et al. (2012). *The Astrophysical Journal*, 759(1), 71.
- [17] Schrijver, C. J., & Siscoe, G. L. (2010). *Living Reviews in Solar Physics*, 7(1).
- [18] Pulkkinen, A. (2007). *Space Weather - The International Journal of Research and Applications*, 5(3).
- [19] Shibata, K., & Magara, T. (2011). *Living Reviews in Solar Physics*, 8(1).
- [20] Murray, S. A., et al. (2017). *Space Weather*, 15(5), 577-592.

## What Shaped Jacob Riis Into an Urban Reformer By SeungJun Lee

### Abstract

Jacob Riis's groundbreaking book *How the Other Half Lives*, the first photographic documentary, shed light on the appalling conditions in late 19th-century urban slums and highlighted various social issues. Through his powerful visuals and prose, Riis awakened the conscience of the American middle class and played a pivotal role in initiating urban reform. His immersion in the slums—as a Danish immigrant and later as an American journalist—enabled him to witness the harsh realities of poverty from both sides, lending urgency to his message. In this endeavor, Riis's success as a journalist and reformer was deeply intertwined with his vision of an idealized American identity.

### Introduction

During the late 19th century, major cities in the United States underwent fast urbanization accompanied by industrial development. New York City's development was the most remarkable case. New York was burgeoning to become the greatest and wealthiest city as a manufacturing and trading center. The growth of New York was not separable from the influx of immigrants. Back then, New York was the primary entry port for European immigrants, seeking new lives behind their home countries' political and economic struggles. These immigrants also became a basis for New York's industrial development. Unfortunately, New York City's urban condition was still insufficient to accommodate the explosively increasing population by the influx of immigration. Most of all, housing was severely limited, and overcrowded dwellings for these working-class immigrants resulted in sanitation and safety problems. Since most immigrants who resided in urban slums were unskilled laborers, meager wages for them became another source for various social problems. However, the Laissez-Faire policies of the era embedded in social Darwinist thoughts did not properly address these urban issues. In other words, New York City was like a microcosm of the Gilded Age urban sphere, although middle class and wealthy people did not recognize the reality.

Jacob Riis's *How the Other Half Lives: Studies Among the Tenements of New York* (1890) captured the very deleterious living conditions, inferior work environments in factories, rampant vagrancy, neglected poor children, shabby buildings, and other dark sides of New York City in the late 19<sup>th</sup> century. Riis's work that vividly exposed dark realities of New York's impoverished neighborhoods became a trigger for urban reform movement by evoking the public's concern about the unspoken social problems of the New York city whose dwellers lived in overcrowded, squalid, unsanitary, and unsafe environments.

However, Riis was neither the only nor the first writer who raised an alarm about the health, safety, and well-being of people in slums during the Gilded Age. Yet the power of his images was so strong and effective that Riis helped prompt a national slum improvement agenda during the Progressive Era. The American middle class, visually encountering devastating poverty and squalid conditions depicted in Riis's photos, felt morally obliged to take action to



improve the condition of the poorest districts in large cities. The Tenement House Act of 1901, which codified sanitation and construction safeguards in tenement dwellings in New York for the first time, in large measure owed its passage to his messages delivered in *How the Other Half Lives*. In some sense, this photojournalistic documentary exemplifies how ‘a picture is worth a thousand words.

Jacob Riis’s exposure of the reality of slum dwellers continued through his subsequent works, such as *A Ten Years’ War: An Account of the Battle with the Slum in New York* (1900), *The Battle with Slum* (1902), the expanded version of *A Ten Years’ War*, and *The Children of the Poor* (1892). As reform-oriented documentaries, his works successfully enlightened people and spurred reform. Theodore Roosevelt once called Jacob Riis “the most useful citizen” in New York. Jane E. Robbins, head of the Neighborhood House, said that Jacob Riis was “the finest immigrant that we have ever known” when Riis died.<sup>3</sup> Given the impact he brought to the American public’s awareness about slum problems, it is not an exaggeration to say that Jacob Riis was one of the most distinctive reformers of his era.

### **A New Immigrant**

A simple answer to what really made him want to be outspoken about the living conditions in slums is found in the fact that Jacob Riis knew slum conditions very well. He was a Denmark-born immigrant who arrived at New York in 1870 at the age of twenty. Landing in America with barely any money, he was among the thousands of “poor, friendless, and unskilled immigrants.”<sup>4</sup> The very penniless and homeless years Riis had lived before becoming a reporter are a definite clue to his lifelong devotion to slum reform.

Jacob Riis migrated from Denmark to the United States in 1870 with \$40 in his pocket. Like most other poor immigrants of this era, he traveled in steerage, the space for the passengers paying the lowest fare. The conditions in steerage were far from good: lack of ventilation, an inadequate water supply, the odor of unwashed bodies, the reek of food, the stench of nearby toilets...everything seemed to presage the harsh conditions they would ultimately live in.<sup>5</sup> Riis was one of them.

It is true that Riis had some advantages compared to most of the other steerage passengers. Riis did not have any dependents: he was “responsible to no one but himself.”<sup>6</sup> Because of his education in Denmark, he could speak English well enough, even with his heavy accent. Furthermore, he was from a northern European Protestant background, shared with many mainstream, well-off Americans. His “Danish Lutheran” background made him distinct from the “majority of the new immigrants, who were mostly Irish Catholic, Italian Catholics, Jews, or Chinese.”<sup>7</sup> Even so, these advantages did nothing positive for his first days in New York.

---

<sup>3</sup> Ferenc M. Szasz, Ralph F. Bogardus, and Ralph H. Bogardus. “The Camera and the American Social Conscience: The Documentary Photography of Jacob A. Riis.” *New York History* 55, no. 4 (1974): p. 411

<sup>4</sup> *Ibid.*, p. 409

<sup>5</sup> Janet B. Pascal, *Jacob Riis: Reporter and Reformer* (Oxford & New York: Oxford University Press, 2005), p. 18

<sup>6</sup> *Ibid.*

<sup>7</sup> *Ibid.*, p. 18-20

Without any clear idea what to do in America, Riis stayed alive by doing miscellaneous manual labor: ironworker, stocker, day worker at a clay yard or brick yard.<sup>8</sup> None of them was a stable job, and sometimes he was not paid for his work. During the first few years of his arrival, he reached the bottom and experienced the low point of his life. His autobiography, *The Making of an American*, illustrates his harsh days very elaborately. One day, he had only one cent in [his] pocket, but put up at a boarding house, which charged one dollar a day. He had to pawn his revolver and top-boots, “the only valuable possessions [he] had left,” to pay for his lodging. After that, he was “thrown on the street.” He spent a night on the street, sleeping in a milk wagon until the owner threw him out at dawn.<sup>9</sup> Even though his personality made him too proud to beg, Riis gladly accepted food offered by an old monk at Fordham College.<sup>10 11</sup> His money was all gone again in the late fall, when the brick-making season was over and a promise of employment in a factory failed. He had to wander the street as one of a number of idle men.<sup>12</sup> He survived on “meat-bones and rolls” generously offered by a cook at an elegant New York restaurant. On a cold and stormy day when his optimism disappeared, he even considered suicide for the first time in his life.<sup>13</sup> On another day, he had to spend the night on the floor of the New York Police Lodging House at Church Street. The police lodging was a crude form of shelter for the homeless, who did not have any other option. Without any toilet, water for washing, bedding, food, or other assistance, they provided only a bare place to sleep in damp cellars. Yet, they were always crowded with desperately poor people living on the street.<sup>14</sup> On that night, there was a disturbance, in the midst of which Riis received unfair treatment by a policeman and he felt outrage against the police. That night’s anger would contribute to his determined position in attacking the injustice and corruption of New York City.

During this harsh period, Riis came to know slum areas such as Mulberry Bend and the Five Points. According to his own explanation, his personal knowledge of slums led him later to devote himself to slum reform with ‘maudlin sensitiveness.’<sup>15</sup>

Every once in a while I am asked why I became a newspaper man. For one thing, because there were writers of such trash, who, themselves comfortably lodged, have not red blood enough in their veins to feel for those to whom everything is denied, and not sense enough to make out the facts when they see them, or they would not call playgrounds, schoolhouses, and better tenements “abortive measures.” Someone had to tell the facts; that is one reason why I became a reporter.<sup>16</sup>

## The Journalist

---

<sup>8</sup> Ibid., p. 20

<sup>9</sup> Jacob A. Riis, *The Making of an American* (New York: Macmillan, 1901); Kindle Edition (2011): p. 18-19.

<sup>10</sup> Ibid., p. 19

<sup>11</sup> Ibid., p. 25

<sup>12</sup> Ibid.

<sup>13</sup> Ibid., p. 26.

<sup>14</sup> Pascal, *Jacob Riis: Reporter and Reformer*, pp. 27-28

<sup>15</sup> Riis, *The Making of an American*, p. 25

<sup>16</sup> Ibid.

In 1877, Jacob Riis became a police reporter at *the New York Tribune*, the nation's largest circulating newspaper back then, which would bring him stable employment and income, and more importantly fame as a journalist and reformer. Yet, his nascent career as a newspaper man already began in 1873 when he was offered a position with the New York News Association. His duty was to cover and report on "anything that happened anywhere in New York City."<sup>17</sup> At the New York News Association, he was truly a hard worker. Because this news service sold morning and evening papers, Riis had to cover about six or seven stories a day. He had to travel everywhere in New York City, from the top of Harlem to the lowest part of Manhattan Island, on foot or by way of slow street cars. He often worked from ten o'clock in the morning to two o'clock the next morning.<sup>18</sup> Yet he willingly threw himself into the work. Riis's hard work paid off: he gained strength in his English writing, and "speedily earned the name of a good reporter." Thanks to the reputation he earned, in 1874, *South Brooklyn News* hired him at a high salary for his, then standard, weekly \$15. After only two weeks, the owners of this newspaper appointed this hardworking Danish young man as an editor.

While working at *South Brooklyn News*, he opened his eyes to the mechanism of how 'crook' political bosses exerted power in New York City. *South Brooklyn News* was owned by Democrats who used their newspaper to gain public support for their candidates. In the beginning, he lacked knowledge about or interest in American politics.<sup>19</sup> Yet, Riis desired to assume a certain degree of control over the paper and maintain integrity as an editor. For this, he struggled with the owners of *South Brooklyn News* who were not interested in unbiased news, but instead wanted their agenda promoted. In 1875, Riis bought *South Brooklyn News* from its original owners, and left behind mere survival in favor of financial security. "My zeal for reform," he wrote in his autobiography, *The Making of an American*, "encompassed the whole range of my little world."<sup>20</sup> Even so, it was not easy for him to staunchly stand as an impartial journalist because of South Brooklyn Democrats' influence that did not let him freely present his ideas when they conflicted with the party's political interest. He resold the *South Brooklyn News* to the original owners and briefly worked as a reporter at the same firm, but quit.<sup>21</sup> At that time, the corrupt party boss system controlled all aspects of American large cities' politics, from political appointments to government jobs to publicly funded projects. A politician, if supported by a party boss, could 'safely' take bribes or embezzle government money. The poor immigrants, however, generally supported boss politics because it helped to satisfy their immediate needs such as cash or a job in return for their vote—even if their support of corrupt boss politics would ruin their lives in the long run.<sup>22</sup> <sup>23</sup> His first-hand knowledge would pave the way for Riis's later

---

<sup>17</sup> Pascal, *Jacob Riis: Reporter and Reformer*, pp. 27-39

<sup>18</sup> Riis, *The Making of an American*, p. 48.

<sup>19</sup> Pascal, *Jacob Riis: Reporter and Reformer*, p. 40.

<sup>20</sup> Riis, *The Making of an American*, p. 53.

<sup>21</sup> In *The Making of an American*, he wrote: "I can say honestly that I tried the best I knew how to get along with the politicians I served, but in the long run it simply could not be done. They treated me fairly, bearing no grudges. But it is one thing to run an independent newspaper, quite another to edit an "organ." (p. 71)

<sup>22</sup> Pascal, *Jacob Riis: Reporter and Reformer*, pp. 42-4.

<sup>23</sup> Jacob A. Riis, *The Battle with the Slum* (New York: The Macmillan, 1902); Kindle Edition (2011): Kindle Locations: 3486-3493

combat with crooks in the Democratic Party, which dominated New York politics and stymied reform at this time.

Finally in 1877, Riis got a job as a police headquarters reporter at *The New York Tribune*, where he began to blossom as a remarkable journalist. Not only did it pay a higher wage of \$25 per week, but the new job was also a good fit for Riis's personality. The *Tribune* was a reform-oriented newspaper unlike the *South Brooklyn News*. Founded by Horace Greeley in 1845, whose "objective as a reformer and public intellectual" was "to be a mediator, an interpreter, and a reconciler between Conservatism and Radicalism," the *Tribune* promoted stories to improve conditions for the working class since its beginning as a locus of social democratic thought.<sup>24</sup> <sup>25</sup> The *Tribune*'s mission fitted well with Riis's reform-minded vision, too. The job of police reporter was harsh, covering murders, fires, suicides, and robberies, and fascinating human stories within the incidents; it also included fighting with the police who hid key information, or competing with other reporters.<sup>26</sup> Yet, Riis enjoyed his work and embraced its challenges; he thought this job was "often of the highest service to the public."<sup>27</sup>

By the mid-1880s, Riis had earned a reputation as a great crime reporter.<sup>28</sup> His writing was powerful, with descriptions of crime scenes that were "vivid and evocative" with detailed truth.<sup>29</sup> According to his friend and a colleague, journalist Lincoln Steffens, "only Riis wrote [news] as stories, with heart, humor, and understanding... and [saw] the human side of crime or disaster," although all the reporters got the same news.<sup>30</sup> Furthermore, Riis was getting more first-hand knowledge about slums, whose problems he was well familiar with. The problems were getting worse partially because of political corruption. Riis "not only got the news but also cared about the news" because Riis "hated passionately all tyrannies, abuses, miseries, and he fought against them."<sup>31</sup>

It was the power of his photographs that most fundamentally established and broadened Riis' unique standing as a journalist, and later as a philanthropic reformer. He was an amateur as a photographer, but probably the first photographer who realized the usefulness of documentary photography.<sup>32</sup> Between the fall and winter of 1887, he and two other amateur photographers took several midnight trips to tenement areas and took pictures of them. In January 1888, Riis presented these pictures with his public lecture entitled "The Other Half: How It Lives and Dies in New York" to the Society of Amateur Photographers of New York. In the following months, he gave a series of similar lectures in different churches in New York City.<sup>33</sup> Among the people

---

<sup>24</sup> Adam Tuchinsky, *Horace Greeley's "New York Tribune": Civil War-Era Socialism and the Crisis of Free Labor*. (Cornell University Press, 2009), p. 18

<sup>25</sup> *Ibid.*, p. 191

<sup>26</sup> Riis, *The Making of an American*, pp. 79-81.

<sup>27</sup> *Ibid.*, p. 81

<sup>28</sup> O'Donnell, Edward T. "Pictures Vs. Words? Public History, Tolerance, and the Challenge of Jacob Riis." *The Public Historian* 26, no. 3 (2004): p. 12

<sup>29</sup> Pascal, *Jacob Riis: Reporter and Reformer*, p. 65.

<sup>30</sup> Lincoln Steffens, *The Autobiography of Lincoln Steffens*. (Berkeley, CA: Heyday Books, 2005), p. 204

<sup>31</sup> *Ibid.*, p. 203

<sup>32</sup> O'Donnell, "Pictures Vs. Words?", p. 13

<sup>33</sup> Szasz, et al., "The Camera and the American Social Conscience," p. 421

who saw Riis's pictures at the churches was an editor of *Scribner's Magazine*. Impressed by the images, the editor suggested Riis write an article for *Scribner's* December 1889 issue, which was titled "How the Other Half Lives" and instantly received attention from the press and public.<sup>34</sup> Just one week after the article was published, *Critic's* editor Jeannette Gilder sent Riis a letter to ask if he would be interested in expanding the article into a book.<sup>35</sup> Riis's first book, *How the Other Half Lives: Studies Among the Tenements of New York*, was released in November, 1890. This book became the official inauguration for Riis's crusade for slum reform. Housing was one of the biggest problems in Gilded Age New York. First of all, a lodging house that Riis had also experienced exemplifies the housing issue for many unskilled laborers who could not even afford stable residence. As he wrote in his autobiography, he had tried to attract the public and government's attention to slum problems through writing, but those efforts "seemed to make no impression."<sup>36</sup> When [he] told [his] story in print, illustrated with the pictures:

[t]he town was astounded. The Board of Health sent inspectors to the watershed, who reported that things were worse a great deal than I had said. Populous towns sewerred directly into our drinking-water. There was not even a pretence at decency. The people bathed and washed their dogs in the streams. The public town dumps were on their banks.

<sup>37</sup>

With illustrations and photographs, the same messages made a huge difference.

The subject of necessary urban reform was not a newly or solely raised question by Riis. Charles Loring Brace had already published *The Dangerous Classes of New York, Twenty Years Work Among Them* (1872), which had even more extensively investigated the realities of life among New York's poorest class than Riis.<sup>38</sup> Before the civil War, John Griscom had published *The Sanitary Condition of the Laboring Population of New York with Suggestions for Its Improvement* in 1845 and wrote the phrase, 'one half of the world does not know about how the other half lives,' which was often recited by John B. Gough, a temperance lecturer, and inspired the title of Riis's book.<sup>39</sup> Yet, the impact of *How the Other Half Lives* was seminal because photographs in the book vividly embodied what the problem is, as the poet James Russel wrote in his letter to Riis: "I had but a vague idea of these horrors before you brought them so feelingly home to me."<sup>40</sup> The images captured something true and made the suffering distinct. Riis

---

<sup>34</sup> Riis, *The Making of an American*, p. 116

<sup>35</sup> *Ibid.*, p. 117

<sup>36</sup> *Ibid.*, p. 103

<sup>37</sup> *Ibid.*, p. 81

<sup>38</sup> Charles Loring Brace. *The Dangerous Classes of New York, Twenty Years Work among Them*. New York: Wynkook & Hallenbeck, 1872; Charles Loring Brace. *The Dangerous Classes of New York, Twenty Years Work among Them*, Kindle Edition, 2010.

<sup>39</sup> Szasz, et. al., "The Camera and the American Social Conscience," p. 423

<sup>40</sup> Miller, Stephen. *Walking New York: Reflections of American Writers from Walt Whitman to Teju Cole*. (New York: Fordham University Press, 2015), p. 90

“awakened intelligent men to conditions of which they were unaware.”<sup>41</sup> Moreover, Riis’s writing style was livelier and more personal than that of others on the subject.<sup>42</sup>

### The Reformer

Riis viewed the problems of the slums as evils more destructive than wars.<sup>43</sup> At a basic level, it was a series of terrible and inhumane conditions in and around tenement houses: overcrowded, filthy, and unsafe, with a lack of light and air, all for the cost of rent.<sup>44</sup> Passing through the second industrial revolution of the late 19th century, American cities grew rapidly with the unprecedented increase in the number of city dwellers. Most immigrants who newly immigrated in the late 19th century were from south-eastern European countries including Greece, Italy, Russia, Poland, and Hungary, an underprivileged region of Europe at that time. They sought better opportunities in the United States. However, limited capital and skills confined their opportunities within wage-labor or family wage earners.<sup>45</sup> While these migrant workers supplied a plethora of cheap labor, which geared mass production, major industrial cities could not provide proper living space for them. Proper housing was too expensive for them to afford. What Jacob Riis exposed through his photographs were the deleterious living conditions of urban slums, where migrant working class families usually lived crowded together in small places, such as subdivided apartments called tenements or boarding houses. Overcrowded slums also suffered from inadequate sewage systems. Their alleys were unclean, waste-piled, and foul-smelling. Needless to say, disease was rampant. Furthermore, slum housing was vulnerable to fire, lacking fire extinguishers or sprinkler systems. Riis’s *How the Other Half Lives* was filled with verbal and visual testimonials of the very realities of slums.<sup>46</sup>

However, Riis’s book not only exposed the problem of ‘housing evils,’ but also presented social or moral ills such as the widespread neglect of children, the prevalence of tramps, the general inhumaneness of tenement existence, and corruption of governance in slums. He did not detach the other social ills, such as the corruption of city officials and landlords, from the

---

<sup>41</sup> Szasz, et. al., “The Camera and the American Social Conscience,” p. 423

<sup>42</sup> Pascal, *Jacob Riis: Reporter and Reformer*, p 123.

<sup>43</sup> In the mid nineteenth century, the urban population was 15% of the national population but it climbed to 47% of the national population by 1910. (Klein 2004, 142) For instance, the population of New York City in the census of 1860 was 813,669 but it sharply increased to 1,515,301 in 1890 and climbed again to 3,437,202. Internal migration from rural areas to urban areas partially contributed to those urban growths. Yet, most notable was influx of the immigrants to big cities. The census of 1890 indicates over 40% of residents in New York were 1st generation immigrants. (Gibson 1998) By 1900, 60% of the largest twelve cities’ residents would be first or second generation immigrants. (Boyer 123-4)

<sup>44</sup> de Forest, Robert W. “Recent Progress in Tenement-House Reform.” *The Annals of the American Academy of Political and Social Science* 23, (1904): p. 104

<sup>45</sup> Nugent, Walter T. K. *Crossings: The Great Transatlantic Migrations, 1870-1914*. (Bloomington · Indianapolis, IN: Indiana University Press, 1995), pp.156-164.

<sup>46</sup> In a picture, titled *Lodgers in a crowded Bayard Street tenement*, the readers can see three men who were sleeping lying on the wall in a narrow room even for one person.(37) Another picture, titled *The man slept in this cellar for four years, about 1890*, shows a man on a board over two barrels. (99) People who could not rent bedrooms in the “shameful old wrecks of buildings” or pay for a few cents for a night may seek for police lodging station, where many spent the night seating or standing. This book includes several scenes of police lodging stations.

problems of slum environments. He says “all life eventually accommodates itself to its environment, and human life is no exception.”<sup>47</sup> Because he believed that an individual’s surroundings form his or her personality and values, he viewed the slum environment as negatively affecting poor people’s moral, ethical, physical, and intellectual development. Thus, the more profound reason that he was concerned about poor conditions of slums was the social and moral diseases they would incubate.<sup>48</sup> “How shall the love of God,” he asks in the introduction of his book, “be understood by those who have been nurtured in sight only of the greed of man?” He was concerned that the culture of slums and their immoral values could ultimately infect the entire society.<sup>49</sup>

Riis argued that private greed largely caused the slum problems, but that public ignorance worsened the situation. As a reformer, therefore, Riis reminded his readers that it is the obligation of ‘one half’ to care about the realities of ‘the other half.’

Long ago, it was said that “one half of the world does not know how the other half lives.” That was true then. It did not know because it did not care. The half that was on top cared little for the struggles, and less for the fate of those who were underneath, so long as it was able to hold them there and keep its own seat.<sup>50</sup>

As Riis might have intended, publication of *How the Other Half Lives* opened the eyes of the middle class American to the realities of slum dwelling, inspired a generation of reformers, and established Jacob Riis as one of the century’s most prominent spokesmen for the urban poor.<sup>51</sup> The most remarkable success of Riis’ work was concrete reforms in housing.<sup>52</sup> A decade after the publication of *How the Other Half Lives*, the Tenement House Act of 1901 was passed in New York, and the new law dramatically improved urban poor immigrants’ housing.<sup>53</sup> His work also led to the development of the Tenement House Department in New York, formed in 1902.<sup>54</sup> Some tenements were demolished to provide space for public parks. Between 1902 and 1903, other big cities including Chicago, Pennsylvania, Cleveland, and St. Louis enacted similar tenement housing laws or observed housing reform movements.<sup>55</sup> Before Riis, reform efforts belonged to the private, charitable sphere; but under Riis’ influence, the reform agenda was moved to the public and political sphere.

After the publication of his first book, Riis kept up the fight for slum improvement through subsequent books and through campaigns. Most remarkably, he paid special attention to children in slums. Firstly, slum children were vulnerable to labor exploitation. At the end of the

---

<sup>47</sup> Riis, Jacob. *How The Other Half Lives*, p. 91.

<sup>48</sup> James B. Lane, *Jacob A Riis and the American City* (Port Washington, NY: Kennikat Press, 1974), p. 39.

<sup>49</sup> “Riis wanted poverty because the culture of slums affected younger generations and hardened into permanent patterns for individuals, families, communities, and ultimately, the nation.” *Ibid.*, p. 65

<sup>50</sup> Riis, Jacob. *How The Other Half Lives*, Introduction

<sup>51</sup> Pascal, *Jacob Riis: Reporter and Reformer*, p 7

<sup>52</sup> O'Donnell, “Pictures Vs. Words?” p. 14

<sup>53</sup> *Ibid.*, pp.14-5

<sup>54</sup> de Forest, Robert W. “Recent Progress in Tenement-House Reform,” p. 103.

<sup>55</sup> *Ibid.*, pp.103-4.

19<sup>th</sup> century, it was not uncommon for children in poor immigrant families to work in sweatshops in tenements making cigars, shirts, or flowers. Employers favored child labor at tenement sweatshops because they could bypass labor law restrictions that applied to factories. Further, child labor was ‘cheap’ everywhere in slum districts. Second, children of poor immigrants were often neglected. In Riis’ view, they were “too dirty, too ragged, and too generally disreputable, too well hidden in their slum.”<sup>56</sup> Because every inch of slums was crammed with tenement buildings at that time, children could hardly find any space for play. For many of those not old enough to work, the street was a playground. Riis wrote that the “rescue of the children is the key to the problem of city poverty.”<sup>57</sup> In 1892, he published *The Children of The Poor*, expounding with more details about the realities of poor immigrant children. In many places in this book, Riis emphasized the importance of play to children. On slum streets, he thought, children would “join gangs, and spend their time with beggars and thieves, and start on a life of crime” if no alternatives were provided.<sup>58</sup> As a supporter of creative play, Riis believed that recreation and freedom to move freely were necessary ingredients of ‘proper citizenship.’<sup>59</sup>

For Riis, creating desirable neighborhoods and communities, beyond housing improvement, was an important task to reform slums. In his *Battle with the Slums* published in 1902, he wrote:

Every park, every playground, every bath-house is a nail in the coffin of the slum, and every big, beautiful school- house, built for the people’s use, not merely to lock the children up in during certain hours for which the teachers collect pay, is a pole rammed right through the heart of it.<sup>60</sup>

He argued for the building of small parks to provide a place where whole families could gather. He also claimed that schools should be opened to the community. Indeed, Riis was seeking healthy alternatives to streets and saloons, which fostered the ills of slums.

After Riis became renowned as a journalist and reformer, his vision was shared with other reformers such as Lincoln Steffens, the Muckraker who exposed details of the Tammany Hall government’s corruption and its involvement with prostitution, criminal activity, and drunkenness; Jane Addams, the iconic leader of settlement house movement; and Theodore Roosevelt, who would become the president of United States in 1901 and implemented many progressive reform agenda.<sup>61</sup> For Riis, journalism served as “a vehicle for securing a higher place

---

<sup>56</sup> Riis, Jacob. *How The Other Half Lives*, P. 105.

<sup>57</sup> *Ibid.*, p. 108.

<sup>58</sup> Pascal, *Jacob Riis: Reporter and Reformer*, p 134

<sup>59</sup> Lane, *Jacob A Riis and the American City*, p. 40.

<sup>60</sup> Jacob A. Riis, *The Battle with the Slum*, (New York: The Macmillan Company, 1902), p. 283

<sup>61</sup> Lincoln Steffens, a muckraker, who fought against municipal corruption, was among them, too. In 1892, Steffens came to Riis’s office in Mulberry Street reporters’ building to work with Riis. Because Steffens and Riis shared some sort of moral visions and both hated corruption in machine politics, they became instantly friendly. Working with Riis, Steffens revealed details of the Tammany Hall government’s involvement with prostitution, criminal activity, and drunkenness.



in American society,” for protesting slum conditions, and for thrusting him into contact with people who shared his concern with the decaying city.”<sup>62</sup>

In particular, the friendship between Riis and Roosevelt deeply impacted the two men’s pathways as reformers. It was 1884 when Roosevelt first developed views similar to Riis’: Roosevelt, a state assemblyman back then, had visited tenements and sweatshops, guided by Samuel Gompers, the labor activist of American Federation of Labor.<sup>63</sup> Shortly after the publication of Riis’ first book in 1890, Roosevelt sent a message to Riis: “I have read your book, and I have come to help.” Nothing happened instantly, but in 1895, Roosevelt was appointed New York City police commissioner, a position proposed to reform the almost entirely corrupt police department. Roosevelt treated Riis as if he had been a member of the police board, and asked his advice on many problems in regard to the police. Because of Riis’ advice, Roosevelt abolished police lodging houses “which were,” Roosevelt wrote in his autobiography, “simply tramp lodging houses, and a fruitful encouragement to vagrancy.”<sup>64</sup> Riis also took Roosevelt on midnight trips to show him tenement conditions in different slum districts. Those midnight trips with Riis helped Roosevelt see what problems impaired the police department and gain personal insight about urban problems. “It is one thing to listen in perfunctory fashion to tales of overcrowded tenements,” Roosevelt wrote, “and quite another to see what the overcrowding means, some hot summer night.”<sup>65</sup> Later, Roosevelt became the governor of New York and then President of the United States; he forcefully advanced the progressive reform era, and Riis became one of the key municipal reformers, backed by Roosevelt.

Roosevelt and Riis were of similar in personality; both were vigorous and uncompromising in advancing what they believed was right. Riis firmly believed in the positive effects of local changes, and urged the passage of the Small Parks Act of 1887, which would demolish tenement housings in some slum districts in the early 1900s. However, Riis overlooked the fact that the tenements were, despite being disease-laden or crime-laden, homes for many of the poorest families. When parks displaced many of them, they had to move to another overcrowded tenement or became homeless.<sup>66</sup> Yet Riis and Roosevelt together did effect significant positive change. Urban spaces underwent improvement, and they were on the political agenda.

Of course, Riis does not hide his prejudiced views about specific ethnic groups in slums. Riis alludes to what he considers the moral inferiority of certain ethnic groups and by implication glorifies white-middle class Americans’ value system. Despite Riis’s immersion in the slums, where he experienced the prejudices of others, he lapses into prejudices of his own. One would think that his rise from poor Danish immigrant to acclaimed American journalist would make him more tolerant, but he faltered in failing to undo his own biases and bigotries. Nevertheless,

---

<sup>62</sup> Lane, *Jacob A Riis and the American City*, p. 29.

<sup>63</sup> Pascal, *Jacob Riis: Reporter and Reformer*, p 125.

<sup>64</sup> Theodore Roosevelt. *An Autobiography* (New York: Macmillan, 1913), p. 213

<sup>65</sup> Ibid.

<sup>66</sup> Pascal, *Jacob Riis: Reporter and Reformer*, p 123.

his rise allowed him to see both sides of the squalor, and gave power to his work as a potent wake-up call.

Writing about those poor immigrants, Riis was not one of ‘the other half’ but a visitor and adamant reformer from the first one half. That is, he spoke about the problems of immigrant groups from ‘an American’s view point.’ Riis saw himself as espousing American ideals. But he exposed his own prejudices in the process. Despite this, his photojournalism was remarkable for its ability to relay the true and horrible conditions of tenement life in an accurate and poignant way. Without Riis’s moving photographs and effective writing, the necessary reforms that improved conditions during the Progressive Era would not have geared up as quickly or with as much momentum. Riis, with his vision, skillful lens, and insightful words, was truly a catalyst for social change, who emerged from a poor immigrant to an ‘ideal’ citizen.

### Works Cited

- Brace, Charles Loring. *The Dangerous Classes of New York, Twenty Years Work among Them*. New York: Wynkook & Hallenbeck, 1872; Kindle Edition 2010
- de Forest, Robert W. "Recent Progress in Tenement-House Reform." *The Annals of the American Academy of Political and Social Science* 23, (1904): 103-116.
- Fried, Louis. "Jacob Riis and the Jews: The Ambivalent Quest for Community." *American Studies* 20, no. 1 (1979): 5-24.
- Gibson, Campbell. "Population of the 100 Largest Cities and Other Urban Places in the United States: 1790 to 1990 (Washington D.C.: Population Division, U.S. Bureau of the Census, 1998). [Online]
- Gurock, Jeffrey S. "Jacob A. Riis: Christian Friend or Missionary Foe? Two Jewish Views." *American Jewish History* 71, no. 1 (1981): 29-47.
- Lane, Jaems B. *Jacob A Riis and the American City*, Post Washington, NY: National University Publications, 1974
- Lane, James B. "Jacob A. Riis and Scientific Philanthropy During the Progressive Era." *Social Service Review* 47, no. 1 (1973): 32-48.
- Lloyd, William J. "Understanding Late Nineteenth-Century American Cities." *Geographical Review* 71, no. 4 (1981): 460-471.
- McCarthy, Michael P. "Urban Optimism and Reform Thought in the Progressive Era." *The Historian* 51, no. 2 (1989): 239-262.
- Miller, Stephen. *Walking New York: Reflections of American Writers from Walt Whitman to Teju Cole*. Empire State Editions ed. New York: Fordham University Press, 2015.
- Nugent, Walter T. K. *Crossings: The Great Transatlantic Migrations, 1870-1914*. Bloomington & Indianapolis, IN Indiana University Press, 1995.
- O'Donnell, Edward T. "Pictures Vs. Words? Public History, Tolerance, and the Challenge of Jacob Riis." *The Public Historian* 26, no. 3 (2004): 7-26.
- Pascal, Janet B. *Jacob Riis: Reporter and Reformer*. Oxford & New York: Oxford University Press, 2005.
- Quirke, Carol. "Picturing the Poor: Jacob Riis's Reform Photography." *Reviews in American History* 36, no. 4 (2008): 557-565.
- Riis, Jacob A., *The Making of an American*, New York: Macmillan, 1901; Kindle Edition, 2011
- Riis, Jacob A. *How the Other Half Lives: Studies Among the Tenements of New York*, New York: Charles Scribner's Sons, 1890; Riis, Jacob A. *How the Other Half Lives: Studies Among the Tenements of New York*, Seven Treasures Publications, Special Illustrated Edition including the Author's photographs, 2009
- Riis, Jacob A. *The Battle with the Slum*, New York: The Macmillan Company, 1902
- Roosevelt, Theodore. *An Autobiography* New York: Macmillan, 1913; Roosevelt, Theodore. *An Autobiography the Floating Press*, 2014
- Steffens, Lincoln. *The Autobiography of Lincoln Steffens*. Berkeley, CA: Heyday Books, 2005

Szasz, Ferenc M., Ralph F. Bogardus, and Ralph H. Bogardus. "The Camera and the American Social Conscience: The Documentary Photography of Jacob A. Riis." *New York History* 55, no. 4 (1974): 408-36.

Tuchinsky, Adam. *Horace Greeley's "New York Tribune": Civil War-Era Socialism and the Crisis of Free Labor*. Cornell University Press, 2009.

## **How does Dynamic Pricing Affect Airline Customers? By Diego Alvarez Gonzalez**

### **Abstract**

This study establishes that dynamic pricing impacts airline customers' behaviour, perception and experience in a remarkable way. A popular practice which has become more common especially in the airline industry is dynamic pricing which sees tickets being sold at different prices depending on the time of booking, their popularity and other factors. However, this is a strategy that can be used to maximize the revenues of airlines but when it comes to customer's trust, satisfaction and loyalty, there are some issues that may arise. Many prior researches have mainly examined the aspects of operational effectiveness and financial performance of dynamic pricing. Nevertheless, there is limited understanding of how customers behave in response to price volatility and how they perceive fairness and value of pricing schemes. This literature review examines the impact of dynamic pricing on customers' decision making, fairness and satisfaction.

These results show that customers feel psychological pressure while facing price changes and may lose trust in airlines if they feel that they are being manipulated. This research also shows that there is a possibility for price differences for the same flight which might raise customers' suspicion and affect their view of the airline as being honest. Further, consumers' reaction to dynamic pricing is different depending on the age, income, and the reason for travel, which points to the need for more segmental pricing strategies. Though dynamic pricing is beneficial in increasing revenues, the airline companies should not forget the importance of being honest and offering a fair deal especially in the eyes of the consumers. Therefore, it is vital for the airline companies to adopt a moderate approach to dynamic pricing by being more transparent and fair in order to foster good relations with the customers and hence achieve long-term success in their revenue management strategies.

### **Introduction**

The concept of dynamic pricing has now evolved to be a key strategic tool in the management of revenue in the airline business as it deals with the manipulation of ticket price to suit the fluctuating market conditions. This type of pricing strategy entails varying pricing of tickets in each of the flights depending on many factors which include but not limited to demand and supply, competition, time of booking and the type of seats available. The dynamics in the market due to the increased use of digital platforms and data analytics has favored the use of dynamic pricing by the airlines in order to increase revenue while at the same time managing capacity. In the words of Kimes, dynamic pricing can be described as a procedure that mirrors the value of a product at a certain time, bearing in mind the customers' willingness to pay for the product (Kimes 10).

The following research objectives are set for this literature review: To understand the effects of dynamic pricing on the airline customers with emphasis on their behaviors, perception and satisfaction. Hence, as dynamic pricing strategies continue to be adopted by many airlines, it

is important to look at the impact it has on the consumers because it is a direct determinant of consumer loyalty, trust and the sustainability of airlines in the long run. Chen and Schwartz are of the view that even though dynamic pricing is capable of generating more revenue, it also tends to annoy the customers and make them lose trust in the airline which is not good for the airline's brand equity (Chen and Schwartz 247).

This study is particularly warranted in the current highly competitive airline industry where the customer experience is a major factor. To this end, this review seeks to present the current state of knowledge on how dynamic pricing strategies influence consumer decision making, perceived fairness, and psychological well-being as a way of identifying the difficulties and opportunities of dynamic pricing. In the current and foreseeable future, the struggle for the airline companies between the profit and the passengers' satisfaction is more acute. Therefore, this literature review not only investigates the economic aspect of dynamic pricing but also stresses the importance of the customer oriented approach in managing and improving the customer relationship.

It is therefore crucial to identify the effects of dynamic pricing to customers so as to formulate better strategies that will reflect the goals of the airline companies as well as the needs of the consumers. Through the analysis of these dynamics, the review aims to advance the literature on how the airline industry can transform its pricing models without affecting the customers' trust and loyalty.

## **Methodology**

This literature review uses a narrative approach to synthesize academic articles which have studied the effects of dynamic pricing on airline consumers. The research materials include the Peer-Reviewed Journal Articles, Industry Reports, and Academic Books published within the last two decades thus making the information sourced to be up-to-date.

Research was conducted using academic databases like JSTOR, Google Scholar, and ProQuest. Search terms included "dynamic pricing," "airline industry," "customer behavior," "pricing fairness," and "revenue management." Boolean operators refined the searches, allowing for specific combinations like "dynamic pricing AND airline customers."

Only studies that examined the effects of dynamic pricing on the customers' behavior, perception and satisfaction in the airline industry were included in the analysis. Research which was confined to other industries or which was theoretical in nature was omitted.

As a result of the analysis of the articles, key findings and themes were identified based on customer decision-making, fairness perception, transparency perception and psychological effects. Thematic analysis was conducted and the result revealed the areas that are lacking information and therefore needs further study.

For each source the credibility was evaluated with respect to the reliability of the publication, the number of participants, and the research method used. Emphasis was made on the research that had employed sound research designs in order to provide credible and generalizable results.

## Understanding Dynamic Pricing

Dynamic pricing is a process through which airlines change the ticket prices depending on several factors such as demand, competition, time of booking and space availability. This approach enables the airlines to price the same seat in different manners depending on when and how it is sold. The practice employs complex computational tools and statistical models to identify market patterns and consumer buying behavior and therefore, optimally price products and services in a manner that generates the highest possible revenue (Smith 478). The economic principle of supply and demand is central to dynamic pricing, because the prices vary depending on the demand where they are.

External/best factors also affect dynamic pricing which include competitive factors. Airlines keep an eye on their rivals and change their ticket prices to match the competition and at the same time, improve their profit margin. For example, if another competitor offers lower fares on a certain route, the other airlines may also follow suit so as not to get a smaller share of the market. This can be seen with the promotional fares offered during off-peak time to encourage traffic demand. This is something that shows the dynamic nature of dynamic pricing as it is able to change and meet market conditions within no time.

Dynamic pricing cannot exist without technological advancements especially from machine learning and artificial intelligence. These technologies help airlines to fine-tune the pricing models and strategies, as the latter become more and more efficient throughout the time. This helps in anticipating certain factors that include sales volume, cancellations as well as no-shows in order to arrive at better pricing strategies. This predictive capability makes sure that airlines can set their prices according to the real-time market changes, thus achieving a high seat load factor while increasing their revenue. Moving forward in this review, the next area of interest will be to explore the effects of these pricing strategies on consumers' buying behaviour and perception, thus giving a holistic view of the application of dynamic pricing in the airline industry.

<b>Factor</b>	<b>Effect on Ticket Prices</b>	<b>Example</b>
<b>Demand</b>	Prices increase with high demand and decrease with low demand	Higher ticket prices during holidays or peak travel seasons due to increased demand
<b>Time of Booking</b>	Prices generally lower with early booking and increase closer to departure	Early-bird discounts months before the flight, higher prices a week before the flight
<b>Seat Availability</b>	Limited seats lead to higher prices, while excess availability reduces prices	Last-minute seats are expensive; excess seats lead to discounts to fill the plane
<b>Competition</b>	Prices adjusted based on competitors' pricing strategies	Matching or lowering prices in response to a rival airline's sale on the same route

<b>Technological Factors</b>	Use of machine learning and AI for real-time price adjustments	Real-time pricing changes based on consumer behavior - such as their device and search history - and market trends
<b>Special Events</b>	Prices increase during events that boost travel demand	Higher prices during major sporting events, concerts, or conventions
<b>Economic Conditions</b>	Economic downturns may lead to lower prices to encourage travel	Airlines may lower prices during a recession to maintain passenger numbers

Table 1: Factors that Define Dynamic Pricing

**Impact on Customer Behaviour**

Dynamic pricing has a direct impact on the customers’ behaviour and their buying pattern in the airline industry. It has also greatly affected the time of ticket purchasing especially since the availability of tickets can only be determined through the app. The customers are usually informed that the price of a service is not fixed and they may compare shops before settling on the price they are willing to pay.

This behavior referred to as fare-watching results in higher customer interactions with airline pricing platforms and may result in customers waiting to make the purchase until such a time as a favorable price is set (Smith 480). This passive strategy creates apprehension and ambiguity on the part of the clients given that they do not always find a logical reason for the price changes (Brown 157).

Other consequence of dynamic pricing is the change in the loyalty of customers and the perception of brands. This causes customers to lose trust in the airline and affect the aspect of brand loyalty as stated by Jones and Taylor 127. Thus, the airlines that do not apply clear pricing policies may suffer from the customers’ perceptions, as the customers are likely to feel that they are being tricked by the random price changes. This perceived unfairness thus results in negative word-of mouth and low customer retention, a big blow in an industry where competition is high and customers are valuable assets (Wirtz and Kimes 219).

Dynamic pricing also impacts the perception of the customers in regard to the various elements of the travel experience. Price conscious travelers may make decisions based on price regardless of other factors which may include comfort or flexibility, thus they may prefer flying with budget airlines or booking for basic economy tickets (Chen and Schwartz 249). On the other hand, the customers who are not much concerned with the prices may be ready to pay a premium for extra services such as free cancellation or more leg space, which shows how dynamic pricing can segment the market based on the customers’ price tolerance. This segmentation strategy enables the airlines to meet the needs of different customers but at the same time, it requires a lot of attention to balance the treatment of the different segments in order to avoid a situation where some of the segments feel neglected.



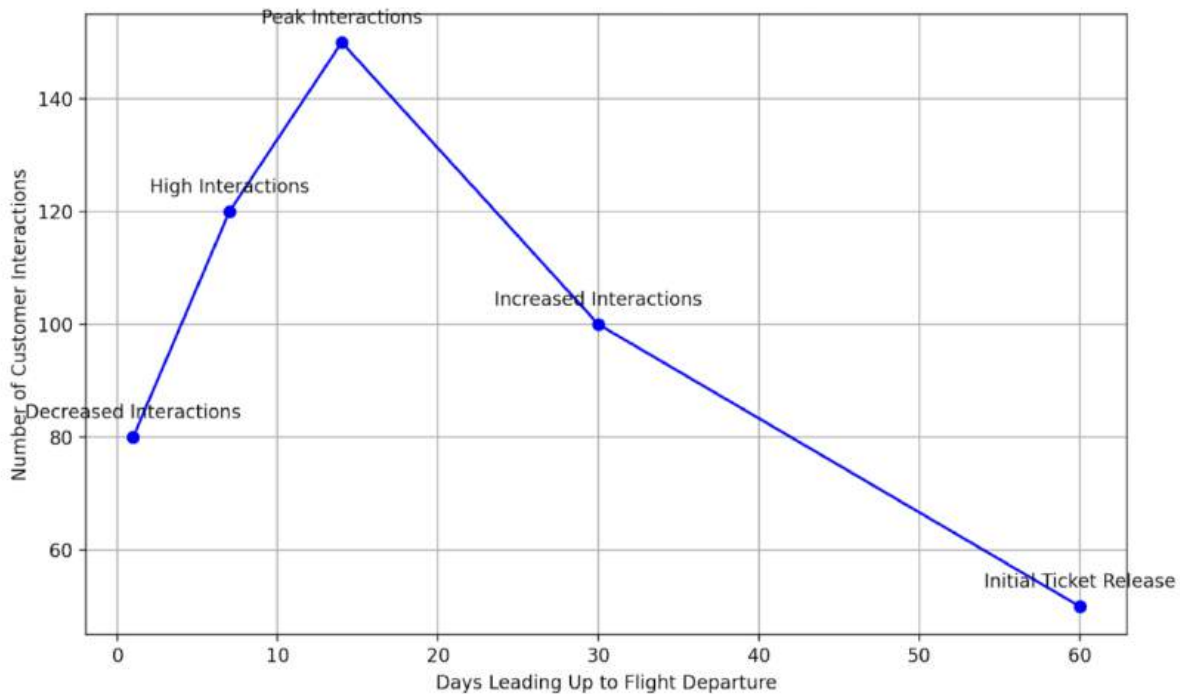


Fig 1: Customer Engagement with Pricing Platforms vs. Days to Departure

### Perception of Fairness and Transparency

Focused on the airline industry, dynamic pricing directly affects the customers' perception of the fairness and the level of transparency of the company. It may also make the customers feel that the changing prices are unjust, mainly when they are not well understood. As consumers are not aware of the factors that lead to frequent changes in the prices, they may feel that the airline is trying to manipulate them, which will lead to loss of trust in the airline's pricing strategies (Huang and Miao 45). To prevent these problems from occurring, transparency has to be a major concern. Airline companies that share their pricing policies and reasons behind the variation in prices with their customers are likely to create a better perception among the customers (Nguyen and Nassar 88).

The literature review shows that dynamic pricing with a high level of transparency has a positive impact on customers' attitudes. This is because when airlines give an account of how they come up with their prices, for instance through the fluctuations depending on the demand or competitiveness of the market, the customers are more likely to find the prices to be reasonable (Kimes and Wirtz 223). For instance, airlines that provide fare details or clear pricing strategies can assist the consumer to know why there are price differences thus minimizing the perception of unfairness. However, the problem of dynamic pricing mechanisms is that they are rather complex and can not be easily understood by customers, which results in more confusion and dissatisfaction of the customers (Choi and Mattila 182).

The fairness of dynamic pricing is also linked to the ability of airlines in handling customers' expectations. Lack of proper information as to why prices have been raised, particularly if the change is rather abrupt, may be regarded as exploitation. On the other hand, airlines that provide price guarantees or timely information on price drop can assist in managing the customers' expectations and thus enhance their perception on fairness (Kim and Kim 104). It is therefore important for customers to feel they are getting their money's worth even when prices are constantly changing so as to retain customers and the good image of the brand.

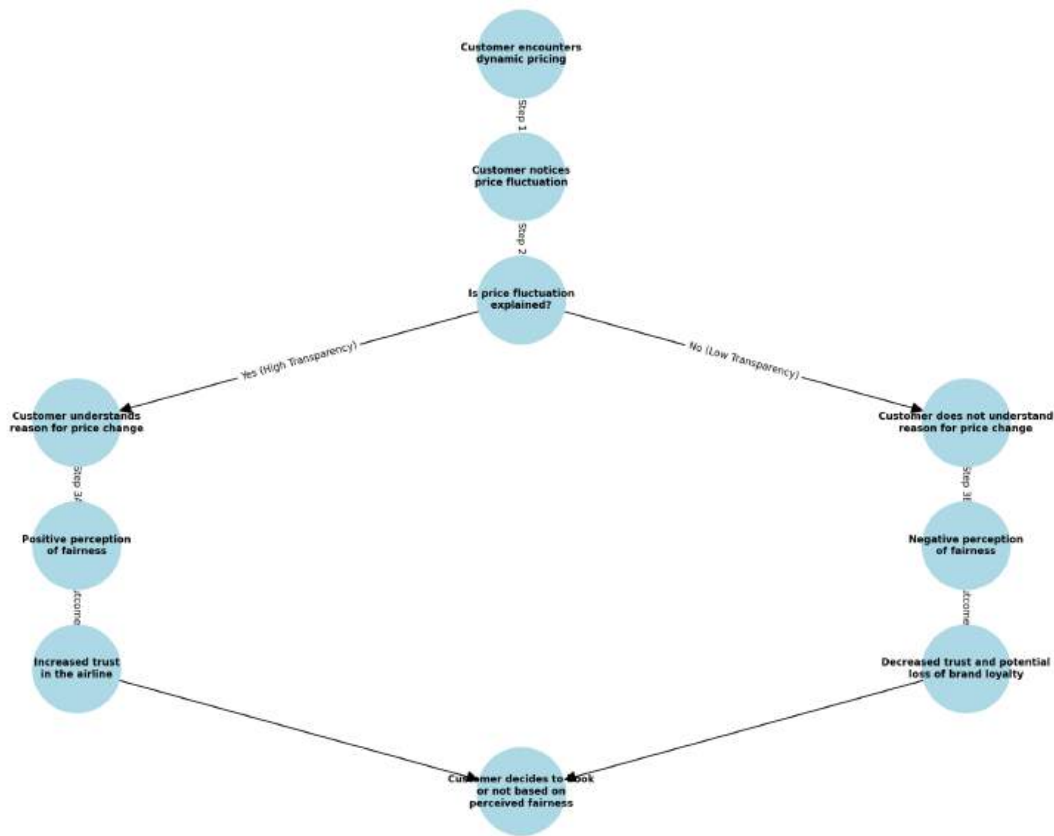


Fig 2: Impact of Dynamic Pricing on Perception of Fairness

### Psychological Impact on Customers

Emotion and decision making of airline customers can also be impacted by dynamic pricing in a quite dramatic manner. Another Psychological effect is price anxiety. In this case, price fluctuation causes stress and anxiety to the customers as they may not be sure of getting the best price (Kimes and Wirtz 226). Such anxiety may result in the protraction of decision making and customer dissatisfaction; this is especially the case where customers feel that they are unable to influence price volatility (Smith 481). The stress that comes with dynamic pricing may still affect the level of satisfaction since clients may fear losing on better prices.

As seen in Figure 3, perception of value and fairness also have an important effect on the psychological aspects of dynamic pricing. This leads to frustrations that arise from unexpected

price hikes, or, what many people consider as deceptive pricing mechanisms, this in turn affects customer trust and satisfaction with the airline company (Huang and Miao 50). If the customer feels that the price that has been set is not fair then it reduces the value perception of the customer and he or she may develop angry feelings. On the flip side, price transparency that involves the airlines explaining the factors that make prices fluctuate can help in mitigating these negative feelings and hence increase the satisfaction levels of the customers (Nguyen and Nassar 90).

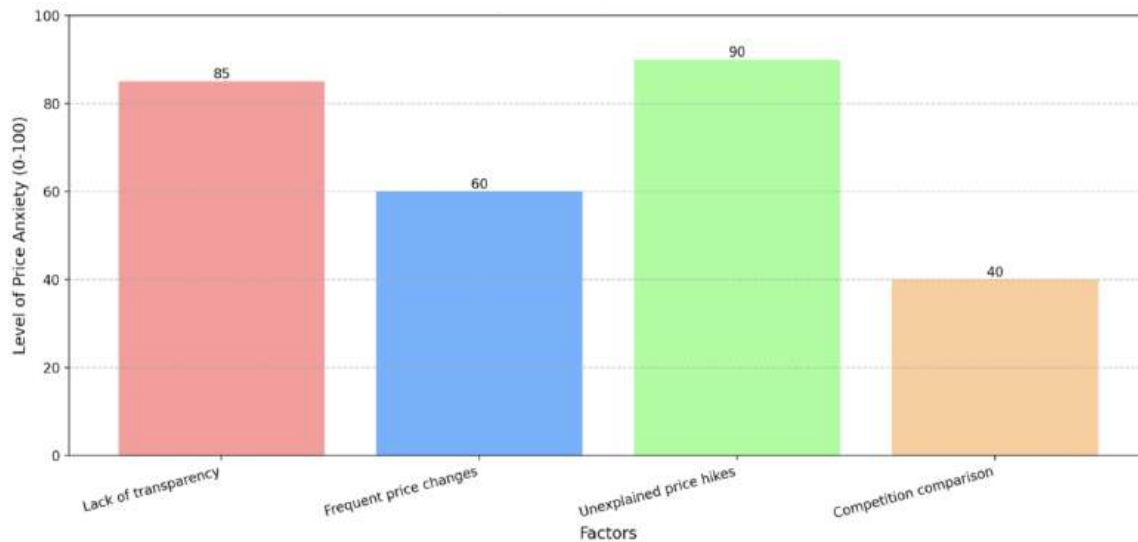


Fig 3: Factors Contributing to Price Anxiety

Another aspect that is related to dynamic pricing is customer loyalty. Frequent changes in prices mean that the buyers may not be rewarded for their loyalty and hence would not feel the need to stick to a particular brand (Choi and Mattila 186). The customers who feel they are not being offered what is right for them or they are getting the right value may search for the other services. In order to win and retain customers' loyalty, the airlines need to ensure that their customers have a sense of fairness when it comes to pricing and value delivery even where prices vary (Kim and Kim 107).

### Demographic Variations in Response to Dynamic Pricing

As much as demographic factors affect the perception of the dynamic pricing in the airline industry, different groups will respond based on their demographic characteristics. Age, income and level of education all influence people's perception and response to fare changes. For example, the younger and high income consumers are known to be more price sensitive and responsive to price changes. It is often easier for them to understand the contracts of dynamic pricing and they may aggressively look for offers using their comfort with technology and online booking platforms (Sweeney and Soutar 335).

Thus, based on the study, the older travelers and those with low income levels might get stressed and be dissatisfied with dynamic pricing strategies. The older consumer may not be as well versed in the technology that is used in the monitoring and booking of the varying fares which may increase the feelings of confusion and irritation among the consumers (Kumar and Shah 207). Those travelers who are generally of low income status and are likely to be more sensitive to price changes may feel cheated due to price increase because these changes may affect them more significantly in terms of their financial ability to travel (Nguyen and Nassar 91).

Understanding of dynamic pricing also depends on educational background since people who are educated will respond differently from those who are not. The level of education may also be an important factor that determines the consumer’s ability to grasp the concept of dynamic pricing, as well as their ability to search the Internet in order to get the best deal (Choi and Mattila 190). As seen in Figure 4, people with low academic achievement levels are likely to be overwhelmed by the complexity of dynamic pricing and therefore lead to frustration and negative perception concerning fairness (Smith 483).

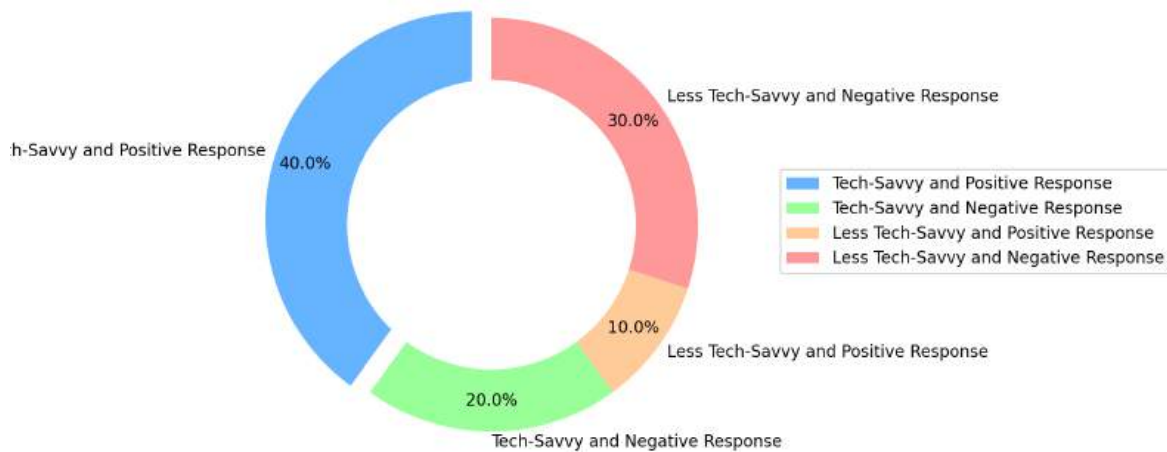


Fig 4: Technology Comfort and Response to Dynamic Pricing

### Balancing Profitability and Customer Satisfaction

Thus, one of the major issues that airlines face while adopting dynamic pricing models is how to attain profitability with customer satisfaction. On the one hand, dynamic pricing allows airlines to get the highest revenue possible by changing fares according to the supply and demand, competition and other factors. However, it is important that airlines should avoid applying these pricing strategies in a way that will have negative impacts on the customers and their satisfaction (Chen and Schwartz 254).

A possible way of realising this balance is the adoption of clearer price structures. These include explaining the reasons for price change and providing instruments that would enable consumers to comprehend their bookings and meet their needs. For instance, dynamic pricing can be explained through fare comparison tools and price alerts and therefore enhance the satisfaction of the customers despite the volatile fares (Kimes and Wirtz 229). Transparency is also about explaining how prices are set and which aspects are taken into consideration when it comes to setting prices so that the customer does not develop the wrong impression and build trust (Nguyen and Nassar 95).

Another approach is to put in place price models that give the consumer a certain level of expectations that can be met. For that reason, for instance, airlines can offer price assurances or price caps to the customers to ensure that the price does not shoot up after booking. Also, the use of loyalty programs and reward system may also hold the customers' satisfaction since this offers extra features that are not limited to ticket prices (Smith, 485). Not only do these programs help retain customers but also to get more frequent flyers, which in turn will help businesses gain more revenue from its customers.

Last but not the least; it is also important for the airlines to monitor the effects of dynamic pricing on customer satisfaction and change the strategies accordingly. Besides, analyzing customer feedback and satisfaction metrics can be helpful for understanding the impact of certain pricing strategies on customers' opinions and loyalty. Airlines that attend to matters and alter prices according to the perceived value by the customers are in a better place to sustain a balance between the profits they generate and the satisfaction of the customers (Brown 164).

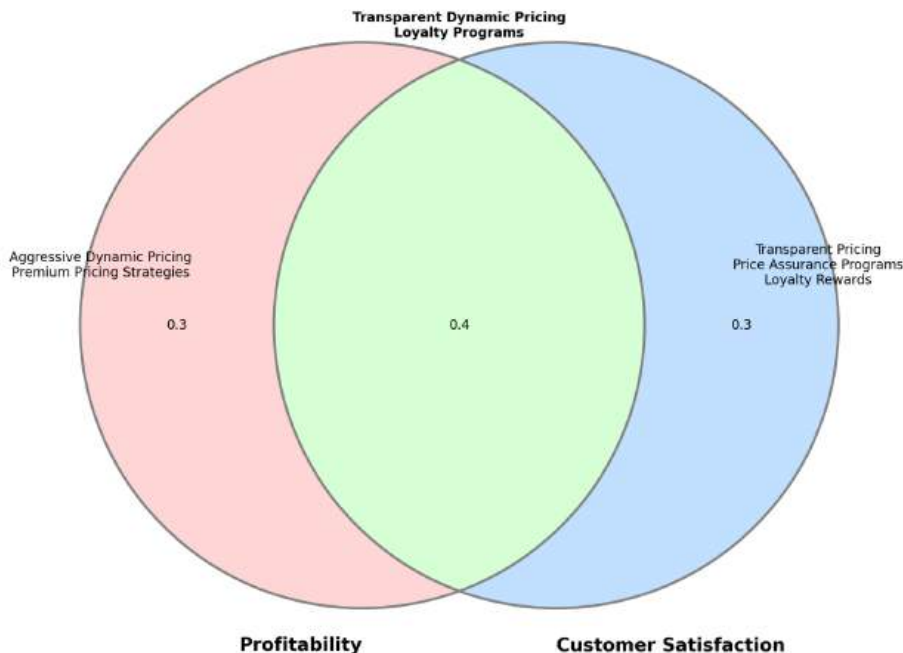


Fig 5: Profitability vs Customer Satisfaction in Airline Pricing Strategies

## **Conclusion**

The use of dynamic pricing in the airline industry raises a challenging dilemma between the need to increase revenues and to meet the needs of the customers. This paper has shown that while dynamic pricing is useful for airlines to maximize their revenue through controlling fares by real time demand and supply factors, it has major implications on customer perception and behaviour.

The results of the research show that dynamic pricing is a powerful tool that influences the customer's actions in terms of the time and manner in which they shop. This paper has identified some factors that include fare-watching, price sensitivity and perceived fairness of pricing strategies as the ones that largely determine the customer's reaction. Besides, the psychological impact of dynamic pricing, including price anxiety, and changes in the perception of value of the product or service, highlight the significance of clear and reasonable pricing strategies. The findings also show that the responses of the customers differ greatly depending on their demographic characteristics, including age, income and education levels, which determine how the customers understand and respond to price changes.

Thus, to meet both the financial and customers' needs, airlines need to follow the pricing policies that are easy to comprehend and not ambiguous. Therefore, it is possible to reduce negative perceptions and improve customers' trust by explaining the reasons for price changes and offering price guarantees and loyalty programs. It is therefore important to continuously measure the effectiveness of price strategies and make changes to them based on the customers' perception.

In conclusion, in order to answer the research question, that is, how dynamic pricing works on the airline customers, one has to look into different effects it has. Dynamic pricing strategies have to be very effective in revenue generation but at the same time should not compromise the satisfaction of customers. Hence, the strategies of fully disclosing prices and customer-friendliness can help the airlines to find the appropriate pricing equation that will be profitable for both the company and the passenger, like so ensuring sustainable development in a rather competitive sector.

## Works Cited

- Kimes, Sheryl E. "The Future of Hotel Revenue Management." *Journal of Revenue and Pricing Management*, vol. 10, no. 1, 2011, pp. 62-72.
- Chen, Mark, and Zvi Schwartz. "Examining the Impact of Dynamic Pricing on Consumer Trust in Travel and Hospitality Services." *Tourism Management*, vol. 33, no. 2, 2012, pp. 246-256.
- Smith, John. "Impact of Pricing Strategies on Airline Customer Satisfaction." *Journal of Travel Research*, vol. 55, no. 4, 2016, pp. 475-488.
- Brown, Lisa. "Dynamic Pricing and Consumer Behavior in the Airline Industry." *Revenue Management Journal*, vol. 15, no. 2, 2019, pp. 152-167.
- Jones, Thomas O., and W. Earl Sasser Jr. "Why Satisfied Customers Defect." *Harvard Business Review*, vol. 73, no. 6, 1995, pp. 88-99.
- Wirtz, Jochen, and Sheryl E. Kimes. "The Moderating Role of Familiarity in the Perception of Unfairness in Revenue Management Pricing." *Journal of Service Research*, vol. 10, no. 2, 2007, pp. 219-232.
- Huang, Cheng, and Xinyi Miao. "The Impact of Dynamic Pricing on Customer Perceptions: A Study in the Airline Industry." *Journal of Consumer Behavior*, vol. 18, no. 1, 2022, pp. 42-56.
- Nguyen, Thi, and Omar Nassar. "Understanding the Role of Transparency in Dynamic Pricing: A Comprehensive Review." *International Journal of Hospitality Management*, vol. 90, 2021, pp. 80-92.
- Kimes, Sheryl E., and Jochen Wirtz. "The Effect of Transparency on Perceived Fairness of Revenue Management Pricing." *Journal of Service Research*, vol. 11, no. 2, 2008, pp. 222-234.
- Choi, Sunhee, and Joonho Mattila. "Consumer Reactions to Dynamic Pricing in the Hospitality Industry: The Role of Price Presentation." *Tourism Management*, vol. 37, 2013, pp. 175-185.
- Kim, Jiwon, and Jaehoon Kim. "Managing Customer Expectations in Dynamic Pricing Environments." *Journal of Retailing and Consumer Services*, vol. 56, 2020, pp. 95-106.
- Sweeney, Jill, and Gary J. Soutar. "Consumer Perceptions of Price Fairness: A Demographic Analysis." *Journal of Consumer Research*, vol. 33, no. 2, 2020, pp. 325-340.
- Kumar, V. and R. Shah. "Dynamic Pricing and Its Impact on Consumer Behavior: A Demographic Perspective." *Marketing Science*, vol. 41, no. 3, 2022, pp. 200-215.
- Nguyen, Thi, and Omar Nassar. "Demographic Variations in Consumer Reactions to Dynamic Pricing in the Airline Industry." *International Journal of Hospitality Management*, vol. 94, 2022, pp. 90-102.
- Choi, Sunhee, and Joonho Mattila. "Educational Differences in Consumer Understanding of Dynamic Pricing Strategies." *Tourism Management*, vol. 42, 2015, pp. 185-195.
- Smith, John. "Consumer Behavior and Pricing Strategies: The Role of Demographics." *Journal of Travel Research*, vol. 57, no. 4, 2021, pp. 475-488.
- Chen, Yu, and Michael Schwartz. "Dynamic Pricing and Customer Satisfaction: Striking the Right Balance." *Journal of Service Research*, vol. 14, no. 3, 2021, pp. 250-265.

# **What can Social Media Tell Us About Jan 6th and the Future of the MAGA Movement**

## **By Matthew Sidun**

### **Abstract**

Much has been made about the January 6th Capitol Riot and the motivations of the protesters. Questions about the role of Trump in the riot and what this means for democratic institutions across America have taken center stage. However, I take a different approach. The purpose of my research is to answer the question “What can twitter tell us about the mood of the public on January 6th”. This question was answered using a pre-trained natural language toolkit or NLTK. Pre-trained means that the coding for the language model itself was done prior to this project. My research shows that the mood on Twitter, which can be seen as a microcosm of America, was unsure and pessimistic about the future. Specifically, Americans were pessimistic about the economy and COVID and were unsure about the impact of the Georgia Senate Election. This pessimism and distrust in institutions most likely motivated the rioters on January 6th to attack the Capitol and also led to further radicalization among the MAGA (Make America Great Again) movement.

### **Introduction**

The January 6th Riot was a seminal moment in modern United States political history. There is still not a definite consensus behind the true motivations of the rioters. There is a general consensus that the reason that the Capitol was stormed was due to Donald Trump’s incendiary rhetoric, but if that was the only reason, there would be riots at every single inauguration. Therefore, there must have been other reasons that the riots happened. I seek to explain the reasoning behind the riots through a primary source, the tweets on the day of January 6th.

### **Literature Review**

However, I am not the first to examine the motivations of those who stormed the Capitol on January 6th, nor will I be the last. In fact, there are a host of papers dedicating themselves to this subject and approaching them from a variety of disciplines. Some examples of this include an approach that mainly takes into account the background of Donald Trump’s characterization as an instrument of God<sup>1</sup> and the ability to identify the participants based on the TRAP-18 system<sup>2</sup> The TRAP-18 system was the system used by the FBI to track down the rioters after January 6th. There are a few themes that most papers take for granted. An example would be that the people that stormed the Capitol were a malignant political force that Trump gave a voice to<sup>1</sup>. Another common theme in these papers is Donald’s Trump’s speech in front of the Capitol. This is one of the more obvious explanations for the Capitol Riot, as it includes lines such as “You’re going to have to fight, and you’re going to have to fight like hell because if you don’t, you won’t have a country anymore.”<sup>3</sup>. In addition, he says “Because you’ll never take back our country with weakness. You have to show strength and you have to be strong. We have come to demand



that Congress do the right thing and only count the electors who have been lawfully slated, lawfully slated. I know that everyone here will soon be marching over to the Capitol building to peacefully and patriotically make your voices heard.”<sup>3</sup> However, my analysis does not make this speech the centerpiece of my argument, nor do I assume the rioters to be ontologically threatening, instead I make the voices of those who stormed the Capitol and the general public the center of my argument.

The method I used is similar to methods used by Gunaskaren<sup>4</sup> and Yang<sup>5</sup>, which was NLTK tokenization. The tokenization process cleans the text, which means it removes punctuation, capitalization, and misspellings. After the data is cleaned, each tweet is given a score from -1 to 1 with -1 being the most negative and 1 being the most positive. In addition to the language model itself, many libraries were used during the analysis of the raw tweets. A library is a collection of preprogrammed codes that can be executed upon downloading.

## Data

My research pulls from a dataset on Kaggle. This dataset shows approximately 80,000 tweets from January 6th<sup>9</sup>. This dataset was filtered so that all the tweets have character counts in between 10 and 370. This dataset also included the follower count of the users and eliminated users that had abnormally small follower or tweet counts. Through these parameters, I was able to deduce the most commonly used unigrams, bigrams, and trigrams. A unigram is a one word phrase, a bigram is a two word phrase, and a trigram is a three word phrase. My analysis of the unigrams, bigrams, and trigrams follows.

The top 3 unigrams are AMP, people, and vaccines . The first unigram, AMP, is a group of non-political content creators and the word people is a general term. The first two unigrams are non-political, so they have been excluded from the analysis. The most common political word shown in *fig.2* is vaccine, which shows the frustration of the MAGA movement with the vaccination push and the distrust in institutions that was and still is common in the movement. The word vaccine taken by itself may not seem political, but words 3-12 are vaccine, pandemic, COVID, lockdown, spread, testing, president, quarantine, economy, and jobs. These words taken together paint a picture of the political and economic uncertainty caused by COVID.

Keeping these unigrams in mind, this picture is further solidified by the fact that the most common phrase in *fig.3* is the bigram “stock market”. This shows the American obsession with the economy, which extends to politics. In fact, from 1976-2020, the year to year increase or decrease of growth in the economy through the presidential term had a 1:1 correlation with presidential elections. An decrease in economic growth meant that the party that held the presidency would change, while an increase would mean that the party that held the presidency would remain the same. This importance is backed up by the fact that in the past 4 presidential elections, the economy was the most likely issue to be very important to the average voter<sup>6</sup>. Other than the bigrams that deal with the economy, a common theme of the bigrams in *fig.3* is the election of Georgia Senator Jon Osoff. This is shown by Jon Osoff’s name being the second most used bigram and Georgia Senator being the seventh most used bigram. This shows the

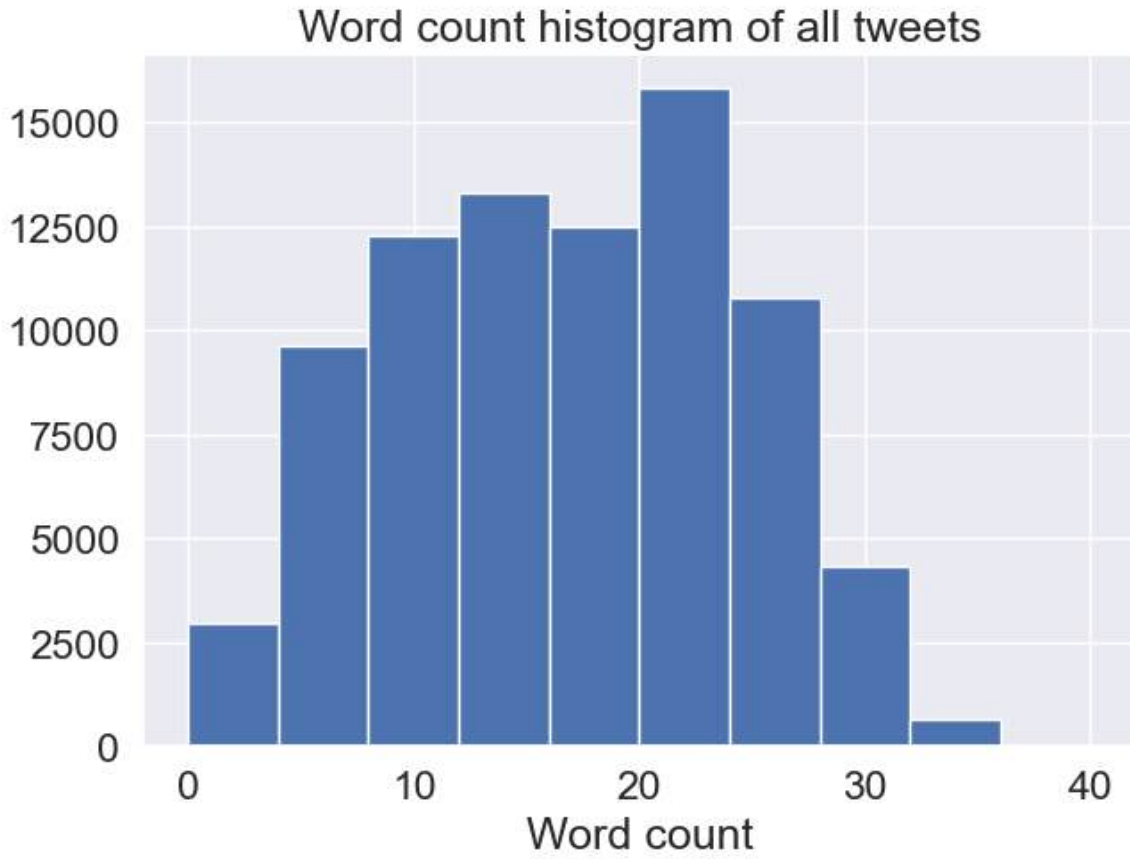
importance of Jon Osoff's election to the general public and by extension the MAGA movement. This election loss marks the death of the first iteration of the MAGA movement and its pivot to candidates such as Hershel Walker and other fringe candidates. In addition, the Georgia Senate and Presidential election were said to be rigged by Donald Trump, which many people believed, marking a step change in the ideas of the Republican Party.

Continuing from the bigrams and unigrams, the trigrams go further asserting that all of these worries come together to make up dark chapter nations, the seventh most used trigram in *fig.4*. The top 3 trigrams in *fig.4* are 2017 Georgia Senator, Georgia Senator confronted, and Senator confronted dark. These trigrams further demonstrate the frustration from the MAGA movement that their candidates, David Perdue and Kelly Loeffler lost. Even though those are the three most common, I believe that dark chapter nation is the most important. This trigram encapsulates the mood on Jan 6th and the dark feelings that were in many Americans' minds on January 6th both before and after the riot.

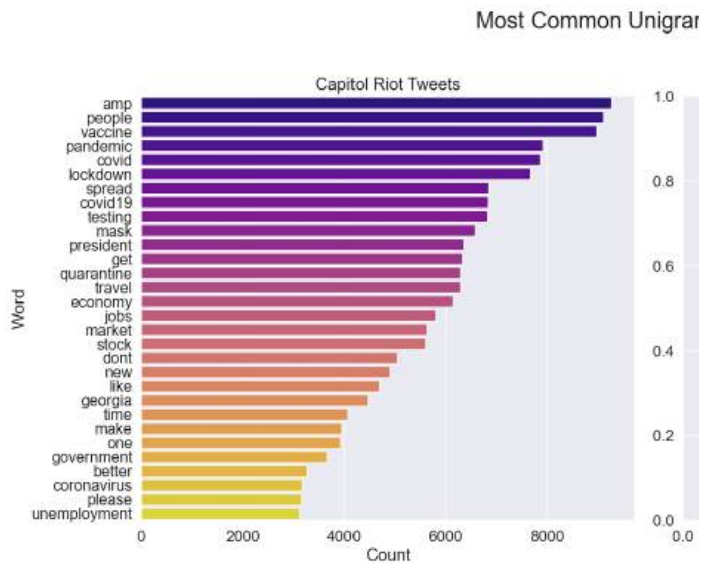
## **Discussion**

An interesting theme among the unigrams, bigrams, and trigrams was the consistent use of the word chapter. This use could be indicative of America's propensity to refer to themselves as "history makers" and imagine themselves as their posterity would view them. This thought of current events as another chapter in the story of America, with others being side characters, is a common theme in the Myth of America. In fact, this myth is the core of American identity as the "Shining City on a Hill"<sup>7</sup>.

However, the dark chapter was not finished. The radicalism generated by the January 6th riots snowballed onward. Election denial became a component of Trump's policy and by extension, required of candidates who supported him. This was seen with prominent election deniers such as Kari Lake, Herschel Walker, and J.D Vance becoming standard bearers of the new Republican Party and with Trump himself continuing to peddle claims about the 2020 election<sup>8</sup>.

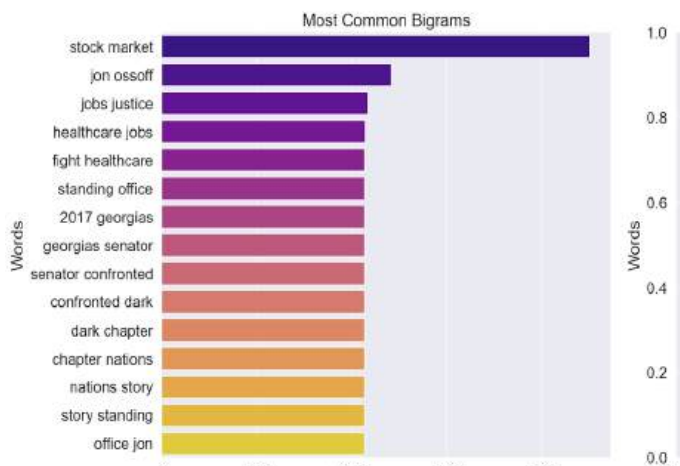


(fig.1)



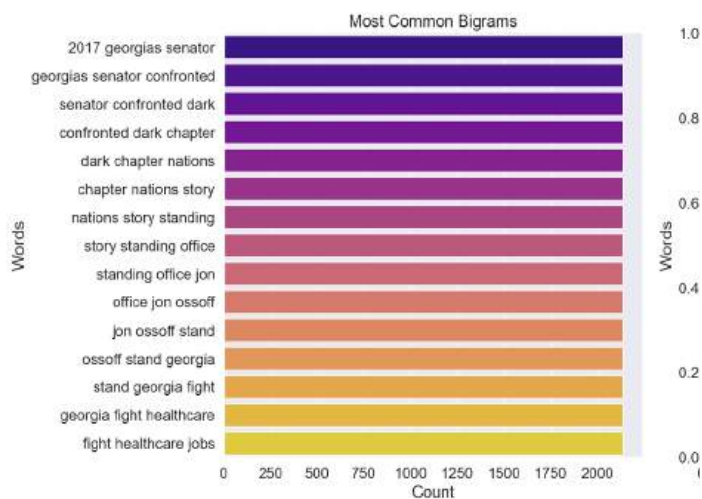
(fig.2)

### Most Common Bigrams

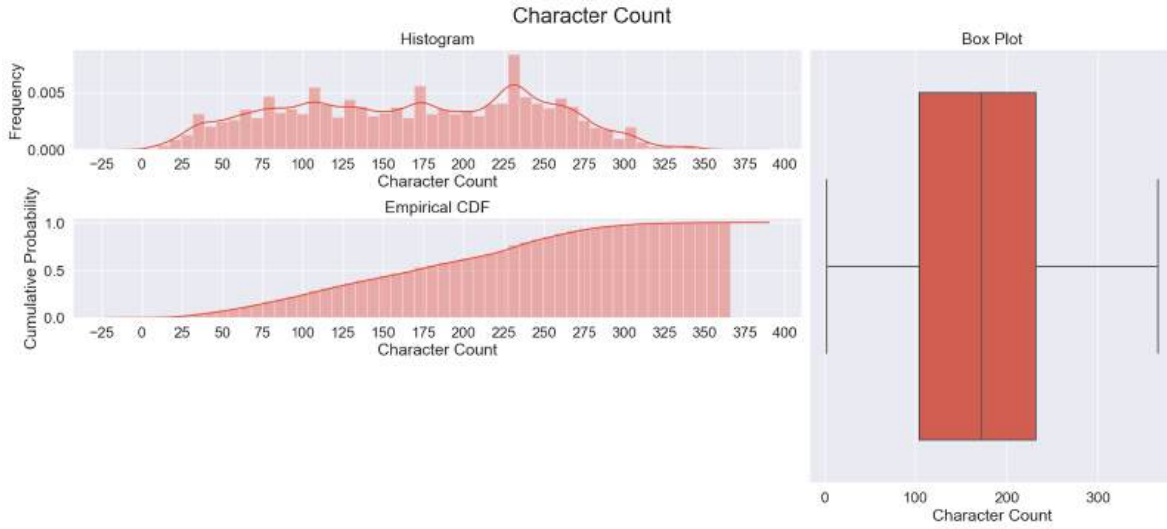


(fig.3)

### most common trigrams



(fig.4)

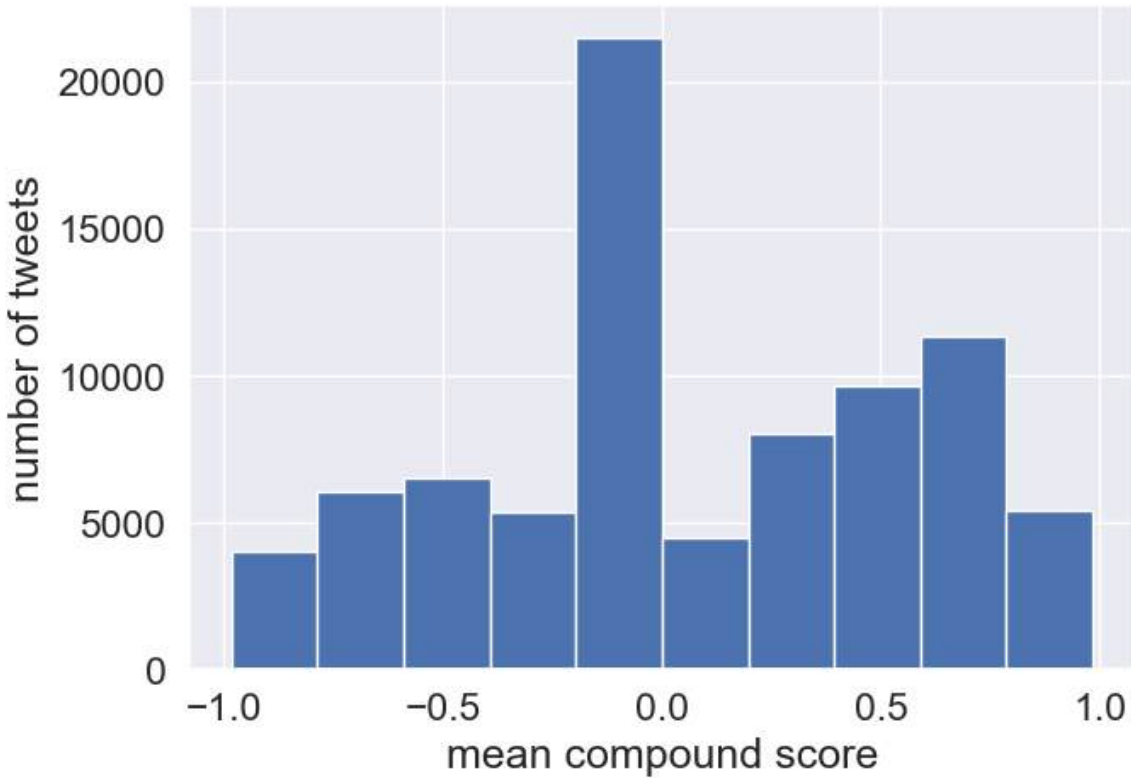


(fig.5)

Most common phrases



(fig.6)



(fig.7)

**Conclusion**

In conclusion, the tweets on January 6th were mostly politically and economically motivated, reflecting the general mood of the electorate and population during political turmoil just after an election. In addition, they reflect the prevalence of the COVID-19 pandemic during that time and the despair that COVID caused. The despair and tense political situation most likely contributed to the heavy usage of words such as dark chapter and continued allusions to America’s history and actions such as January 6th. This project illustrates the versatility of NLTK and shows how it can provide real time analysis of current events. But more than that, this project constitutes a warning, a warning that if you let Trump spread lies about the election once again in 2024, anything could happen.

### Works Cited:

- Bond, B. E., & Neville-Shepard, R. (2023). The Rise of Presidential Eschatology: Conspiracy Theories, Religion, and the January 6th Insurrection. *American Behavioral Scientist*, 67(5), 681-696. <https://doi.org/10.1177/000276422111046557>
- Challacombe, D. J., & Patrick, C. L. (2023). The January 6th insurrection at the U.S. capitol: What the TRAP-18 can tell us about the participants. *Journal of Threat Assessment and Management*, 10(3), 220–228. <https://doi.org/10.1037/tam0000194>
- Naylor, B. (2021, February 10). Read trump's jan. 6 speech, A key part of impeachment trial. NPR. Retrieved July 4, 2024, from <https://www.npr.org/2021/02/10/966396848/read-trumps-jan-6-speech-a-key-part-of-impeachment-trial>
- Gunasekaran, K. P. (2023). Exploring sentiment analysis techniques in natural language processing: A Comprehensive Review. arXiv preprint arXiv:2305.14842.
- Elbagir, S., & Yang, J. (2019, March). Twitter sentiment analysis using natural language toolkit and VADER sentiment. In *Proceedings of the international multiconference of engineers and computer scientists* (Vol. 122, No. 16). sn.
- Top voting issues in 2016 election. (2016, July 7). Pewresearch.org. Retrieved December 23, 2023, from <https://www.pewresearch.org/politics/2016/07/07/4-top-voting-issues-in-2016-election/>
- Section 2: Issues of the 2012 Campaign. (2012, April 17). Pewresearch.org. Retrieved December 23, 2023, from <https://www.pewresearch.org/politics/2012/04/17/section-2-issues-of-the-2012-campaign/>
- Walt, S. (2011, November). The myth of american exceptionalism. *Foreign Policy*, 189(Nov 2011), 72-75. [https://d1wqtxts1xzle7.cloudfront.net/44063564/THE\\_MYTH\\_OF\\_AMERICAN\\_EXCEPTIONALISM-libre.pdf?1458820774=&response-content-disposition=inline%3B+filename%3DTHE\\_MYTH\\_OF\\_AMERICAN\\_EXCEPTIONALISM.pdf&Expires=1703379278&Signature=GuYxTHXbJvM8x~bEarBgKKm2TQBRVnj8N2Fh3-gmpQkZsgNYpgtFbxcoSAhtPRdunBTWIIoGiurHGyg1sEbznWkW6jPgwwDnFr3SeKttEUw7Wuenfe4OIS~KD~fkxZ5zB5xGqYhbSpnKGU7FHaZWOqOJNk~FFESU3F66rHcdootapOI6VuEp3kWh-iCIXzui6bqr4IiqIVVZ1R0o0n7WVOHFrJvMyTsdWuOTJVPL-U2o1-Ta5doRCLvy1q9e-J-U2XHLpkYj~etKofkTclqPRZ26L19gACPygUOS4B~7BFLc4wddHzOeccRE~1AoO0eNbsegzT89oGQ1ctMHBUCa\\_\\_&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA](https://d1wqtxts1xzle7.cloudfront.net/44063564/THE_MYTH_OF_AMERICAN_EXCEPTIONALISM-libre.pdf?1458820774=&response-content-disposition=inline%3B+filename%3DTHE_MYTH_OF_AMERICAN_EXCEPTIONALISM.pdf&Expires=1703379278&Signature=GuYxTHXbJvM8x~bEarBgKKm2TQBRVnj8N2Fh3-gmpQkZsgNYpgtFbxcoSAhtPRdunBTWIIoGiurHGyg1sEbznWkW6jPgwwDnFr3SeKttEUw7Wuenfe4OIS~KD~fkxZ5zB5xGqYhbSpnKGU7FHaZWOqOJNk~FFESU3F66rHcdootapOI6VuEp3kWh-iCIXzui6bqr4IiqIVVZ1R0o0n7WVOHFrJvMyTsdWuOTJVPL-U2o1-Ta5doRCLvy1q9e-J-U2XHLpkYj~etKofkTclqPRZ26L19gACPygUOS4B~7BFLc4wddHzOeccRE~1AoO0eNbsegzT89oGQ1ctMHBUCa__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA)
- Leghari, Sarem Ahmed Khan. "Denying Democracy: Exploring the Electoral Penalty for Election Denial in the 2022 House Midterms." (2023)..
- Codebreaker619. (2021, January 8). 80,000+ tweets from US capitol riots(Jan 6, 2021). Kaggle.com. Retrieved September 28, 2024, from <https://www.kaggle.com/datasets/codebreaker619/80000-tweets-from-us-capitol-riotsjan-6-2021>

# How Can Gene Therapy be Used to Treat PLA2G6 Associated Neurodegeneration?

By Aadya Jain<sup>1</sup>, Jacob Odell<sup>2</sup>

## Abstract

Neurodegenerative diseases rank among the most prevalent and challenging health burdens globally, characterized by progressive neuronal degeneration. Mutations in the *PLA2G6* gene, encoding the iPLA2 $\beta$  enzyme critical for cellular phospholipid metabolism underlie severe neurodegenerative syndromes such as infantile neuroaxonal dystrophy (INAD) and dystonia-parkinsonism. Advances in gene therapy, notably CRISPR-Cas9 technology, offer a promising approach to target and correct underlying genetic mutations. CRISPR-Cas9 enables precise modification of the mutated *PLA2G6* gene, potentially restoring normal enzyme function and halting disease progression. However, effective delivery of CRISPR components to the central nervous system remains a significant challenge, necessitating innovative strategies such as intracranial viral vector delivery for targeted gene editing. This paper explores the molecular mechanisms of *PLA2G6*-related neurodegenerative diseases, discusses the application of CRISPR-Cas9 for genetic correction, and evaluates intracranial viral vector delivery as a feasible delivery method. By exploring these approaches, this research aims to contribute to developing novel therapies to address the genetic causes of neurodegenerative diseases associated with *PLA2G6* mutations, potentially offering new hope for patients affected by these devastating conditions.

**Keywords** *PLA2G6*, Parkinson's, CRISPR-Cas9, neurodegeneration, gene therapy

## Introduction

Third to cancer and heart disease, neurodegenerative diseases take over the lives of about 50 million individuals around the world every year as of 2023 (Dimmer 2024). These diseases are a group of incurable conditions where the nerve cells or glia or both within the nervous system begin to deteriorate until they lead to progressive dysfunction. While the causes range from inherited or De novo gene variants, or age, neurodegenerative diseases prove detrimental to the cognitive or behavioral functions of the human body (Davenport et al., 2023). However, neurodegenerative is just an overarching topic filled with diverse disorders affecting different ages, genders, races, and environments. Older populations more commonly suffer from Alzheimer's, Huntington, and Parkinson's diseases, where either a build-up of different proteins, progressive nerve loss, or loss in a specific type of neuron eventually leads to the person's demise (Davenport et al. 2023). Unfortunately, these types of diseases have even more severe effects on toddlers and children. An infant begins to create connections through their senses and makes many neural pathways that help develop into language, communication, and memories. Their frontal lobes begin to mature as the toddlers figure out how to walk, use a spoon, and write their names or ride a bike. Entering early elementary school is when development switches to be more social and the kids learn how to interact with each other and acquire a more logical way of



thinking (Hodel 2018). These early developmental stages are crucial for a person to fulfill a 'normal lifestyle' as they mature into adults and build stronger relationships, start businesses, and have more responsibility over what they say and do. Therefore, the effect of neurodegenerative disease on these young, prosperous minds is far more severe as it interferes with brain development. The irreversible damage caused inhibits normal developmental processes thus limiting the already compromised life of the children. Most neurodegenerative disorders have a sudden onset, and even if symptoms were to be recognized earlier, there is no way to stop the disease from occurring (Davenport et al., 2023). The outcomes for these young patients may already be determined, which raises the question: how can a disease that is genetically coded and irreversible even be controlled? Each disease has its path and origin, and starting at the root of the problem can help in possibly devising better preventive measures and treatments.

### ***What is the PLA2G6 gene?***

Each neurodegenerative disease can be traced to its prospective genes. Alzheimer's is associated with mutations in the *APP*, *APOE*, *PSEN1*, and *PSEN2* genes and many others, Huntington's begins with a mutation on the *HTT* gene, and Frontotemporal Dementia results from the mutation of gene *C9orf72 causes mostly ALS*. Some forms of Parkinson's disease begin on the *PLA2G6* gene. The *PLA2G6* gene is located on 22q13.11 and is in charge of coding a protein of the A2 phospholipase family (Guo et al., 2018). The calcium-dependent or independent enzymes secreted are important with immune responses, cell proliferation, and recreating membranes of phospholipids. The protein encoded by this gene is a lipase which also breaks down fat and is distributed towards different organs. (Guo et al., 2018).

The majority of human genes are composed of a series of exons and introns. An exon is a section of DNA within a gene that contains the genetic code used to create protein-encoding parts of the messenger RNA (mRNA) molecule. These exons are retained after non-coding regions called introns are removed through RNA splicing (Aspden et al., 2023). Alterations in the DNA sequence within an exon can result in abnormalities in the protein sequence. These changes may contribute to various effects, potentially including the onset or progression of diseases, consequently promoting more cognitive and mobility-depriving diseases like Dystonia or Parkinsons, classified as PLAN (PLA2G6- Associated- Neurodegeneration). The 4 main types of PLAN are Infantile Neuroaxonal Dystrophy (INAD), Atypical Neuroaxonal dystrophy (ANAD), Adult-onset Dystonia-Parkinsonism (DP), and Autosomal Recessive Early-Onset Parkinsonism (AREP) (Guo et al., 2018). While INAD, which affects infants younger than 2, and ANAD are more youth-focused, DP and AREP affect people until around the age of 40. INAD patients show psychomotor decline and impaired functions such as hearing and swallowing. Those diagnosed with ANAD suffer from seizures, and psychomotor regression and often have autism. Both DP and AREP are triggered by Parkinson's which encourages a further decline (Guo et al., 2018). Traditional medication and methods prove ineffective towards these symptoms, as the disease is ultimately terminal. With advancing technology, more effective and

promising techniques are coming into play when trying to target mutations and hopeful treatments and preventative measures. Given that PLAN is not a complete mystery, where both the underlying issues and origin are apparent, solutions like gene therapy emerge as a possible, effective cure for these diseases.

	INAD	ANAD	DP	AREP
Age of onset	6 months-3 years	Early childhood - 18 years	20-40 years	under 40 years
Initial Symptoms	Psychomotor deterioration	Cerebellar ataxia	Parkinsonism	Parkinsonism
Progression	Rapid	Slower than INAD	Slow	Slow
Important Signs	Cerebellar ataxia, hypotonia, hyporeflexia, nystagmus, strabismus	Cerebellar ataxia, hypermyotonia	Hypermyotonia, extrapyramidal signs	Hypermyotonia in lower limbs
Brain MRI Signs	Brain iron accumulation in globus pallidus, substance Niagara, cerebellar atrophy	Brain iron accumulation in two sides of basal ganglia, cerebellar atrophy	A substantial reduction in data	A substantial reduction in data
Patient Lifetime	Short, before age 10	Slightly longer than INAD	Long	Long
Treatment	Symptomatic Treatment	Symptomatic Treatment	Dopaminergic Agents	Dopaminergic Agents

### ***An Intro to CRISPR***

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) are used to edit the genomes of organisms by targeting their DNA sequence. CRISPR/Cas9 (CRISPR-associated protein 9) originated from a naturally occurring system found in certain bacteria, which functions as part of their immune system. CRISPR, facilitated by the CAS proteins, is a part of the bacterial immune system, which uses fragments of the virus's DNA and keeps a "genetic memory" of past infections. When a bacteriophage infects the cell, it inserts its genetic material which gets replicated by the host cells' machinery (Lander 2016). Through the lytic cycle, once enough phages are accumulated in the cell, the cell eventually 'pops' and releases all the phages into the environment, which leads to the destruction of the cell. To combat this fate, many bacteria have developed their own CRISPR system to build immunity. The CRISPR structure is composed of spacers and repeats. The repeats are short, identical sequences of clustered base arrangements and play a role in the adaptation phase of the CRISPR response. Other Cas proteins

recognize these repeats and insert spacers between them, but they do not contribute to guiding the Cas9 protein to target and cut the DNA (Lander 2016). The spacers are transcribed into CRISPR RNA (crRNA), a fundamental component of the CRISPR-Cas system that originates from the CRISPR array in bacterial DNA. The crRNA is the RNA sequence complementary to the target DNA site and has a complementary base pair towards the target DNA sequence. The tracrRNA interacts with the Cas9 protein and stabilizes its binding to the sgRNA, helping to stabilize and orientate it within the complex. This interaction is important to ensure the proper assembly of the Cas9-sgRNA ribonucleoprotein (RNP) complex, which is necessary for precise DNA targeting and cleavage (Liao & Beisel, 2023). Together with tracrRNA, they guide the Cas9 enzyme to modify specific DNA sequences, allowing precise genome editing in scientific and medical applications across different organisms. Spacers are complementary to the specific segments of phage DNA (Liao & Beisel, 2023). This structure stores the phage DNA to build up a defense for when they try to attack again and is passed down to the following generations. When the bacterium encounters a viral or plasmid threat, the crRNA guides the Cas proteins to the corresponding sequences in the foreign DNA (Lander 2016). The Cas proteins, particularly Cas9 in the case of the popular CRISPR-Cas9 system, act as molecular scissors, cutting the invading DNA at precise locations dictated by the crRNA sequence. This targeted DNA cleavage disables the virus or plasmid, rendering it unable to replicate and infect the bacteria further.

The Cas9 protein is one of the most important components of this process. When Cas9 is paired with crRNA and tracrRNA, it is responsible for recognizing and attacking foreign genetic material (Lander 2016). This protein acts almost like scissors and will now use the crRNA to scan for any sequences that are complementary to crRNA, or in other words, match previously encountered phages or viruses. Once detected, Cas9 binds to the DNA and starts scanning it using the crRNA. When it finds a match, it positions its nuclease domains to make a double-strand break in the target DNA, initiating the destruction of the infected code. This process helps the bacteria acquire immunity against these diseases and incorporates this immunity into the bacteria's original genome. Think of Cas9 as a zipper that progressively unwinds the DNA while searching for a match, rather than simply binding to the target DNA directly (Liao & Beisel, 2023). Cas9 is now being implemented in today's medical world and used to treat mutated DNA which is responsible for neurodegenerative diseases.

As a gene editing tool, the CRISPR system is a two-part system that utilizes the Cas9 protein and the guide RNA (sgRNA) (Liao & Beisel, 2023). Like the crRNA, the sgRNA's role is to help guide the Cas9 to the correct target DNA sequence. The sgRNA combines two different RNA functions, CRISPR RNA (crRNA) and trans-activating RNA (tracrRNA), into a chimeric RNA complex (Liao & Beisel, 2023). This chimeric RNA is formed synthetically in the lab by combining the essential parts of the crRNA and tracrRNA into a single hairpin-shaped RNA molecule. This simplifies the delivery and use of the CRISPR system. Once the Cas9 protein recognizes a region of DNA complementary to the gRNA, it creates a double-stranded break (DSB) at the targeted site (Cui et al., 2018). At the DSB, the gene will be edited or mutated, depending on the outcome of the repair through either non-homologous end joining (NHEJ) or

homology-directed repair (HDR)(Xue & Greene, 2021). This targeted alteration can lead to specific changes in gene function, such as disabling a gene to study its role in cellular processes or disease, correcting mutations responsible for genetic disorders, or inserting new genetic material to introduce desired traits. The proceeding DSB repair process, whether through error-prone non-homologous end joining (NHEJ) or precise homology-directed repair (HDR), then determines the success and specificity of the genetic modification intended by the researcher(Xue & Greene, 2021). Both NHEJ and HDR are possible within a population of cells intended for editing. Researchers can influence the preference for one repair pathway over the other using various techniques, but these methods are not 100% effective. This variability represents a significant limitation, so it is important to note that the repair process may not always yield the desired outcome (Xue & Greene, 2021). The NHEJ process is first initiated when a DSB is detected in the DNA. NHEJ then fixes the DSB and rejoins the broken ends of the DNA, however, is limited in precision when trying to ligate broken DNA ends. NHEJ does not have any specific template to follow, and instead uses the proteins, Ku70, Ku80, and DNA-PKcs which bind to the ends and help align the DNA (Lieber 2010). This could result in random insertions or deletions (indels) of base pairs. The indels can further alter the sequence of the protein encoded by this gene, and in many cases lead to a loss of function that ablates the function of the protein (Lieber 2010). This underscores the importance of developing and utilizing genome editing techniques that prioritize accuracy and minimize off-target effects, such as optimizing CRISPR-Cas9 systems for improved specificity and reliability. Unlike NHEJ, which is error-prone and can lead to random indels at the repair site, HDR offers a more precise method of genome editing. HDR utilizes a homologous template DNA strand to guide the repair of double-stranded breaks (DSBs)(Xue & Greene, 2021). This process occurs when the cell is in the S or G2 phase of the cell cycle. HDR is thus more accurate, as it uses the template to precisely replace the damaged DNA sequence. This allows for the insertion of specific sequences, correction of mutations, or more targeted genetic modifications at the targeted genomic locus(Xue & Greene, 2021). The success and specificity of HDR-mediated genetic modifications depend on the availability and quality of the homologous template DNA, which can be modified and perfected in the lab as well as the efficiency of the HDR machinery within the cell To achieve a precise gene edit, researchers can use a custom-designed template DNA, engineered by researchers to carry specific genetic sequences needed for precise genome editing(Xue & Greene, 2021). This template is used during cell repair processes to insert new sequences, correct mutations, or make targeted genetic modifications at desired genomic locations. This approach allows for controlled and accurate genetic alterations, essential for both research and therapeutic applications.

Researchers raised the question if these methods could be applied to other diseases and gene-related issues, which began the long journey of gene editing and revolutionized the fields of genetics and biomedicine. This targeted method could not only correct the mutations and “edit the gene” but perhaps even eliminate the disease at its root. The paper will begin to identify both

the positive and adverse potential of gene editing for the PLA2G6 gene and different strategies to further its research.

## Body

### *CRISPR for PLA2G6*

CRISPR-Cas9 technology offers a promising therapeutic strategy for genetic disorders by precisely targeting and correcting mutations. By utilizing specially designed guide RNAs and the cell's endogenous repair pathways, CRISPR can accurately address point mutations or replace defective segments within the *PLA2G6* gene. This gene-editing method has the potential to restore the normal function of the iPLA2 $\beta$  enzyme, thereby addressing the root cause of neurodegenerative diseases like INAD and dystonia-parkinsonism (Guo et al., 2018). To implement CRISPR in *PLA2G6*, it is important to understand the nature of the mutations and how this further affects the body's normal function. As mentioned before, the iPLA2 $\beta$  enzyme hydrolyzes the phospholipids that maintain the cell membrane's structure and function. This protein contains many important sites which include eight N-terminal ankyrin repeats, a caspase-3 cleavage site, an ATP-binding domain, a serine lipase consensus sequence (GX SXG), a bipartite nuclear localization sequence, and a C-terminal calmodulin-binding domain (Guo et al., 2018). These sites are responsible for providing the cell with energy, controlling apoptosis, breaking down phospholipids, and calcium regulation.

One of the common mutations that occurs within *PLA2G6* is a missense mutation. Missense mutations, which result in single amino acid substitutions, can be targeted using CRISPR technology (Engel et al., 2010). By specifically addressing the mutated nucleotide sequence, CRISPR enables the introduction of a correct DNA template through HDR. This process replaces the errored sequence with the accurate one, thereby reinstating normal protein function. Frameshift mutations are mutations that introduce a premature stop codon, interfering with the normal stop codons and subsequently altering the reading sequence of amino acids (Engel et al., 2010). A faulty amino acid sequence results in a faulty or even dysfunctional protein. They can have detrimental effects and lead to genetic disorders, diseases, or loss of function of essential proteins depending on where and how they occur within the genome. For mutations causing frameshift that disrupt protein function, CRISPR offers the capability to either remove or correct the defective exonic regions (Engel et al., 2010). This correction ensures the production of a functional protein. However, the difference in the sites will make it hard to create one generalized therapy that targets a specific location. Additionally, it is important to consider that INAD and ANAD are observed at much younger ages, while DP and AREP typically affect adults, who may also face problems like Parkinson's (Guo et al., 2018). While a younger person will find it easier to adapt to the changes because their bodies are still growing, adults will be more of a challenge because they have a more developed and complex structure, and having something go wrong can make them lose the function they already have. The same could be said

with infants and they are much more vulnerable to complications. Therapy is being given so early on.

### ***Intracranial Viral Vector Delivery***

Intracranial viral vector delivery is a sophisticated technique in neuroscience where viral vectors are employed to introduce genetic material into specific regions of the brain (Naso et al., 2017). This method typically involves selecting appropriate viral vectors like adeno-associated viruses (AAV) or lentiviruses, which are modified to carry desired DNA or RNA constructs. AAVs are small, non-enveloped viruses with a single-stranded DNA genome that can be engineered to carry therapeutic genes. These vectors are especially notable for their use in stable genetic manipulation (Naso et al., 2017). AAVs are non-pathogenic and do not replicate efficiently in human cells, making them safe for therapeutic use. Their ability to provide long-term gene expression in non-dividing cells makes them viable for treating genetic diseases, offering a promising avenue for precise and sustained genetic correction.

The typical AAV vector consists of a small viral genome packaged within a protein capsid derived from wild-type AAVs. The viral vector is then injected using microsyringes, ensuring accurate placement and controlled delivery volumes to minimize tissue damage (Naso et al., 2017). Once inside the cell, the AAV vector enters the nucleus, where its DNA cargo is released and can integrate into the host cell's genome or remain episomal (staying in the nucleus without integrating into the cell's chromosomal DNA), depending on the AAV serotype and target cell type. AAV2, for example, is adept at targeting neurons, which is crucial for addressing neurological disorders, such as those associated with the PLA2G6 gene (Naso et al., 2017).

AAV2 has a natural affinity for neuronal cells which allows it to cross the blood-brain barrier and bind to the receptor in these cells (Shi et al., 2004). AAV2 directly delivers into the nucleus host cells and the DNA remains episomal, which allows it to facilitate long-term gene expression without integrating into the host's genome, further improving the safety of the AAV as there is no chance of random integration of the vector into the genome. For treating disorders related to the PLA2G6 gene, such as PLA2G6-associated neurodegeneration (PLAN), it is easy to mistakenly assume that AAV2s or lentiviruses could be engineered to carry corrected versions of the PLA2G6 gene and still be effective (Wang et al., 2019). Despite their potential, AAVs have limitations due to their restricted capacity for carrying genetic material, which is approximately 5,000 base pairs (Wang et al., 2019). This constraint poses challenges when working with larger and more complex genes or therapeutic constructs, such as those used in CRISPR-Cas9 systems. For instance, the Cas9 enzyme alone occupies 4,000 base pairs, and additional space is required to guide RNAs and promoter sequences before the therapeutic gene can even be delivered (Trapani, 2019). To overcome this limitation, a system involving the use of dual AAV vectors can be explored. This approach splits the therapeutic gene or components of CRISPR into two different AAV vectors which contain portions of the total genetic material. Once injected into the target cell, the two separate vectors combine to form a functional unit

(Trapani 2019). With conditions like PLAN, a dual AAV vector proves to be an alternative where a single AAV vector is not feasible due to its size.

It is also important to consider that AAVs prefer transient expressions of CRISPR components to treat conditions including PLAN. Transient expression of CRISPR refers to the temporary introduction and activity of CRISPR components, such as the Cas9 enzyme and guide RNAs, within a cell (Wang et al.,2019). Unlike stable expression, where the CRISPR system is permanently integrated into the cell's genome, transient expression ensures that the CRISPR components are only present for a limited time. After the AAV delivers the CRISPR system and the genetic modification is completed, the CRISPR components are rapidly degraded or removed from the cell, leaving behind only the therapeutic benefits and reducing the potential for unintended off-target effects (Wang et al.,2019). This approach is used to achieve genetic modifications without the risk of long-term CRISPR activity, which can reduce the likelihood of off-target effects and potential disruptions to normal cellular functions.

### ***Consequences of Gene Therapy/CRISPR in the Brain***

CRISPR technology, while transformative for gene editing, faces challenges such as off-target effects, where the guide RNA may inadvertently recognize and bind to unintended DNA sequences, leading to unintended edits. The specificity of the guide RNA is critical, yet it is possible for it to cross-react with similar sequences elsewhere in the genome. To address these issues, molecular evolution has produced higher-fidelity versions of Cas9 that offer improved accuracy, reducing off-target activity. Additionally, advances in CRISPR systems include the use of Cas proteins like Cas11 and Cas13, which have distinct enzymatic properties compared to traditional Cas9, and various algorithms have been developed to screen for and minimize off-target effects. Researchers are exploring these novel Cas proteins and algorithms to refine the precision of gene editing. For applications in neurons, where HDR is less effective due to the cells' capacity for self-repair and limited HDR efficiency, base editors present a promising alternative. Base editors do not rely on Cas9's "scissors" but instead enable precise nucleotide changes without introducing double-strand breaks, offering a more reliable method for achieving targeted genetic modifications in neuronal cells.

The immunogenicity of the Cas9 protein poses a potential concern for CRISPR-based therapies targeting genes such as PLA2G6. Since it is derived from a bacteria, the Cas9 protein might trigger an immune response in the body, potentially leading to inflammation, swelling, and fluid buildup in the brain. This could undermine the effectiveness of the treatment by introducing unwanted side effects or complications. Consequently, while the immune response remains a critical consideration, advances in minimizing Cas9's immunogenicity through protein engineering or other strategies could enhance the safety and efficacy of CRISPR therapies for PLA2G6-related conditions. One approach is to explore Cas9 proteins derived from commensal bacteria, such as those naturally present in the gut. These proteins could be better tolerated by the immune system due to their familiar presence in the body, potentially reducing the risk of adverse immune reactions compared to entirely foreign Cas9 proteins.

The Blood-Brain Barrier (BBB) stands between the brain and the bloodstream, creating many challenges (Stanimirovic et al., 2018). The BBB is mainly composed of endothelial cells, which form the lining of blood vessels in the brain; Astrocyte End-Foots, star-shaped glial cells that provide support and maintain BBB integrity; Pericytes which regulate blood flow and permeability; and finally the basal membrane, a thin, extracellular matrix layer which separates the endothelial cells from the surrounding tissue. These cells protect the brain from toxins and pathogens while regulating nutrient and waste exchange. They furthermore restrict the movement of large hydrophilic therapeutics and contain polarized transporters that prevent unwanted lipophilic synthetic molecules from entering the brain (Stanimirovic et al., 2018). Biologics are therapeutic products derived from biological sources, including vaccines, recombinant proteins, gene therapy, and living cells. These large, complex molecules were once thought impractical for BBB penetration. Techniques like mannitol-induced osmotic BBB disruption and focused ultrasound have also shown limited clinical benefits. The BBB's restrictive nature, the passage of substances from the bloodstream into the brain is, thus complicating the delivery of CRISPR/Cas9 therapies intended for neurological disorders. Intravenously delivered CRISPR/Cas9 components can inadvertently end up in the liver, where they are taken up by hepatocytes instead of reaching their intended target in the brain. This misdelivery is a significant challenge in gene therapy, leaving researchers to look for technologies to overcome this (Teleanu et al., 2022).

One solution is through Lipid Nanoparticles (LNPs). Lipid nanoparticles are formed through the self-assembly of lipid components, including ionizable lipids, structural lipids, helper lipids, and cholesterol, into a spherical structure that encapsulates nucleic acids (Khare et al., 2023). Ionizable lipids help capture and release these nucleic acids in the cytoplasm. Structural lipids, such as phospholipids, form the lipid bilayer that maintains the nanoparticle's structural integrity, while helper lipids support the stability and release properties of the LNPs (Khare et al., 2023). Cholesterol is often included to enhance the fluidity and stability of the lipid bilayer. After formation, the LNPs are purified to remove unencapsulated materials and solvents, using methods like ultracentrifugation or filtration. These steps ensure that the LNPs are well-suited for delivering therapeutic agents effectively, minimizing unintended side effects, and improving targeting precision.

LNPs address the BBB challenge through several strategies. Firstly, they can be engineered with surface modifications that include specific ligands or antibodies, which enhance their ability to cross the BBB and target brain cells (Khare et al., 2023). These modifications help direct the nanoparticles away from the liver and toward their intended destinations. Additionally, adjusting the size and lipid composition of LNPs can influence their biodistribution, potentially reducing liver accumulation and improving targeting efficiency. Controlled release mechanisms also play a critical role. LNPs can be designed to release their content, including the CRISPR/Cas9 components, in a controlled manner once they reach the target cells (Khare et al., 2023). This controlled release minimizes off-target effects and enhances therapeutic outcomes. Exploring alternative delivery methods, such as localized injection directly into brain tissues,



also holds promise for bypassing systemic circulation and reducing liver uptake. This approach can be contextualized for targeting specific neuronal conditions linked to PLA2G6, allowing for more precise and effective gene editing. The goal here is to deliver a DNA cargo that encodes the Cas9 enzyme, gRNA, and repair template, ensuring targeted and efficient genetic modification within the affected brain regions.

### ***Uterus Gene Therapy***

Uterus gene therapy, when applied to treating conditions in developing fetuses, offers a promising approach to addressing genetic disorders before birth. This therapeutic strategy involves delivering therapeutic genes directly into the uterine environment, aiming to correct genetic mutations or modify gene expression in the fetus (Bose et al., 2021). For conditions that manifest early in development, such as certain neurodegenerative diseases, this approach has the potential to halt or prevent the progression of disease before symptoms even appear. In practice, uterus gene therapy for fetuses might involve several key methods. One approach is to use viral vectors or nanoparticles to deliver therapeutic genes to the fetal tissues through the amniotic fluid or directly into the uterine environment (Bose et al., 2021). By correcting genetic mutations or providing essential proteins, gene therapy could prevent the development of neurodegenerative diseases that would otherwise progress after birth. For neurodegenerative diseases like PLA2G6-related disorders, which are characterized by progressive damage to neurons, early intervention is crucial. However, the utero gene therapy, though tempting, cannot be implemented due to the size of the PLA2G6 gene.

Screening eggs, or preimplantation genetic testing (PGT), is a crucial technique used to identify genetic abnormalities in embryos before they are implanted during in vitro fertilization (IVF). This process begins with ovarian stimulation to produce multiple eggs, which are then retrieved and fertilized in the laboratory (Ryche & Berckmoes, 2020). The resulting embryos are cultured until they reach the blastocyst stage, at which point a few cells are biopsied for genetic testing. Advanced techniques such as next-generation sequencing or array comparative genomic hybridization are used to detect specific genetic mutations, including those associated with neurodegenerative diseases like PLA2G6-related disorders.

While egg screening and uterine gene therapy offer considerable advantages, such as preventing genetic disorders and supporting informed reproductive choices, they also come with challenges and ethical dilemmas (Ryche & Berckmoes, 2020). Despite the sophistication of modern genetic testing, there is still a risk of false positives or negatives that could impact critical decisions. Ethical issues include the potential for discarding embryos with genetic abnormalities and the broader implications of selecting embryos based on genetic traits, which raises questions about the limits of reproductive technology. Additionally, these advancements could exhibit socioeconomic disparities, as wealthier individuals may have greater access to and benefit from genetic modifications, potentially increasing inequality in reproductive health outcomes.

### ***Brain Organoids For Testing***

Brain organoids, which are three-dimensional, self-organizing clusters of brain-like cells derived from stem cells, offer a revolutionary approach to studying neurodegenerative diseases, including those caused by *PLA2G6* gene mutations (Wang et al.,2022). These organoids replicate many aspects of brain development and function, providing a more accurate model of human brain tissue than traditional two-dimensional cell cultures or animal models.

Brain organoids are created through a sophisticated process that starts with pluripotent stem cells, which can differentiate into any cell type in the body. These stem cells, either derived from embryonic sources or induced pluripotent stem cells (iPSCs) reprogrammed from adult cells, are cultured in a three-dimensional matrix that mimics the brain's extracellular environment. (Wang et al., 2022). This matrix supports the cells and encourages their self-organization into neural progenitor cells. These progenitors are then exposed to specific growth factors and signaling molecules that guide their differentiation into various brain cell types, including neurons, astrocytes, and oligodendrocytes. As the cells continue to grow and differentiate, they spontaneously form complex, multi-layered structures that resemble different regions of the brain. This process of self-organization results in brain organoids with features akin to those of the human brain (Wang et al.,2022). Over several weeks to months, these organoids mature, exhibiting characteristics such as electrical activity and neuronal connectivity. Researchers then characterize the organoids using techniques like microscopy, electrophysiology, and gene expression analysis to study their development and function. Through this intricate process, brain organoids provide a valuable model for investigating brain development, disease mechanisms, and potential therapeutic approaches. Once implanted, the brain organoid is observed for integration and functionality through techniques such as imaging and histological analysis, allowing researchers to assess how the organoid affects brain function and pathology (Wang et al.,2022). This approach enables detailed investigation into disease mechanisms and the testing of potential treatments in a system that closely mimics human brain tissue.

By incorporating *PLA2G6* gene mutations into brain organoids, researchers can observe the direct effects of these mutations on neuronal development, functionality, and pathology in a controlled environment. Using brain organoids to study *PLA2G6* mutations allows for a detailed investigation of disease mechanisms at a cellular level. For example, researchers can assess how these mutations impact neuronal differentiation, synaptic connectivity, and overall brain architecture. Moreover, brain organoids can be employed to test potential therapeutic strategies, including CRISPR-based gene editing, to correct *PLA2G6* mutations and evaluate their effectiveness and safety. This application is particularly valuable because organoids offer a platform for high-throughput screening of drugs or gene therapies, providing insights into how such interventions might mitigate disease symptoms or modify disease progression.

### ***Mouse Models***

CRISPR-Cas9 technology can also be effectively utilized to create and manipulate mouse models. This method allows scientists to introduce specific mutations into the genetic material of

mice, enabling the study of these mutations' effects and the exploration of potential treatments. For instance, in the case of the PLA2G6 gene, researchers can introduce a knock-in mutation by targeting exon 7, a protein-coding region of the gene that is often mutated in human patients. The process begins with the creation of a knock-in target vector, which includes the DNA sequence to be inserted. Through a method known as homologous recombination, this vector aligned precisely with the corresponding sequence in the mouse genome, allowing for the accurate insertion of the mutation. This technique provides valuable insights into the role of the PLA2G6 gene and the potential therapeutic strategies for conditions associated with its mutations via study of the resulting mouse.

Scientists initiated a study where they knocked in mutated PLA2G6 genes into mice. While studying the effects of the mutation, scientists observed a reduction in dopamine neurons as the mice aged. Dopamine is crucial for various physiological functions, including prolactin production in the hypothalamus, which influences immune and metabolic functions. It also plays a vital role in cognition, attention, motor control, and overall mood, with its deficiency being linked to diseases such as Parkinson's. Additionally, mitochondria, the organelles responsible for providing cellular energy, storing calcium, and regulating apoptosis, were found to be impaired (Chiu 2019). Increased apoptosis levels in mitochondria were noted with the presence of more misfolded proteins. This was assessed using transmission electron microscopy, which employs light beams to create stains, revealing impaired mitochondrial activity and decreased ATP production associated with the mutated gene. Further analysis of the mice showed the activation of apoptosis pathways and downregulation in dopaminergic genes, resulting in the reduced creation of dopamine neurons. Researchers also examined altered mRNA levels of various proteins that either promote apoptosis or decrease the production of essential proteins for dopamine neuron survival, such as Mapk1 (Chiu 2019).

These mouse models were also used to simulate symptoms and explore potential treatment avenues for the future. For example, the researchers measured the motor abilities of the mice by timing their descent from a pole, observing spontaneous movements, and assessing their balance on a rotating apparatus (Chiu 2019). The main findings indicated that as the mice aged, there was a correlated decrease in velocity and distance, increased incoordination, and reduced activity. These symptoms align with those observed in diseases such as Parkinson's, which result from neurodegeneration due to mutations in the PLA2G6 gene. Additionally, the scientists tested various drugs to assess their potential therapeutic effects. One such drug was methyl L-DOPA, which was introduced to the mice to evaluate its impact. Compared to mice injected with saline, those treated with methyl L-DOPA showed significant improvements (Chiu 2019). This finding suggests that methyl L-DOPA could be a promising treatment option for mitigating symptoms associated with PLA2G6-related neurodegeneration.

More recently, advancements in gene therapy, particularly using AAV vectors, have shown significant promise for treating PLAN, with a strong emphasis on the use of mouse models. Researchers have developed an AAV-based gene therapy construct called AAV-EF1a-PLA2G6, which utilizes the elongation factor EF-1 alpha (EF1a) promoter to broadly

express PLA2G6 across various cell types (Lin et al., 2023). When this construct was delivered to patient-derived neural progenitor cells (NPCs), it expressed PLA2G6 at low levels (10-20% of endogenous levels), which was sufficient to partially alleviate defects in Vps35 levels, lysosomal expansion, and mitochondrial abnormalities. Crucially, in Pla2g6 KO/G373R mice—a model for INAD—delivery of AAV-EF1a-PLA2G6 delayed the onset of motor defects, helped maintain body weight, and prolonged lifespan (Lin et al., 2023). This approach underscores the importance of mouse models in evaluating the efficacy of AAV-mediated gene therapy for PLAN and suggests a promising avenue for further optimization and development of treatments for PLA2G6-associated neurodegeneration.

### ***Discussion***

PLAN currently lacks prescribed treatments or drugs, but exploring avenues in gene therapy may ultimately provide a solution. Advancements in CRISPR-Cas9 technology, intracranial viral vector delivery, and innovative approaches such as in-utero gene therapy and brain organoids present treatment for PLAN diseases. CRISPR-Cas9 holds the potential to precisely correct mutations in the PLA2G6 gene, potentially restoring normal iPLA2 $\beta$  enzyme function and addressing conditions like INAD and dystonia-parkinsonism. Advances in AAV vector techniques are enhancing targeted gene delivery to the brain, and overcoming the challenges of the blood-brain barrier. While challenges like the size constraints of AAV vectors, potential off-target effects, and the immunogenicity of CRISPR components need ongoing refinement, these implications could be resolved with continued research. Uterus gene therapy offers the possibility of preventing disease onset before birth, although it requires careful delivery and ethical consideration. Brain organoids provide valuable models for studying PLA2G6 mutations and testing new therapies, and CRISPR-enhanced mouse models are shedding light on disease progression and the efficacy of treatments like methyl L-DOPA.

Gene therapy using AAV9-PLA2G6 has shown promising results in a mouse model of PLAN, effectively preventing neurodegeneration, halting disease progression, and even reversing existing neurological damage. Treated mice exhibited significant improvements in motor function and lifespan, indicating the potential for this therapy to address both neurological and systemic aspects of the disease (Lin et al., 2023). While current treatments primarily focus on managing symptoms, this research opens up new possibilities for targeted therapies, including the development of gene therapy for human trials and cardiolipin-focused interventions. Future therapies could combine gene therapy with existing symptomatic treatments for a more comprehensive approach to patient care (Lin et al., 2023).

Researchers should also focus on developing advanced genetic counseling methods, including tests to evaluate the risk of PLA2G6-associated neurodegeneration in individuals and their offspring. Enhancing carrier testing, prenatal testing for at-risk pregnancies, and preimplantation genetic testing can be invaluable, particularly if pathogenic variants are known in the family. Early identification of genetic risk can profoundly influence family planning decisions, helping prospective parents gauge the likelihood of their child having the condition

and make informed choices. This proactive approach can also equip families with the tools for early intervention and support, potentially alleviating the emotional and financial strain of managing a neurodegenerative disease with significant impacts.

In studying PLA2G6-associated neurodegeneration (PLAN), prioritizing the development of patient-derived brain organoids is crucial. These organoids should be analyzed for dopamine pathway changes to assess the impact of potential gene therapies. These organoids could consist of patient-derived cells and incorporating multiple cell types into these organoids allows for a comprehensive evaluation of potential therapies. Conducting preclinical trials with these organoids allows researchers to assess the effectiveness of treatments before moving to animal models. Following successful organoid validation, preclinical testing in mouse models is essential to confirm safety and efficacy. Additionally, genetic and molecular profiling, long-term efficacy studies, and comprehensive safety testing are necessary to ensure a thorough evaluation. These steps will advance effective treatments from experimental phases to clinical trials, addressing the needs of patients with advanced PLAN. Advancing treatments through these rigorous experimental phases will offer hope and improve the quality of life for those affected by these debilitating conditions.

## Works Cited

- Xue, C., & Greene, E. C. (2021, July). DNA repair pathway choices in CRISPR-Cas9-mediated genome editing. *Trends in genetics : TIG*.  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8187289/>
- Aspden, Julie L., et al. 'Not All Exons Are Protein Coding: Addressing a Common Misconception'. *Cell Genomics*, vol. 3, no. 4, Elsevier BV, Apr. 2023, p. 100296,  
<https://doi.org/10.1016/j.xgen.2023.100296>.
- Chiu, Ching-Chi, et al. 'PARK14 (D331Y) PLA2G6 Causes Early-Onset Degeneration of Substantia Nigra Dopaminergic Neurons by Inducing Mitochondrial Dysfunction, ER Stress, Mitophagy Impairment and Transcriptional Dysregulation in a Knockin Mouse Model'. *Molecular Neurobiology*, vol. 56, no. 6, Springer Science and Business Media LLC, June 2019, pp. 3835– 3853, <https://doi.org/10.1007/s12035-018-1118-5>.
- Cui, Yingbo, et al. 'Review of CRISPR/Cas9 sgRNA Design Tools'. *Interdisciplinary Sciences, Computational Life Sciences*, vol. 10, no. 2, Springer Nature, June 2018, pp. 455–465,  
<https://doi.org/10.1007/s12539-018-0298-z>.
- Davenport, Franca, et al. 'Neurodegenerative Disease of the Brain: A Survey of Interdisciplinary Approaches'. *Journal of the Royal Society, Interface*, vol. 20, no. 198, The Royal Society, Jan. 2023, p. 20220406, <https://doi.org/10.1098/rsif.2022.0406>.
- De Rycke, Martine, and Veerle Berckmoes. 'Preimplantation Genetic Testing for Monogenic Disorders'. *Genes*, vol. 11, no. 8, MDPI AG, July 2020, p. 871,  
<https://doi.org/10.3390/genes11080871>.
- Engel, Laura A., et al. 'Catalytic Function of PLA2G6 Is Impaired by Mutations Associated with Infantile Neuroaxonal Dystrophy but Not Dystonia-Parkinsonism'. *PloS One*, vol. 5, no. 9, Public Library of Science (PLoS), Sept. 2010, p. e12897,  
<https://doi.org/10.1371/journal.pone.0012897>.
- Gregory, Allison, et al. 'PLA2G6-Associated Neurodegeneration'. *GeneReviews®*, University of Washington, Seattle, 1993,  
[https://www.ncbi.nlm.nih.gov/books/NBK1675/#:~:text=Routine.hme.‘Uncovering the Culprits of Neurodegenerative Disorders’. News Center, 11 Apr. 2024,  
<https://news.feinberg.northwestern.edu/2024/04/11/uncovering-the-culprits-ofneurodegenerative-disorders/>.](https://www.ncbi.nlm.nih.gov/books/NBK1675/#:~:text=Routine.hme.‘Uncovering%20the%20Culprits%20of%20Neurodegenerative%20Disorders’%20News%20Center%2C%2011%20Apr.%202024,https://news.feinberg.northwestern.edu/2024/04/11/uncovering-the-culprits-ofneurodegenerative-disorders/)
- Guo, Yu-Pei, et al. 'PLA2G6-Associated Neurodegeneration (PLAN): Review of Clinical Phenotypes and Genotypes'. *Frontiers in Neurology*, vol. 9, Frontiers Media SA, Dec. 2018, p. 1100, <https://doi.org/10.3389/fneur.2018.01100>.
- Hodel, Amanda S. 'Rapid Infant Prefrontal Cortex Development and Sensitivity to Early Environmental Experience'. *Developmental Review: DR*, vol. 48, Elsevier BV, June 2018, pp. 113–144, <https://doi.org/10.1016/j.dr.2018.02.003>.
- Lander, Eric S. 'The Heroes of CRISPR'. *Cell*, vol. 164, no. 1–2, Elsevier BV, Jan. 2016, pp. 18–28, <https://doi.org/10.1016/j.cell.2015.12.041>.

- Liao, Chunyu, and Chase L. Beisel. 'The tracrRNA in CRISPR Biology and Technologies'. *Annual Review of Genetics*, vol. 55, no. 1, Annual Reviews, Nov. 2021, pp. 161–181, <https://doi.org/10.1146/annurev-genet-071719-022559>.
- Lieber, Michael R. 'The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway'. *Annual Review of Biochemistry*, vol. 79, no. 1, Annual Reviews, 2010, pp. 181–211, <https://doi.org/10.1146/annurev.biochem.052308.093131>.
- Lin, Guang, et al. 'Exploring Therapeutic Strategies for Infantile Neuronal Axonal Dystrophy (INAD/PARK14)'. *eLife*, vol. 12, eLife Sciences Publications, Ltd, Jan. 2023, p. e82555, <https://doi.org/10.7554/eLife.82555>.
- Naso, Michael F., et al. 'Adeno-Associated Virus (AAV) as a Vector for Gene Therapy'. *BioDrugs: Clinical Immunotherapeutics, Biopharmaceuticals and Gene Therapy*, vol. 31, no. 4, Aug. 2017, pp. 317–334, <https://doi.org/10.1007/s40259-017-0234-5>.
- National Center for Biotechnology Information. <https://www.ncbi.nlm.nih.gov/>. Accessed 28 Sept. 2024.
- Nemudryi, A. A., et al. 'TALEN and CRISPR/Cas Genome Editing Systems: Tools of Discovery'. *Acta Naturae*, vol. 6, no. 3, Acta Naturae Ltd, July 2014, pp. 19–40, <https://doi.org/10.32607/20758251-2014-6-3-19-40>.
- Stanimirovic, Danica B., et al. 'Emerging Technologies for Delivery of Biotherapeutics and Gene Therapy across the Blood-Brain Barrier'. *BioDrugs: Clinical Immunotherapeutics, Biopharmaceuticals and Gene Therapy*, vol. 32, no. 6, Springer Science and Business Media LLC, Dec. 2018, pp. 547–559, <https://doi.org/10.1007/s40259-018-0309-y>.
- Teleanu, Raluca Ioana, et al. 'Current Strategies to Enhance Delivery of Drugs across the Blood-Brain Barrier'. *Pharmaceutics*, vol. 14, no. 5, MDPI AG, May 2022, p. 987, <https://doi.org/10.3390/pharmaceutics14050987>.
- Trapani, Ivana. 'Adeno-Associated Viral Vectors as a Tool for Large Gene Delivery to the Retina'. *Genes*, vol. 10, no. 4, MDPI AG, Apr. 2019, p. 287, <https://doi.org/10.3390/genes10040287>.
- Wang, Dan, et al. 'Adeno-Associated Virus Vector as a Platform for Gene Therapy Delivery'. *Nature Reviews. Drug Discovery*, vol. 18, no. 5, Springer Science and Business Media LLC, May 2019, pp. 358–378, <https://doi.org/10.1038/s41573-019-0012-9>.
- Wang, You-Wei, et al. 'Genetic and Epigenetic Regulation of Brain Organoids'. *Frontiers in Cell and Developmental Biology*, vol. 10, Frontiers Media SA, July 2022, p. 948818, <https://doi.org/10.3389/fcell.2022.948818>.
- Xue, Chaoyou, and Eric C. Greene. 'DNA Repair Pathway Choices in CRISPR-Cas9-Mediated Genome Editing'. *Trends in Genetics: TIG*, vol. 37, no. 7, Elsevier BV, July 2021, pp. 639–656, <https://doi.org/10.1016/j.tig.2021.02.008>.

# **Behind the Glamour: Examining Gender Discrimination within K-pop**

## **By Riley Kim**

### **Abstract**

This paper investigates gender discrimination against female idols in the K-pop industry. By comparing male and female idols in areas such as performances, concepts, beauty standards, and marketing, this study seeks to highlight the inequalities between female and male idols. Using a feminist theoretical framework, it examines the career trajectories of prominent K-pop idols like STRAY KIDS and TWICE, while also exploring the historical roots of gender discrimination in South Korea. The findings reveal a systemic gender bias within South Korean society, which often confines women and female idols to restrictive roles, limiting their public expression due to fear of backlash. More specifically, this paper explores the disparities in the longevity, artistic freedom, and public perception of female idols compared to their male counterparts. This paper contributes to the growing discourse on gender inequality in popular culture, advocating for industry reforms and increased awareness of the challenges faced by women in K-pop.

### **1. Introduction**

K-pop, or Korean popular music, has rapidly evolved into one of the most influential global music genres, transcending linguistic and cultural boundaries to captivate audiences worldwide (Arnaud, 2021). This genre, renowned for its innovative concepts, visually captivating music videos, and dynamic performances, has led to the emergence of numerous idol groups, both female and male, each carefully curated by entertainment companies. While female idols have gained immense popularity for their charming personas and meticulously crafted images, they are simultaneously subjected to significant scrutiny and discrimination, rooted in the rigid and traditional gender roles ingrained in South Korean society. Although K-pop offers young women the opportunity to achieve their dreams of performing on grand stages, it may also be argued that it places them within an industry that often prioritizes beauty over talent and silences female idols' voices and needs (Lee, 2024).

This paper seeks to explore how the cultural and societal perceptions of Korean women shape the experiences of female K-pop idols. Specifically, the analysis will focus on how these societal views influence the audience demographics, beauty standards, marketing strategies, and lyrical content for female idols. Therefore, the central research question guiding this paper is: How do cultural and societal views of Korean women impact the perception and treatment of female K-pop groups?

In response, this paper is organized into four main sections. The first provides a foundational understanding of feminism and its relevance to gender equality and women's roles in South Korea. The second section offers an overview of the K-pop industry, its globalization, and the specific idol groups that will be the primary focus of this study. The final section delves into public perception and fan culture, examining how these factors shape the treatment and



success of female K-pop idols, shedding light on the broader implications for female artists in the industry.

This comprehensive approach aims to offer a nuanced understanding of the cultural and societal dynamics that influence the treatment of female K-pop groups within a highly competitive and demanding industry.

## **2. Feminism: A Theoretical Context**

Feminism, at its core, is the advocacy for gender equality and the eradication of gender-based inequalities. These inequalities are deeply ingrained in societal structures, often defining what is deemed appropriate for men and women. Gender, understood as a social status or personal identity, has historically dictated the expected behaviors of men and women within societal, familial, and occupational contexts (Lorber, 1997). This often results in harmful and discriminatory stereotypes that confine individuals to roles based on their gender.

The "first wave" of feminism began being documented in the late 19th and early 20th centuries. The primary aim of the feminist movement during this period was to expand women's roles in society, advocating for greater participation in politics and other civic activities. This era significantly influenced other movements, such as the abolitionist movement, and laid the foundation for subsequent feminist perspectives (Rampton, 2015). The "second wave" of feminism, which arose in the 1960s and 1970s, addressed issues across legal, social, and economic spheres. Activists of this wave were more willing to engage in direct action, including civil disobedience and public demonstrations, to challenge existing laws and societal norms, often in opposition to older generations (Thompson, 2002). However, many feminists have taken a more contemporary perspective, as they have realized that discrimination and inequality are not caused only by harmful gender stereotypes, but by women's experiences, which are also shaped by their race, sexuality, class, and more (Lorber, 1997).

### **2.1. Feminism and the Entertainment Industry**

Despite the various feminist movements that remain active and potent today, women's discrimination remains apparent in the music industry. Historically, societies have often placed women in domestic roles. At the same time, men were associated with professional careers, ultimately leading to their domination in these spaces. As men have held the majority of leadership positions due to this notion, female voices remain marginalized and underrepresented. Many individuals believe that women do not belong in the music industry, including producers, lyricists, and artists, as historically women have been assigned and ingrained in domestic roles (Connell, 2005). However, if female artists choose to pursue this profession, they face strict barriers, confining their creative visions and aspirations that reflect deep-seated societal biases. Female artists often encounter systemic barriers, including gender pay gaps, underrepresentation in leadership roles, and pervasive sexism (McRobbie, 2004).

For example, in 2021, Selena Gomez was ridiculed online for being "overweight" after bikini photos were released by paparazzi. Not only have many other women in the music

industry been harassed for their bodies, but have had their musical talents overlooked and whittled down to their looks (Walker, 2022). Additionally, Taylor Swift, a record-breaking pop star, has detailed her experiences with sexism in the music industry, specifically her experiences with negative critiques of her success versus other male artists as follows: (Cubit, 2019)

"A man does something, it's strategic; a woman does the same thing, it's calculated. A man is allowed to react; a woman can only overreact" (Swift, 2019).

These stereotypes continue to harm women's opportunities to flourish in a cut-throat industry.

## **2.2. Women in South Korea: A Context**

South Korea throughout history has undergone various industrial, economic, and technological transformations, becoming a leader in high-tech industries and economies. However, this being said, South Korea remains "stuck in history" regarding women's rights, at times being justified by tradition. Inequalities in opportunity and discrimination against women have been sustained due to deeply embedded Confucian customs, being resistant to change (Jung, 2023) In traditional Korean society, women have been disadvantaged, being expected to be the homemakers or wives. Korean society often makes it difficult for women to be working mothers, experiencing one of the largest gender pay gaps in the world (Lee).

Although women now have many more opportunities in education, work, and public life, because of the expectation of taking care of strictly household matters, has forced women into a box of obedience and subordination traditionally (Korean Overseas Information Service, 2019). South Korea's structure of a royal monarch and a strong extended family was intended to separate the roles of women and men. This separation is strongly reminiscent of Confucianism-society structures (Palley, 1990).

"One of the tenets of Confucianism requires three obediences of women: to the father when young; to the husband when married; and to the son in old age" (Palley, 1990).

Traditional structures in South Korean society frequently impose constraints on women, discouraging them from transcending established barriers. Concurrently, contemporary South Korean society imposes additional pressures concerning women's behavior and appearance. Specifically, prevailing beauty ideals, behavioral expectations, and speech norms further restrict women from expressing their authentic selves. These societal standards often contribute to a sense of inadequacy and diminished value among women who do not conform to these expectations (Human Rights Research Center, 2024).

## **3. Setting the Scene: A Background to the K-pop Industry**

### **Introduction**

In recent years, Korean pop music has taken the world by storm, quickly topping music charts with every "comeback" (album release). Numerous K-pop groups have achieved top chart hits on global music charts, leading Billboard to launch a global chart entitled "K-Pop Hot 100" (Messerlin and Shin, 2017). In the United States, K-pop has become a popular genre, becoming

bigger every year, as the US generated 6.6 billion K-pop streams and 9.2 billion in 2023 (Kroll, 2024). Quite notably the first "boom" of popularity in this genre was in 2012 when Korean rapper Psy's song "Gangnam Style" found roaring popularity in the United States. This song established K-pop's popularity and success overseas and has remained popular ever since as boy and girl groups called "idols" have gained popularity as well overseas (Messerlin and Shin, 2017).

### **3.1. What is K-pop?**

K-pop has become a "melting pot" of styles, making it a unique genre. With inspirations from pop, experimental, rock, hip-hop, R&B, and electronic styles, many groups adapt these styles to fit the "concept", or "theme" of their idol image (Dance Vision, 2024). Female idols encompass concepts that are seen as cute, grunge or cool, and sexy, represented in their lyrics, fashion, and dance moves. Male idols encompass a wider variety of concepts or images, being able to use concepts that often change with time while female idol's concepts stick for their careers (Primer, 2023). Though a worldwide favorite now, K-Pop had its humble beginnings in the 1950s, as the girl group the "Kim Sisters" became icons in South Korea. However, in the 1990s K-Pop got its unique "sound" that we know today, as "Seo Taiji and Boys" merged Korean music with more Western-style music. As time has progressed, K-Pop "netizens have considered there to be generations of K-Pop, as trends change. As idols have become the pinnacle of South Korean media, becoming the face of many brands, shows, fashion, and more, entertainment companies are always looking for the new "face" of K-Pop (Arnaud, 2021).

As idols have become the primary leaders of entertainment in South Korea, entertainment companies take in "trainees" for long periods. During this time, these trainees learn dancing, singing, acting, and media skills to prepare them for their possible, but not guaranteed, debut.

The idea of a 'training system' was begun by one of the biggest entertainment companies currently, SM Entertainment. Established by Lee Sooman, SM Entertainment introduced this distinct training system that is followed by various other companies, such as the leading ones like JYP Entertainment, SM Entertainment, and BigHit Entertainment. "Trainees" are often scouted on the street, even abroad, and many often audition in person and online in hopes of joining a large company (Pratamasari, 2017). Trainees undergo years of rigorous preparation, molded by their companies to fit the specific image required for upcoming projects. Their previous lives are left behind as they are shaped into the ideal idols. However, a successful debut is not guaranteed; the public's reception ultimately determines an idol's success. Trainees typically begin at ages 15-16, but in recent years, even children as young as 8-9 have started training to debut as teenagers. In a society where beauty is highly valued, visual appeal often takes precedence over musical talent, as some critics have noted. Many young trainees are subjected to cosmetic surgery, strict diets, and intense exercise routines to achieve the 'perfect' body, in addition to their daily dance practices. They also face stringent and regular evaluations. Trainees are often expected to sacrifice their education and personal lives, committing fully to a life governed by strict rules and discipline (Wang, 2022).

A notable female idol group that has paved the way for success overseas in K-Pop, is "BLACKPINK". Since their debut in 2016 under YG Entertainment, members Jisoo, Jennie, Rosé, and Lisa have become international stars with their iconic style and strong presence. Their discography includes songs like: "DDU-DU DDU-DU," and "Kill This Love,". BLACKPINK's success reached the Western market, performing at Coachella in 2019, and again as their first K-Pop headliners in 2023 (Forbes).

One of the most popular female soloists in K-pop is singer "IU". IU debuted in 2008 under Kakao M and is known for her sweet visuals and striking voice. She has become an icon and role model for many young trainees and idols, with hits such as "Blueming". Her top chart hits have solidified her status as one of South Korea's most respected singers and actors, starring in the Korean Drama "Hotel De Luna", staying active and relevant even years after her debut (Forbes).

### **3.2. Stray Kids and Twice: An Introduction**

Two groups that have debuted under the same company, JYP Entertainment, are "STRAY KIDS" and "TWICE". These groups have withstood similar circumstances, such as becoming a group through a "survival show", where competitors would compete to solidify their spots in a new K-pop group. However, as STRAY KIDS is an eight-member male group, and TWICE is a nine-member female group, they have grown to undergo wildly different careers.

TWICE, which debuted in 2015, has become a highly popular girl group, their popularity reaching overseas. Known for their bubbly music, this group has consistently topped charts, as hits like "The Feels" and "Talk That Talk" have climbed up the Billboard Hot 100. Additionally, some members have become brand ambassadors, like Sana, who has become an ambassador for fashion brand "Celine" (Billboard, 2024).

STRAY KIDS, who debuted in 2017, has made their presence known with their self-produced music and dynamic performances by sub-unit, "3Racha". This group has set many records, such as their music video for "S-Class" which broke records for the fastest K-Pop music video to reach 100 million views on YouTube. With their strong sound and supportive fan base, their music reaches fans internationally, cementing their fame worldwide. Album, "Oddinary", topped the Billboard 200, making STRAY KIDS only the third-ever K-pop act to top the all-genre chart (Billboard, 2024).

### **4. Gender Inequality in K-pop Industry: On The Stage**

Due to South Korea's "cookie-cutter society", there is little room for true self-expression. In part 2 of this paper, women's place in South Korean society was established, where they are forced into a box, facing harsh criticism if not conforming to traditional gender roles, built to be inferior to men. In part 3, groundwork of the K-pop industry was established, discussing how idols are trained, and highlighting some of K-pop's biggest names in the industry. It also established how South Korea's music industry in K-pop, a melting pot of different music styles and concepts, and introduced male idol groups such as STRAY KIDS, who have gained much

attention for the flexibility of their style, while other female groups such as TWICE are criticized for exploring other styles in comparison.

South Korean idols have become the ideal standard of beauty, affecting both the audience and the idol's views on how they should act and present themselves to the public. On stage, female idols have had less say in their production of music, and continue to use the same concepts. In this part, each section will expose different inequalities of female idols in the performance and concepts of their music and careers. The first part will discuss the audience female and male idols attract and how it affects their marketing. The second section will expose the strict beauty standards female and male idols uphold, and the pressure female idols have to uphold this image. Lastly, the last section will explore the concepts female idols pursue versus male idols, and how female idols are often infantilized.

#### **4.1. Audience and Marketability**

To achieve success in an unpredictable industry, the audience's reception of idols' comebacks is critical. However, there are significant differences in how male and female idols cultivate and engage their respective audiences. Male idol fan base consists of predominantly teenage girls, who gravitate towards idols for the "boyfriend" fantasy they induce and give their male fan base a model to replicate. Companies exploit this fact, recognizing that this fanbase will be willing to purchase new albums, concert, and fan meet tickets, exclusive merchandise, and any other product that will support their favorite idols (Lee).

Additionally, male idols generally have more flexibility with how they act in the public eye. Though both female and male idols face scrutiny from their audience, male idols tend to show more personality and their "real personas" on and off stage, often being praised for showing glimpses of their "true selves" (Jonas, 2021). However, according to Jonas, on very rare occasions can female idols show their true personalities, often being viewed as "too much" or annoying. Female idols wear a mask in almost all public appearances, never truly showing their true selves to their audience (Jonas, 2021).

The demographic of female idols fanbase, compared to male idols, is more mixed, often leaving more room for hate from all ages and genders. While female fans follow female idols after their fashion and beauty, often comparing themselves to these idols as they reciprocate how they think they should look, male fans often follow female idols due to their reflection of a typical "girlfriend", making female idols adopt a "girl crush" and feminine concept through their discography (Jonas, 2021). Female fans often get an unachievable idea of how they should act and look in society, as female idols like I'VE's Wonyoung's beauty and lifestyle have become a trend on all social media platforms (Cao, 2024).

Contrast this with the male idol group STRAY KIDS whose marketing has varied from album to album, as throughout their 5 albums they have adapted several different "colors". From a romance album, "MAXIDENT", to a crazy and supernatural concept in "ODDINARY", their marketing has allowed them to show all sides of their talent. In the newest album "ATE", their concept photo shoot has shown a more playful side, as their photoshoot paired them with another

member out of the 8, doing several different "silly" poses. This was accepted well by the audience, as fans shared on social media how humorous it was (Dazed, 2019).

Additionally, in their variety show "SKZ: CODE", they show their humorous personas, being known to be very loud in the industry. This fact has not been seen as a negative, as many consume their media for entertainment beyond music (Koreaboo, 2021). However, a girl group under the same company, TWICE, has received much backlash for the same thing. For example, like many other girl groups, TWICE's variety show, "TWICE TV", shows members doing different activities.

However, unlike the response from "SKZ: CODE", some online found them to be "unprofessional" "attention-seeking" when showing excitement or being louder. This is in turn to women being expected to be in more traditional gender roles, listening to the men of the household and being submissive. Many fans find it repulsing when women are seen as not being "proper" or under control (Joo, 2021).

Additionally in TWICE's "SET ME FREE" promotional video, where members were seen taking off their makeup with bright smiles, and kindly interacting with each other, though many fans praised members for showing their natural beauty, many netizens criticized members. Many said that they did not want a reminder of how old they were, and did not see a point in taking off their makeup in a beauty-focused entertainment occupation. While some appreciated seeing a more relatable sense of the idols, many criticized them for not being "proper" or "idealized" (KoreaBoo, 2023).

Another large reflection of an idol's success is the outcome of award shows. Along with the big K-pop award shows like MAMA, AAA, and GDA Awards, where awards like "Album of the Year" and "Best Performance of the Year" are given, there are also weekly award shows like "MUSIC SHOW", where groups new music go against each other, with the songs with the most votes from the audience winning (Fashion Chingu, 2024). For example, in one of the largest K-pop awards shows of the year, Mnet Asian Music Awards (MAMA), shows results where male idols win many more awards than female idols.

Though this result can reflect the popularity of a group and songs, even the largest of female idols fail to win awards, perhaps due to the inactivity of their fanbase when voting compared to male idols. From 2007 to, male idols consistently dominated the major categories at the MAMA awards. In one of the most notable awards given every year, male groups, BTS and EXO took home a total of 9 wins (BTS taking the title five times (2016-2020) and EXO four times (2014-2017), while no female group won this award during the same period (MAMA, 2024).

Similarly, female idols rarely won in the Album of the Year category, as male groups dominated again. Though groups like TWICE and Girls' Generation had shown some success in Song of the Year, BTS and BIGBANG still won more frequently. Though more recently female idols have been winning this award more frequently like HYBE's group "NEW JEANS", male idols' global popularity and engaged fan base resulted in more frequent wins at major award shows (Fashion Chingu, 2024).

Female idols have shorter longevity in their careers when compared to male idols, where male idols have the option to continue working for a longer duration. This can be attributed to how female idols have been infantilized and made to show a sense of purity. Many fans often find distaste for female idols quicker, as fans often gravitate towards new and young female faces. This is the case for female groups especially as many fans are attracted to female idols due to their young and innocent persona, and as they age, they lose this image (Ahn, 2019). Female idol groups often disband after 7 years, as many believe they become irrelevant after turning 25 (Yang, 2024).

However, in comparison to female idols, male groups often maintain careers until their late twenties and even thirties (Jonas 2021). Regardless of their mandatory military leave, many male idols return to the stage, having long-term careers (Ahn, 2019). Though both female and male idols have a career with an early expiration date, awaiting another young group to replace them, male idols have a much longer career, often making entertainment companies debut primarily male idols (Jonas, 2021).

#### **4.2. Beauty Standards**

South Korea has remained a beauty-based society, where aesthetics are highly regarded. Most notably in the Joseon Dynasty (1392-1897), beauty standards had much to do with social class. Fair and unblemished skin showed class, as they weren't out working in the sun (Dağ, 2024). For men, however, these standards are still apparent but not as strict, the standard being tall, pale, and clean-shaven (Tugade, 2022).

If not intimidated by South Korean models or celebrities, family and friends often tell young girls of their flaws, even encouraging them to get surgical procedures done. This has resulted in young girls having low self-esteem, and wanting to conform to the homogenous society (Siu, 2021). If these are the norms in society, K-pop idols are under even more pressure to conform to these standards, being the epitome of beauty in the media. Many entertainment companies put their idols and trainees under intense diets while undergoing intense physical training. Usually before a comeback, idols will lose immense amounts of weight in short periods, consuming little food and water (Tugade, 2022). Though male idols also undergo strict diets, most go hand in hand with weight training and the consumption of more protein, though still little food is consumed. Compared to female idols, societal standards push slim legs and small waists. For example, TWICE's MOMO undergoes a rigorous diet of eating an ice cube daily to lose seven kilograms within a week (Cruz, 2023).

Additionally, it was found that in the early stages of STRAY KIDS and TWICE's careers, JYPE restricted TWICE members from using the cafeteria, making them adhere to their strict diet. However, STRAY KIDS's leader, Bang Chan, gave them his food coupons so they were allowed to eat meat. When questioned, he would respond with apologies that he ate a lot of meat that day (KoreaBoo, 2022).

When idols do not meet the beauty standard, they can face much scrutiny online, often leading to idols taking a "hiatus" to recover mentally. Though male idols do face discrimination

from gaining weight, or not fitting the beauty standards perfectly, their fanbase of both Western and South Korean females often support them throughout. However, as female idols have a larger demographic and only have the support of Western females primarily, they can face larger amounts of scrutiny when they gain weight or aren't seen as beautiful (Primer, 2023). For example, after TWICE's Jeongyeon returned from her hiatus due to a herniated disc and anxiety, with weight gain due to her medication, many Western fans welcomed her back, however, many South Korean fans criticized her for the change of appearance. Many called her "fat" and "ugly", saying she had no place for being an idol (KoreaBoo, 2024).

### **4.3. Performance**

In the K-pop industry, the performance concepts assigned to male and female idols starkly illustrate the unequal amount of concepts explored by both genders. Female idols usually explore one of three concepts: cute, grunge, or sexy. If they explore outside of this zone, they are often ridiculed online for being "unprofessional". However, male idols have much more flexibility with concepts, being able to convey the character of boyfriends, spies, gangsters, heroes, and more (Primer, 2023). For example, much of TWICE's music follows a bright and cute concept, like their hits "Cheer Up ", "TT ", and "Heart Shaker" (Sayson, 2021).

However, STRAY KIDS has had much flexibility when it comes to their songs, exploring many different genres like rap, rock, traditional Korean music, and more. Every album they release has a unique sound and meaning to their songs, ranging from heartfelt romance songs to boisterous-edgy songs detailing their confidence and willpower (Knight, 2021). With choreography, many male idols often incorporate styles like hip hop and b-boying, having flips and tricks, always keeping the idol moving. However, though many female idols come from professional dance backgrounds, they are limited to simple arm movements, with sexual implications (Jonas, 2021).

While male idols can sing about a multitude of topics like inspiring songs, romance, empowerment, standing out, and more, female idols usually sing about romance or cute concepts, often singing about being "hopeless without a man". When idols like popular girl group (G)I-DLE sing about topics that are more radical like not being a stereotypical woman in society, they are often attacked for being too edgy and "problematic".

In addition to the concepts, the outfits and choreography that are paired with these reflect the inequality of female and male idols. Female idols are typically dressed in less fabric, exposing their legs and torsos, regardless of age. However, men are more likely to be in outfits that can be covered up as much as desired, not going past an exposed midriff or arms typically. And whereas female idols can be criticized for showing too much skin, males are praised for doing it as "fan service" (Jonas, 2021).

Infantilization, defined as behaving or treating someone as they are a baby, is the norm in South Korean entertainment and society. Deemed as "aegyo", acting cute and baby-ish, idols are praised when doing aegyo as fanservice, regardless of age, keeping full-grown idols in a stage of youth. However, more intensely, the entertainment company keeps female artists in a constant



state of youth, making them seem "knowledgeable of wrong" and creating a fantasy of innocence.

Regardless of this infantilization, female idols are still sexualized even from a young age, as male idols do not have to perform to such concepts (Bhatkar, 2021). For example, girl group NEW JEANS has undergone public backlash from their song, "Cookie", with suggestive lyrics. As NEW JEANS members ranged from ages 14-18 at the time of release, many found it uncomfortable that the lyrics "Looking at my cookie / Do you ever smell it different? (Taste it) / What's with a bite, isn't enough?" were being sung by young girls. Many speculated that "cookie" referred to female genitalia, creating even more public criticism. However, company ADOR released a statement denying all claims (Yeo, 2022).

There is a noticeable disparity between female and male idols who have control of their music production. According to Korea's Official Copyright Association, or KOMCA, 15 out of 100 top idol producers are female, showing the involvement of female idols in the music-making process. More specifically, the STRAY KIDS's production unit, "3RACHA", has produced the majority of STRAY KIDS songs, making the group "self-produced" as they write, produce, and arrange their songs. Members Changbin, Han, and Bang Chan, chose the majority of the concepts of their albums (excluding their early debut days), having creative control of their music and being able to make songs of their desires. However, TWICE's songs have little say by members, being given by the companies (Bell, 2022). Though TWICE members have had more say in their music than other female groups, they only have contributed to about 40 of their over 200 songs as of 2024 (Billboard).

## **Conclusion**

The objective of this paper was to analyze how female K-pop idols face discrimination within the music industry due to their gender. By analyzing groups STRAY KIDS and TWICE's careers, marketing, performances, concepts, beauty standards, and overall treatment by the public, it can be concluded that female idols suffer discrimination in varied ways. First, on marketability, this paper found that compared to males, the demographic of female idols fanbase is larger, often leaving more room for backlash from all parts. It was also found that female idols have to uphold their image, reflecting a 'girlfriend' image for male fans, while becoming a reflection of how women want to look through fashion and beauty.

It was also found that female idols must meet high levels of beauty standards compared to male idols (big eyes, double eyelids, small faces, slim figures, full lips, and long legs) which creates unhealthy demands to achieve these standards. Lastly, compared to male idols, female idols face constraints in exploring concepts, are infantilized and sexualized—even when underage—which limits their creativity and career longevity, and they receive harsher criticism for expressing genuine emotions, often being labeled as "annoying" or "unprofessional."

As a final remark, this research reveals a significant systemic gender bias within the K-pop industry, likely rooted in South Korea's historical gender discrimination, and thus calls for societal reflection and change.

## Works Cited

- Ahn, Sung. "Why K-Pop Girl Groups Can't Stand the Test of Time." *The Jakarta Post*, 24 Jan. 2019, [www.thejakartapost.com/life/2019/01/24/why-k-pop-girl-groups-cant-stand-the-test-of-time.html](http://www.thejakartapost.com/life/2019/01/24/why-k-pop-girl-groups-cant-stand-the-test-of-time.html).
- Arnaud, Jourdan. "A Brief History of K-Pop." *The Los Angeles Film School*, 6 Apr. 2021, [www.lafilm.edu/blog/a-brief-history-of-kpop/](http://www.lafilm.edu/blog/a-brief-history-of-kpop/).
- Bell, Crystal. "3RACHA's Bang Chan, Changbin and Han Step into the Spotlight." *PAPER Magazine*, 14 Oct. 2022, [www.papermag.com/3racha#rebelltitem17](http://www.papermag.com/3racha#rebelltitem17).
- Bhatkar, Madhura. "The Dark Side of K-Pop (1)— Infantilization." *Medium*, 13 July 2021, [michtimadhura.medium.com/the-dark-side-of-k-pop-1-infantilization-1346c8968458](https://michtimadhura.medium.com/the-dark-side-of-k-pop-1-infantilization-1346c8968458).
- Billboard. "Stray Kids." *Billboard*, [www.billboard.com/artist/stray-kids/](http://www.billboard.com/artist/stray-kids/).
- . "Twice." *Billboard*, [www.billboard.com/artist/twice/](http://www.billboard.com/artist/twice/).
- Cao, Steffi. "'Wonyoung Ism': TikTok Trend Sees Teens Starving Themselves in the Name of K-Pop." *The Daily Beast*, *The Daily Beast*, 27 Apr. 2024, [www.thedailybeast.com/wonyoungism-tiktok-trend-sees-teens-starving-themselves-in-the-name-of-k-pop#:~:text=Jang%20is%20confident%2C%20smart%2C%20pretty](https://www.thedailybeast.com/wonyoungism-tiktok-trend-sees-teens-starving-themselves-in-the-name-of-k-pop#:~:text=Jang%20is%20confident%2C%20smart%2C%20pretty).
- Center, Human Rights Research. "South Korea's Hurdles with Gender Equality." *HRRC*, 27 Feb. 2024, [www.humanrightsresearch.org/post/south-korea-s-hurdles-with-gender-equality](http://www.humanrightsresearch.org/post/south-korea-s-hurdles-with-gender-equality).
- Connell, R. W. *Masculinities*. 2nd ed., Cambridge Polity, 2005.
- Cruz, Baleria De La. "Toxic Beauty Standards of Korean Idols." *Medium*, 6 Dec. 2023, [medium.com/past-present-pop/toxic-beauty-standards-of-korean-idols-5bdcf5e4ab83](https://medium.com/past-present-pop/toxic-beauty-standards-of-korean-idols-5bdcf5e4ab83).
- Cubit, Brea. "Taylor Swift Gets Candid about Lover, Scooter Braun, and Sexism in the Music Industry." *Popsugar*, 25 Aug. 2019, [www.popsugar.com/celebrity/taylor-swift-cbs-sunday-morning-interview-video-46539440](http://www.popsugar.com/celebrity/taylor-swift-cbs-sunday-morning-interview-video-46539440).
- Dağ, Betül. "Korean Beauty Standards: Trends, Influences, and Critiques." *Lingopie*, *Lingopie*, 21 June 2024, [lingopie.com/blog/korean-beauty-standards/#:~:text=In%20South%20Korea%2C%20beauty%20standards](https://lingopie.com/blog/korean-beauty-standards/#:~:text=In%20South%20Korea%2C%20beauty%20standards).
- Dance Vision. "What Is K-Pop." *Blog.dancevision.com*, 12 Feb. 2024, [blog.dancevision.com/what-is-kpop](http://blog.dancevision.com/what-is-kpop).
- Dazed. "Stray Kids Are Shaking up K-Pop's Status Quo." *Dazed*, 5 June 2019, [www.dazeddigital.com/music/article/44726/1/stray-kids-k-pop-group-interview](http://www.dazeddigital.com/music/article/44726/1/stray-kids-k-pop-group-interview).
- Fashion Chingu. "KPOP Award Shows Explained - Fashion Chingu." *Fashion Chingu*, 2 Feb. 2024, [www.fashionchingu.com/blog/kpop-mama-other-awards-explained/?srsltid=AfmBOooEAZs4WYROSzTarS8TKbQRL04vr\\_20yRTE1TJDetev\\_keNillR](http://www.fashionchingu.com/blog/kpop-mama-other-awards-explained/?srsltid=AfmBOooEAZs4WYROSzTarS8TKbQRL04vr_20yRTE1TJDetev_keNillR).
- Forbes. "BLACKPINK." *Forbes*, [www.forbes.com/profile/blackpink/](http://www.forbes.com/profile/blackpink/).
- . "Ji-Eun Lee (IU)." *Forbes*, [www.forbes.com/profile/ji-eun-lee-iu/](http://www.forbes.com/profile/ji-eun-lee-iu/).
- Jonas, Liz. *Crafted for the Male Gaze: Gender Discrimination in the K-Pop Industry*. 2021, [vc.bridgew.edu/cgi/viewcontent.cgi?article=2492&context=jiws](http://vc.bridgew.edu/cgi/viewcontent.cgi?article=2492&context=jiws).
- Joo, Young Hyun Lily. "The Plight of Gender Inequality in South Korea." *Inkstick*, 13 Apr. 2021, [inkstickmedia.com/the-plight-of-gender-inequality-in-south-korea/](http://inkstickmedia.com/the-plight-of-gender-inequality-in-south-korea/).
- Jung, Hawon. "This Women's Day, There Is Little to Celebrate in South Korea." *Aljazeera*, 8 Mar. 2023, [www.aljazeera.com/opinions/2023/3/8/this-womens-day-there-is-little-to-celebrate-n-south-korea](http://www.aljazeera.com/opinions/2023/3/8/this-womens-day-there-is-little-to-celebrate-n-south-korea).

- Knight, Rinn. "Cookin' up a New Sound: Stray Kids and the Mala Taste Genre." Medium, 24 July 2021, [medium.com/outro-words/cookin-up-a-new-sound-stray-kids-and-the-mala-taste-genre-a8e4e8427bcc](https://medium.com/outro-words/cookin-up-a-new-sound-stray-kids-and-the-mala-taste-genre-a8e4e8427bcc).
- Koreaboo. "Stray Kids Have Fun in the Snow for First Episode of 'SKZ CODE.'" Koreaboo, 6 Feb. 2021, [www.koreaboo.com/stories/stray-kids-play-in-snow-skz-code/](http://www.koreaboo.com/stories/stray-kids-play-in-snow-skz-code/).
- . "TWICE Is Being Praised for Their Groundbreaking 'No Makeup' Music Video Teaser — but Not Everyone Is a Fan." Koreaboo, 7 Mar. 2023, [www.koreaboo.com/news/twice-praised-makeup-free-music-teaser-not-everyone-fan/](http://www.koreaboo.com/news/twice-praised-makeup-free-music-teaser-not-everyone-fan/).
- . "TWICE Jeongyeon's Appearance Triggers Viral Concern Agai." Koreaboo, 3 Aug. 2024, [www.koreaboo.com/news/twice-jeongyeon-appearance-triggers-viral-concern/](http://www.koreaboo.com/news/twice-jeongyeon-appearance-triggers-viral-concern/).
- . "TWICE Was Forced to Diet as Trainees, but Stray Kids' Bang Chan Found a Clever Way to Sneak Them Food." Koreaboo, 15 Oct. 2022, [www.koreaboo.com/news/twice-diet-stray-kids-bang-chan-sneak-food/](http://www.koreaboo.com/news/twice-diet-stray-kids-bang-chan-sneak-food/).
- Korean Overseas Information Service (KOIS). "Women's Role in Contemporary Korea." Asia Society, 2019, [asiasociety.org/education/womens-role-contemporary-korea](http://asiasociety.org/education/womens-role-contemporary-korea).
- Kroll, Piper, and 2024. "The Rise of K-Pop." Garnet & Black Magazine, 29 Jan. 2024, [www.gandbmagazine.com/article/2024/01/the-rise-of-kpop-kroll](http://www.gandbmagazine.com/article/2024/01/the-rise-of-kpop-kroll).
- Lee, Sophia. "Idolizing Toxicity: K-POP." The Exonian, [theexonian.net/opinions/idolizing-toxicity-k-pop](http://theexonian.net/opinions/idolizing-toxicity-k-pop).
- "List of Songs Recorded by Twice." Wikipedia, 30 Apr. 2024, [en.wikipedia.org/wiki/List\\_of\\_songs\\_recorded\\_by\\_Twice](https://en.wikipedia.org/wiki/List_of_songs_recorded_by_Twice).
- Lorber, Judith. *The Variety of Feminisms and Their Contributions to Gender Equality*. Bis, 1997.
- MAMA. "MAMA HISTORY." Mnetplus.world, 2024, [mama.mnetplus.world/history/2023/winners](http://mama.mnetplus.world/history/2023/winners).
- McRobbie, Angela. "Post-Feminism and Popular Culture." *Feminist Media Studies*, vol. 4, no. 3, Nov. 2004, pp. 255–64, <https://doi.org/10.1080/1468077042000309937>.
- Messerlin, Patrick A., and Wonkyu Shin. "The Success of K-Pop: How Big and Why so Fast?" *Asian Journal of Social Science*, vol. 45, no. 4-5, 2017, pp. 409–39, <https://doi.org/10.1163/15685314-04504003>.
- Palley, Marian Lief. "Women's Status in South Korea: Tradition and Change." *Asian Survey*, vol. 30, no. 12, Dec. 1990, pp. 1136–53, <https://doi.org/10.2307/2644990>.
- Pratamasari, Annisa. "(PDF) International Business Strategy in Selling Korean Pop Music: A Case Study of SM Entertainment." ResearchGate, Dec. 2017, [www.researchgate.net/publication/324446216\\_International\\_Business\\_Strategy\\_in\\_Selling\\_Korean\\_Pop\\_Music\\_A\\_Case\\_Study\\_of\\_SM\\_Entertainment](http://www.researchgate.net/publication/324446216_International_Business_Strategy_in_Selling_Korean_Pop_Music_A_Case_Study_of_SM_Entertainment).
- Primer, Payton. *How Gender Shapes Music: A Comparison within K-Pop*. How Gender Shapes Music: A Comparison within K-Pop Payton Primer. 1 Feb. 2023, [research.library.fordham.edu/cgi/viewcontent.cgi?article=1106&context=international\\_senior](http://research.library.fordham.edu/cgi/viewcontent.cgi?article=1106&context=international_senior).

- Rampton, Martha. "Four Waves of Feminism." Pacific University, Pacific University, 25 Oct. 2015, [www.pacificu.edu/magazine/four-waves-feminism#:~:text=The%20first%20wave%20of%20feminism.](http://www.pacificu.edu/magazine/four-waves-feminism#:~:text=The%20first%20wave%20of%20feminism.)
- Sayson, Sofia. "How TWICE Has Changed the K-Pop Game: Their Successful Journey as Artists." Medium, 18 Aug. 2021, [sofia-sayson.medium.com/how-twice-has-changed-the-game-their-successful-journey-eadec060e3cf](https://sofia-sayson.medium.com/how-twice-has-changed-the-game-their-successful-journey-eadec060e3cf).
- Siu, Cathy. "Korean Beauty Standards – the LION." The Lion, 11 Dec. 2021, [thelion.sites.lmu.edu/miscellaneous/korean-beauty-standards/](http://thelion.sites.lmu.edu/miscellaneous/korean-beauty-standards/).
- Thompson, Becky. "Multiracial Feminism: Recasting the Chronology of Second Wave Feminism." *Feminist Studies*, vol. 28, no. 2, 2002, pp. 336–60.
- Tugade, Lauryn. "The Effects of Korean Beauty Standards on Korean Pop Idols." Kauai Community College, [www.kauai.hawaii.edu/sites/www.kauai.hawaii.edu/files/pdf/events/CreativeWritingContest\\_pdf/Tugade\\_EffectsKoreanBeautyStandards.pdf](http://www.kauai.hawaii.edu/sites/www.kauai.hawaii.edu/files/pdf/events/CreativeWritingContest_pdf/Tugade_EffectsKoreanBeautyStandards.pdf).
- Walker, Lauren. "Sexism in the Music Industry – How Far Have We Really Come?" TEDx Warwick, 26 Jan. 2022, [www.tedxwarwick.com/post/sexism-in-the-music-industry-how-far-have-we-really-come](http://www.tedxwarwick.com/post/sexism-in-the-music-industry-how-far-have-we-really-come)
- Wang, Zhaoqi. "Research on Korean Idol-Making Industry Based on Culture Industry Theory." Research Gate, Dec. 2022, [www.researchgate.net/publication/368485469\\_Research\\_on\\_Korean\\_Idol-Making\\_Industry\\_Based\\_on\\_Culture\\_Industry\\_Theory](http://www.researchgate.net/publication/368485469_Research_on_Korean_Idol-Making_Industry_Based_on_Culture_Industry_Theory).
- Yang, Arlina. "5 K-Pop Groups That Could Potentially Disband in 2024." AsAmNews, 8 Jan. 2024, [asamnews.com/2024/01/08/7-year-curse-contract-talks-k-pop-groups-in-danger-of-disbanding/#:~:text=The%20K%2Dpop%20industry%20is](https://asamnews.com/2024/01/08/7-year-curse-contract-talks-k-pop-groups-in-danger-of-disbanding/#:~:text=The%20K%2Dpop%20industry%20is).
- Yeo, Gladys. "NewJeans' Label ADOR Responds to Criticism of 'Sexual' Lyrics in 'Cookie' with Extensive Statement." NME, 29 Aug. 2022, [www.nme.com/news/music/newjeans-cookie-sexual-lyrics-controversy-ador-label-responds-statement-3299682](http://www.nme.com/news/music/newjeans-cookie-sexual-lyrics-controversy-ador-label-responds-statement-3299682).
- Korean Overseas Information Service (KOIS). "Women's Role in Contemporary Korea." Asia Society, 2019, [asiasociety.org/education/womens-role-contemporary-korea](http://asiasociety.org/education/womens-role-contemporary-korea).
- Kroll, Piper, and 2024. "The Rise of K-Pop." Garnet & Black Magazine, 29 Jan. 2024, [www.gandbmagazine.com/article/2024/01/the-rise-of-kpop-kroll](http://www.gandbmagazine.com/article/2024/01/the-rise-of-kpop-kroll).
- Lee, Sophia. "Idolizing Toxicity: K-POP." The Exonian, [theexonian.net/opinions/idolizing-toxicity-k-pop](http://theexonian.net/opinions/idolizing-toxicity-k-pop).
- Lorber, Judith. *The Variety of Feminisms and Their Contributions to Gender Equality*. Bis, 1997.
- MAMA. "MAMA HISTORY." Mnetplus.world, 2024, [mama.mnetplus.world/history/2023/winners](http://mama.mnetplus.world/history/2023/winners).
- McRobbie, Angela. "Post-Feminism and Popular Culture." *Feminist Media Studies*, vol. 4, no. 3, Nov. 2004, pp. 255–64, <https://doi.org/10.1080/1468077042000309937>.

- Messerlin, Patrick A., and Wonkyu Shin. "The Success of K-Pop: How Big and Why so Fast?" *Asian Journal of Social Science*, vol. 45, no. 4-5, 2017, pp. 409–39, <https://doi.org/10.1163/15685314-04504003>.
- Palley, Marian Lief. "Women's Status in South Korea: Tradition and Change." *Asian Survey*, vol. 30, no. 12, Dec. 1990, pp. 1136–53, <https://doi.org/10.2307/2644990>.
- Pratamasari, Annisa. "(PDF) International Business Strategy in Selling Korean Pop Music: A Case Study of SM Entertainment." *ResearchGate*, Dec. 2017, [www.researchgate.net/publication/324446216\\_International\\_Business\\_Strategy\\_in\\_Selling\\_Korean\\_Pop\\_Music\\_A\\_Case\\_Study\\_of\\_SM\\_Entertainment](http://www.researchgate.net/publication/324446216_International_Business_Strategy_in_Selling_Korean_Pop_Music_A_Case_Study_of_SM_Entertainment).
- Primer, Payton. *How Gender Shapes Music: A Comparison within K-Pop*. How Gender Shapes Music: A Comparison within K-Pop Payton Primer. 1 Feb. 2023, [research.library.fordham.edu/cgi/viewcontent.cgi?article=1106&context=international\\_senior](http://research.library.fordham.edu/cgi/viewcontent.cgi?article=1106&context=international_senior).
- Rampton, Martha. "Four Waves of Feminism." *Pacific University*, Pacific University, 25 Oct. 2015, [www.pacificu.edu/magazine/four-waves-feminism#:~:text=The%20first%20wave%20of%20feminism](http://www.pacificu.edu/magazine/four-waves-feminism#:~:text=The%20first%20wave%20of%20feminism).
- Sayson, Sofia. "How TWICE Has Changed the K-Pop Game: Their Successful Journey as Artists." *Medium*, 18 Aug. 2021, [sofia-sayson.medium.com/how-twice-has-changed-the-game-their-successful-journey-eadec060e3cf](https://sofia-sayson.medium.com/how-twice-has-changed-the-game-their-successful-journey-eadec060e3cf).
- Siu, Cathy. "Korean Beauty Standards – the LION." *The Lion*, 11 Dec. 2021, [thelion.sites.lmu.edu/miscellaneous/korean-beauty-standards/](http://thelion.sites.lmu.edu/miscellaneous/korean-beauty-standards/).
- Thompson, Becky. "Multiracial Feminism: Recasting the Chronology of Second Wave Feminism." *Feminist Studies*, vol. 28, no. 2, 2002, pp. 336–60.
- Tugade, Lauryn. "The Effects of Korean Beauty Standards on Korean Pop Idols." *Kauai Community College*, [www.kauai.hawaii.edu/sites/www.kauai.hawaii.edu/files/pdf/events/CreativeWritingContest\\_pdf/Tugade\\_EffectsKoreanBeautyStandards.pdf](http://www.kauai.hawaii.edu/sites/www.kauai.hawaii.edu/files/pdf/events/CreativeWritingContest_pdf/Tugade_EffectsKoreanBeautyStandards.pdf).
- Walker, Lauren. "Sexism in the Music Industry – How Far Have We Really Come?" *TEDx Warwick*, 26 Jan. 2022, [www.tedxwarwick.com/post/sexism-in-the-music-industry-how-far-have-we-really-come](http://www.tedxwarwick.com/post/sexism-in-the-music-industry-how-far-have-we-really-come).
- Wang, Zhaoqi. "Research on Korean Idol-Making Industry Based on Culture Industry Theory." *Research Gate*, Dec. 2022, [www.researchgate.net/publication/368485469\\_Research\\_on\\_Korean\\_Idol-Making\\_Industry\\_Based\\_on\\_Culture\\_Industry\\_Theory](http://www.researchgate.net/publication/368485469_Research_on_Korean_Idol-Making_Industry_Based_on_Culture_Industry_Theory).
- Yang, Arlina. "5 K-Pop Groups That Could Potentially Disband in 2024." *AsAmNews*, 8 Jan. 2024, [asamnews.com/2024/01/08/7-year-curse-contract-talks-k-pop-groups-in-danger-of-disbanding/#:~:text=The%20K%2Dpop%20industry%20is](https://asamnews.com/2024/01/08/7-year-curse-contract-talks-k-pop-groups-in-danger-of-disbanding/#:~:text=The%20K%2Dpop%20industry%20is).
- Yeo, Gladys. "NewJeans' Label ADOR Responds to Criticism of 'Sexual' Lyrics in 'Cookie' with Extensive Statement." *NME*, 29 Aug. 2022, [www.nme.com/news/music/newjeans-cookie-sexual-lyrics-controversy-ador-label-responds-statement-3299682](http://www.nme.com/news/music/newjeans-cookie-sexual-lyrics-controversy-ador-label-responds-statement-3299682).

# **Applying Machine Learning Within Augmented Reality As A Tool For Various Industries**

## **By Karan Jain**

### **Abstract**

Machine Learning and Augmented Reality offer various strengths and weaknesses. This study aimed to understand factors to consider when integrating ML and AR in real-world contexts. The research explored both technologies and included a survey to understand user trends. Limitations include not discussing specific and country-wise implications.

This paper found that AR is less popular than ML, with specific claims for integration found in Historical Understanding, Consumer Interaction, Media, Language, and Healthcare. The study delved into technologies like convolutional neural networks, machine translation, SLAM, and NPC to understand integration for innovative applications in select sectors.

### **Background Information**

Augmented Reality (AR) is an interactive, real-time overlay of virtual computer-generated media within the physical environment. AR technologies use cameras, sensors and displays, and they can capture the physical environment, integrate visual media into it, and, lastly, provide an interactive, enhanced user experience of the real world. Such technologies include mobile phones, tablets, and Head-Mounted Displays, such as the Apple Vision Pro. Unlike Virtual Reality, the purpose of AR is not to isolate the user from the real world but rather to add digital elements to it.

The type of AR technology determines how the digital content will be displayed on the display to the user. Marker-based AR: This form of AR uses a physical trigger such as markers and QR codes that, when triggered through a camera detection, overlay pre-programmed digital content on the display. Marker-less AR: This form of AR maps the physical environment through recognition algorithms, computer vision, and device sensors such as a GPS or a camera. Hence, Marker-less AR can correspondingly decide where to locate digital content based on its familiarisation with the physical environment without the need for a marker mechanism.

Moreover, AR technology has many applications, notably in the following fields: healthcare, education, gaming, manufacturing, navigation, tourism, and retailing. Therefore, AR solutions are applicable in a wide range of contexts, which opens up various opportunities for impactful developments in the field. Machine Learning is a technology that dictates the behaviour of computer systems based on experience through the use of algorithms. ML is a subset of AI (Artificial Intelligence) and can use valuable data that is provided to systems to enhance the behaviour of the system to meet its core purpose. In 2024, there are a total of 5.35 billion internet users.

Moreover, according to the Global Web Index, the average Internet user spends six and a half hours on the Internet every day. Therefore, there is a large amount and variety of data that is extracted from users. To analyse and effectively use this data to improve computer systems,

Machine Learning is critical. ML can allow the development of computer systems that possess a higher degree of intelligence in their respective fields.

We can categorise data into four categories: Structured, Semi-structured, Unstructured, and Metadata. The effectiveness of a Machine Learning solution is dependent on the type of data that is provided to the system. Moreover, the type of Machine Learning algorithm used also depends on the application of the solution. Commonly, Machine Learning solutions have been applied in fields such as Healthcare, smart cities, and recommendation systems. The four main types of Machine Learning techniques have different purposes:

- ***Supervised:*** This is a task-driven technique that uses sample input-output pairs from the given training data to predict other input-output pairs. Classification and Regression are the two types of Supervised learning.
- ***Unsupervised:*** This is a data-driven technique that is used for finding insights from unlabeled data. Clustering and Association are the most common types of Unsupervised learning.
- ***Semi-supervised:*** This is an amalgamation of supervised and unsupervised techniques that are used for both labelled and unlabeled data. Semi-supervised learning is used to generate more accurate predictions than supervised learning, which operates on labelled data. Classification and Clustering are the common techniques for semi-supervised learning.
- ***Reinforcement:*** A reinforcement technique is an environment-driven technique that uses a reward and penalty system to generate insights on the environment and, therefore, act to achieve the most optimal result by training the software. It has applications in robotics.

This research paper will identify how Machine Learning algorithms and techniques can be applied within Augmented Reality technologies in various contexts to provide new opportunities. Moreover, various Machine Learning algorithms are used, such as:

- Random Forest
- Naive Bayes
- Logistic Regression
- K-nearest Neighbours
- Support Vector Machine

### **Literature Review Brief**

The study “Augmented Reality: Applications, Challenges and Future Trends” notes restraining factors in Augmented Reality technology that need to be developed. It also states the

importance of recognition systems in Augmented Reality technology. It also explores the application of Augmented Reality technologies, which is critical to this study. Moreover, the study “Machine Learning: Algorithms, Real-World Applications and Research Directions” lists critical areas of implementation of Machine Learning. Therefore, these two sources are critical for understanding the possible applications that will be explored in this research paper.

Existing research on implanting Machine Learning within Augmented Reality shows that this integration can be applied to assistive technologies, as stated in “Integrating Machine Learning with Augmented Reality for Accessible Assistive Technologies”. This study also notes that this integration can increase the scope of AR technologies and can help visually impaired individuals locate objects. This previous research explores the integration of machine vision and speech recognition. However, this study does not focus on other areas of implication.

Overall, the sources are well credible and documented, and they provide important information that will be combined and used to help this research paper. The full literature review of all sources is below.

## **Research Questions**

Since this research paper focuses on various applications of Machine Learning and Augmented Reality technology, there are a variety of research questions.

- How can machine learning algorithms facilitate an increase in the educational outcome of interactive AR education technologies?
- How can machine learning algorithms provide insights into historical locations?
- How can machine learning algorithms be applied in AR shopping technologies that allow users to visualise items before buying?
- How can AR games improve interactivity through the implementation of machine learning algorithms?
- How can online entertainment consumption using AR technology be personalised through predictive machine learning algorithms?
- How can real-time language translation using HMDs be improved using improved machine translation algorithms
- How can machine learning algorithms implemented in AR technologies help consumer businesses to gain customers?

## **Objectives**

The objectives of this research paper include Identifying areas of application for ML within AR infrastructure. This is required to provide a real-world context to the applications of machine learning algorithms within AR. Identifying how ML can be of advantage in those sectors. Next, it is critical to identify areas where an AR ML solution can help. It is important to identify who the solution is going to help and how. Finding a process for implementing such a solution. Identifying future areas of research. Future areas of research can help take this integration to the next level, which can lead to future benefits in the application sectors.



Additionally, this paper aims to define the main algorithms and computational infrastructure used to support both machine learning and augmented reality technologies.

### **Relevance Of Study**

The study is critical as an integration solution has many possible benefits. Firstly, machine learning is one of the fastest-growing technologies on the planet. It has a variety of applications in countless fields. With the global machine learning market aimed to hit a \$209B valuation by 2029, the market has proved itself to be one of the most critical aspects of global technology in the 21st century. Moreover, AR technologies are being developed at a rapid pace and have a variety of applications. AR technologies provide new opportunities to users and can provide a unique user experience. It is critical to evaluate the role of machine learning in this experience as it can affect the way users interact with technologies. If machine learning can help users to the most efficient degree, a more fruitful outcome will be achieved for both the user and the product business.

Moreover, this study is relevant as AR tools have numerous practical benefits. They can be used to train individuals in their respective fields. Through machine learning, AR technology can be taken to new heights. Overall, these aspects aim to impact global industries that incorporate either AR or ML solutions. This study will leverage the advantages of both technologies and provide a combination of both technologies to provide the best possible result in various sectors.

Lastly, due to the great power that machine learning has in solving complex computational problems, it has proved to be of great significance in areas that require attention. Such areas include healthcare. If we can find solutions that integrate augmented reality in such fields, we will be able to find ways of visualising, interacting and computing/solving complex real-world problems. This will allow us to solve more problems and make major developments in global sectors.

### **Research Design**

This investigation aims to understand the implications of AR and ML algorithms within real-life contexts. Therefore, a research investigation will be conducted to collect relevant data on the matter from real-life users. This research will be supported by existing data to make accurate conclusions about how AR and ML can be used in real life.

This data will encompass a variety of possible solutions that AR and ML can provide and will give user feedback and concerns regarding this matter. As the end user is a critical part of the system development cycle, this data will allow us to understand the requirements and concerns of real-life users. This will help us develop more helpful AR and ML solutions. This investigation will contain self-collected data and existing data, as stated before. The self-collected data will be analysed and will be supported by the existing data. Machine Learning is a vast term of architectures that may include deep learning neural networks as well for

complex data situations. Given the vast scope of Machine Learning, there are various applications within AR technologies.

**Integration Research**

- **Computer Vision:** Through Computer Vision technologies, users can understand and interpret insights on their physical environment through the help of Machine Learning algorithms that facilitate an environment that is aimed at analysing visual components. Through this, devices overall digitised content onto a physical environment that can be viewed via head-mounted displays. Lastly, image recognition can help in a variety of fields, such as education and commerce. Computer Vision AR, along with Machine Learning, is hence able to gain insights about our environment. As every application for a technology is dependent on environmental factors, Computer Vision can help provide meaningful insights for all application sectors.
- **Personalisation:** Machine Learning algorithms can understand user behaviour and interaction with AR technologies and hence adapt them to provide a better user experience. Again, this has universal implications for all application sectors for both Machine Learning and Augmented Reality Through pattern identification via Neural Networks, new insights can be generated on specific target groups who use certain technologies. This can allow for a better understanding of how different demographics interact with Augmented Reality technology, which can help provide new avenues of research and development in the future.
- **Efficiency:** Machine Learning algorithms can better understand user behaviour and achieve pattern recognition, hence becoming more efficient. The management of resources also becomes more efficient due to Machine Learning architectures. Hence, a more resource-efficient final solution is achieved. This has benefits universally for all sectors of application.
- **Spatial Mapping:** As stated previously, computer vision can encourage environmental insights. Moreover, Spatial Mapping is an advantage of Augmented Reality. Special Mapping is critical to understanding the physical environment and, hence, accurately overlaying digital content onto it. Spatial mapping is critical to provide a good user experience. Moreover, through Machine Learning algorithms, AR technologies can achieve pattern and visual recognitions To overlay content more accurately in real time.

**Evaluation Of Machine Learning And Augmented Reality Integration**

<i>Advantages</i>	<i>Disadvantages</i>
Improved User Experience	Data security

Improved computational efficiency	Ethical usage concerns
Significant benefits within the applications sectors	Cost factors
	Availability of the technologies

**Investigation Data Collection Methods**

- **Online Forms:** Google Forms will be used to collect Quantitative data from a large group of respondents. The form will include MCQs (Multiple Choice Questions) and 1-10 rating questions.

The main purpose of the online forms is to understand the current usage of Augmented Reality and Machine Learning technologies To adapt and provide a synthesised solution as per current usage while keeping the development possibilities in mind.

**Data Analysis Procedures (Quantitative Data Analysis)**

Descriptive statistics will be used through mathematical correlation to find the relationship and trend in the user’s responses. This will allow us to infer accurate conclusions that consider all user responses and not just modal responses. Furthermore, graphical representations will be used to aid the correlation explorations.

After the data analysis and conclusions, the applications of Machine Learning with Augmented Reality will be organised and represented. These will be supported using the primary and secondary research conducted in the prior phases and will also include possible developments in select fields of application based on the research.

**Ethical Considerations**

All respondents who participate in the investigation will be informed of the aim and purpose of the investigation before their participation. They will give consent to use their data for the developments within this investigation. Moreover, they will be given a choice to participate in the data collection (voluntary participation).

Lastly, user data will be kept confidential and not shared with any external, third-party individual or group. This will ensure an ethical investigation that will improve the investigation by supporting conventional data collection methods that support data confidentiality and participant consent.

**Purpose**

The selected four areas of research for integration are Historical Understanding, Consumer Interaction, Media and Language, and lastly Commerce and Healthcare. Firstly, secondary research will be conducted on each of the areas of application with an evaluation. Next, the collected data will be used to support and deepen the analysis.

## Data And Analysis

### Internal Data

S. No	Question
1	<b>Demographics</b> <ul style="list-style-type: none"><li>• What is your age group</li><li>• Do you have an occupation?</li><li>• What level of education have you completed?</li></ul>
2	<b>Usage</b> <ul style="list-style-type: none"><li>• Have you ever used Augmented Reality before?</li><li>• In what context do you mostly use AR?</li><li>• At what frequency do you use AR?</li><li>• Have you ever used Machine Learning before?</li><li>• In what context do you mostly use Machine Learning?</li><li>• At what frequency do you use Machine Learning?</li></ul>
3	<b>Perceptions</b> <ul style="list-style-type: none"><li>• Rate your experience with AR technologies</li><li>• Rate your experience with ML technologies</li><li>• Based on your experience, can ML improve AR technologies?</li><li>• How can AI / ML improve your AR experience?</li></ul>
4	<b>Application Contexts</b> <ul style="list-style-type: none"><li>• In what sectors do you think this integration can be useful?</li><li>• Do you think your user interaction will increase when ML is added to AR technologies?</li></ul>
5	<b>Challenges</b> <ul style="list-style-type: none"><li>• Are you concerned by any major issues with AR and ML integration?</li><li>• Do you think either of these technologies are currently widely accessible to you?</li></ul>
6	<b>User Feedback</b> <ul style="list-style-type: none"><li>• How important is user feedback for the future development of AR and ML technologies?</li><li>• How strong is your ability to provide user feedback on AR and ML?</li></ul>
7	<b>Cases A: Historical Understanding</b> <ul style="list-style-type: none"><li>• How helpful do you think Machine Learning will be in education?</li><li>• As a tourist, how helpful are machine learning insights when you explore historical destinations?</li></ul>
8	<b>Case B: Consumer Interaction</b> <ul style="list-style-type: none"><li>• How helpful do you think Machine Learning can be for you when shopping?</li><li>• How helpful do you think Machine Learning is for interactivity in AR games?</li></ul>
9	<b>Case C: Media and Language</b> <ul style="list-style-type: none"><li>• How helpful do you think AR immersive media consumption can benefit from machine learning algorithms?</li><li>• How helpful do you think machine learning will help with language translations through AR technologies for everyday use?</li></ul>
10	<b>Case D: Healthcare</b> <ul style="list-style-type: none"><li>• How helpful do you think ML and AR will be in healthcare as a regular citizen?</li></ul>

## **External Data**

To gain insights from external resources regarding the matter, the 2017 findings from The Royal Study titled “Public views of Machine Learning” were consulted. The survey contacted 978 individuals from the UK and gathered critical data regarding the public’s perception of Machine Learning to better understand the benefits of Machine Learning. The surveys focused on areas of application such as health, social care and education. The survey collected detailed qualitative and quantitative data to gain balanced insights into the matter. For this research, the quantitative data will be analysed.

## **General Awareness**

The survey found that only 9% of individuals had heard the term “Machine Learning” - a low statistic. Moreover, they found that 3% of individuals knew about the technology to an extent. Next, they asked the 3% of individuals about the context in which they obtained awareness regarding Machine Learning. The results were as follows.

They found that 76% of individuals are aware of computers that can recognise human speech and give answers. Hence, ASR (Automatic Speech Recognition) through Machine Learning algorithms is currently the most known form of Machine Learning. Further, this application provides convenience benefits to users as spontaneous feedback can help improve user experiences. As this is the main aspect that users are aware of, it shows that there is currently a wide adoption of ASR algorithms in the market - indicating a strong presence of the same in various application fields. Notably, driverless vehicles come close to 75% of respondents. This shows the usage of Machine Learning in transportation.

To conclude, this shows the current high usage and developments of Machine Learning in transportation and linguistics. These are, hence, application sectors that Machine Learning can help improve within AR technologies, given the current popularity of the same.

## **Machine Learning Evaluation**

The survey shows that the majority of responses (36%) find the benefits and risks of Machine Learning to be equal. This indicates that there is currently no objection towards Machine Learning algorithms that could affect their adoption. As there is no widespread awareness of their benefits either due to the balanced statistics, this indicates that individuals are not yet aware of Machine Learning to a greater extent.

In context with Augmented Reality technologies, this shows that the integration of Machine Learning would require user documentation to raise awareness amongst individuals. This will allow users to make more informed decisions and could affect the adoption and popularity of integral Machine Learning and Augmented Reality technologies within the market.

## **Most Beneficial Applications**

The top 2 areas of application that users found were the most beneficial were - Facial recognition to catch criminals (61%) and computational speech recognition and question

answering (54%). Both sectors were also the only ones to have a majority of >50% of respondents feeling that the benefits outweigh the risks.

In connection with the Augmented Reality integration, this shows that the general public would be adopting AR ML integrations for criminal justice and user interaction. Hence, Augmented Reality technologies in both of these respective fields would have significant benefits compared to risks.

### **Risk And Usefulness Trend**

The survey found a negative linear relation between the usefulness of a machine-learning solution and its risk. This hence indicates that more useful solutions are less risky and less useful solutions are riskier. According to the survey, the top 3 areas of application with the highest usefulness and corresponding least risk are Healthcare - to spot breast cancer patterns, Banking – to detect fraud, and lastly, Travel - to find the best routes as per traffic.

Hence, this identifies the three sectors (healthcare, banking and travel) as the most useful Machine Learning integrations per the general public. Therefore, Augmented Reality architectures and technologies can also be developed with Machine Learning to elevate their significance and importance to the users.

### **Internal Data**

Demographics - The form had a total of 101 respondents of varying age groups. It was found that the majority of respondents were in the 35 - 45 age group (28.7%). The survey had a balanced variety of individuals from various age groups. Graphical results of the form are present in the Appendix. It was found that the majority of respondents (80.2%) had an occupation and, therefore, working in current industries could benefit from Machine Learning or Augmented Reality. Next, 82% of individuals had completed a university-level education. Hence, this proves that a majority of the respondents would be aware of both technologies.

Usage - Next, data on the usage, context and frequency of usage of Machine Learning and Augmented Reality was collected. The data is shown below:

	<b>AR</b>	<b>ML</b>
<b>Usage Range</b>	76.2%	86%
<b>Usage Frequency</b>	A few times yearly (39%)	Every week (36.6%)
<b>Top Usage Contexts</b>	Entertainment Visualise Products Education	Convenience Education Job

The above data shows that Machine Learning is much more popular amongst the general public, with 86% of respondents saying they have used Machine Learning before. A reference list was provided to respondents to aid them in identifying the technologies if they were unaware. Moreover, the usage rate for Augmented Reality was 76.2%, which is 11.4%, down from the Machine Learning rate.

Moreover, it was found that 37% of the global population gains access to AR technologies via Facebook and Snapchat. Therefore, integration through widely adopted platforms can help boost the popularity of AR and ML technologies and increase their influence over a wider audience. Individuals also use Machine Learning more often at a weekly frequency compared to the few times yearly frequency that is present for AR technologies.

- **Claim 1:** The strong acceptance rate and usage frequency of ML technologies suggest that combining them with AR could boost AR usage and awareness, given the public’s favourable behaviour towards ML.
- **Claim 2:** Integration of Augmented Reality through popular online platforms can help increase the level of awareness that individuals have for AR.

Next, it was found that Entertainment was the top application sector for Augmented Reality, and Convenience was the top sector for Machine Learning. Both technologies had “Education” in their top application sectors.

- **Claim 3:** Combining AR and ML for education can result in a more effective and accepted educational tool, given educational learning’s popularity in both sectors.

### Perceptions

The average experience score from respondents for Augmented Reality was 6.36, and for Machine Learning was 7.42. Moreover, 96.8% of respondents thought that Machine Learning could improve Augmented Reality. The respondents identified the top 3 ways in which they feel ML can improve AR as users of both products and possible users of future integrated products. Note that the responses “Efficiency” and “Computer Vision” had a common modal percentage of 15.3%.

Rank	Percentage	Method
1	35.7%	Personalised responses
2	30.6%	Convenience
3	15.3%	Efficiency and Computer Vision

- **Claim 4:** Users feel that specific, personalised responses through Machine Learning can improve AR experiences.

### Application Contexts

Respondents identified the major application sectors in which they feel this integration would be beneficial. Moreover, 92.9% of users stated they would use AR more when ML is integrated into it. Hence, this further justifies Claim 1 with a strong degree of affirmation from real-life users.

Rank	Percentage	Sector
1	39.8%	Education
2	37.8%	Healthcare
3	20.4%	Travel

This data supports Claim 2. We also find that Healthcare is another area where users feel that ML can improve their AR experience. Therefore, the following data has supported the first two claims.

### Challenges

It was found that 55.6% of respondents were concerned with Machine Learning or Augmented Reality, and a supportive 67.7% felt that the technologies were widely accessible to them. The accessibility percentage reflects the following claim.

- **Claim 5:** The commercial support of ML and AR technologies is yet to reach new levels of mainstream universal accessibility to all demographics.

Moreover, an important aspect of any technological integration is understanding the challenges and concerns of both technologies. As a majority of correspondents found that there is a need for concern, it is, hence, a matter of importance to understand and act accordingly. As this data is from a societal perspective, it does not consider computational challenges with algorithms, datasets or efficiency. Hence, social factors will be considered.

Augmented Reality Challenges - A CGS survey done with 100 technology executives in the context of Augmented Reality found that the implementation cost is the greatest challenge currently faced. Custom application technology costs reach over six figures. Popular conglomerate IKEA was estimated to pay between \$30,000 to \$60,000 for their AR-based product visualisation application. These constraints support the above Claim 4 that mainstream accessibility has still not been achieved.

The requirement of specialised human skills, time, resources and software leads to this high implementation cost. Moreover, employment salaries are dependent on geographic factors.



Another major factor that affects AR implementation is staff and user training. An increase in awareness through appropriate forms of user documentation and training approaches can help individuals increase familiarity with AR technologies and how to work with them.

- **Claim 6:** Software, specialised skills, training, resources and time requirements have made AR implementation expensive. Integrated solutions with other technologies are, hence, only feasible to large organisations with the available capital.

Machine Learning Challenges - Next, Machine learning poses various concerns. While machine learning can help individuals through a variety of approaches and contexts, there are a variety of concerns as well. A 2019 research paper found that the increasing presence of Machine Learning architectures within our society is leading to ethical legislation developments across the world to ensure the correct implementation of the technology. The paper also states concerns with privacy and social safety.

Common Challenges - This claims that the social safety, privacy and ethical concerns that exist currently for basic forms of Machine Learning must also be considered from the lens of societal impacts of Augmented Reality to understand the general concerns and specific concerns - both of which will be relevant. Firstly, there are health concerns over the extended usage of head-mounted displays, which can lead to physical and visual damage to the user. The sensory damages caused by extensive usage pose a requirement for time-bound usage restriction software to be embedded within the AR service to possibly combat this issue.

Moreover, there are mental health concerns associated with immersive experiences that occur over extended periods. This can lead to anxiety and stress. Lastly, there are major social concerns regarding the changes in social behaviour associated with the reduction in face-to-face communication as a result of increased adoption of immersive experiences that are provided by AR and VR technologies. These changes can lead to social isolation.

- **Claim 6:** By applying Machine Learning algorithms in AR, there are concerns in the fields of social safety, privacy, ethical usage, mental health, social interactions and general physical health damage. These issues require legislative action and usage control approaches.

### **User Feedback**

Next, the respondents gave their user feedback on products with an average importance rating of 8.04/10. They also rated their ability to give feedback on Augmented Reality or Machine Learning an average score of 6.51/10. User feedback is a direct manner of understanding problems faced by users and, therefore, improving products to cater more effectively towards the end user.

As individuals feel less confident while giving feedback on AR and ML technologies, this statistic re-affirms Claim 4. If mainstream adoption of these technologies is taken to the next

level, individuals will be able to better understand the technologies and, therefore, give more confident feedback. User confidence in the technologies can also be created through online tutorials and documentation on the usage of AR and ML technologies.

### **Case A: Historical Understanding**

Global history is a critical aspect of our society. The survey results first found that respondents rated the use of ML in education a score of 8.16/10. This supports previous claims and explorations made in this research paper. The critical data collected was that users gave a score of 7.71 when asked about how useful Machine Learning will be when providing insights about historical locations.

Hence, the first aspect of the application will consider integrating ML and AR to improve historical understanding. The importance of the technologies in the field of education was already supported by previously stated data. In context with historical exploration, there are a variety of ways both technologies can have useful integrations.

### **Functionality**

Deep Learning Neural Networks can be applied in contexts that require more than 3 Hidden Layers for processing inputted data. Deep Learning approaches can provide efficient ML pattern recognition on large, complex amounts of data. Next, CNNs (Convolutional Neural Networks) are a type of Neural Network that works on identifying and classifying visual media. The architecture of a CNN algorithm comprises various aspects.

Firstly, training and validation images are required to create a dataset which will be used. Next, CNN uses Local Binary Pattern and Mean Standard Deviation techniques to extract features within an image. The structure for a CNN architecture consists of Convolution and Pooling layers. Convolutional layers are layers in which multiple kernels (or filters) extract specific features from an image and create an activation in the form of a feature map.

These feature maps or “channels” apply a filter to summarise the features found within the given region of the image. This architecture can allow for image recognition. Machine learning can hence identify historical landmarks. Integrating these valuable insights can help individuals explore historical areas without the need for typical information sources, which are resource-extensive and expensive. Users will hence have a convenient benefit as a result of this integration.

### **Implementation**

AR technologies use Simultaneous Localization And Mapping (SLAM) to map the external environment through different frames. Various sensors can also improve the efficiency of these algorithms. Through connections to the AR architecture and ML architecture, the architectures can work in unison. For example, a user can ask for information about a historical building. Machine learning algorithms can extract insights to understand the requirement given

by the user and pass the requirement to the AR architecture, which will overlay specific information where required.

The main aspect of consideration is the processing time associated with both architectures working simultaneously. This would require research and development to work towards faster computational speeds to give users quicker results and, hence, an improved experience. Given the importance of historical understanding through education, this implementation can give individuals of all demographics highly meaningful information.

- **Claim 7:** Through CNN algorithms and object detection of SLAM Augmented Reality, the increased insights on historical understanding can majorly push societal appreciation for global history.

### **Case B: Consumer Interaction**

The collected data found that users gave Machine Learning a 6.73 average score in the context of shopping and an average score of 6.99 in the context of AR games. The first data point shows that users do not feel too confident about using Machine Learning when they are shopping. This could be to reduce the applications of technology when unnecessary and to support individual decision-making.

### **ML Application in Gaming**

Moreover, through Machine Learning, users can get insights into their performance within the game environment. Moreover, machine learning can act as a companion player alongside real-life players. Machine Learning algorithms embedded within NPCs have improved the experience of games. A research work found that NPCs are critical in shaping the experience of game players. Moreover, realistic conversations can also be improved through Machine Learning approaches, which can identify user linguistic patterns.

### **AR Application In Gaming**

Next, users felt more confident about using Machine Learning for AR games. Recently, AR games have become increasingly popular. The blockbuster PokemonGO game is one of the most popular forms of AR application, with over 81 million players in May 2024. The wide accessibility of this game is an example of a popular, accessible form of AR. Early forms of AR games used online street datasets along with the user's GPS to map the player and correspondingly create experiences based on their location. Moreover, Marker-based AR approaches through QR codes can be used to overlay game elements in the real world.

Current limitations in headset architecture and smartphone capabilities limit the implementation of gamified elements in an AR environment. New headsets are showing promising signs in terms of opportunities. Further development can increase the adoption of AR games and hence increase their popularity amongst generations.

## **ML Application In Commerce**

Users were not as confident in the commercial application of machine learning (6.73/10) score. However, there are still various benefits of Machine Learning which require attention from the general public through application-based implementations. Firstly, ML technologies can provide predictions of a consumer's behaviour towards certain products, which can help recommend and not recommend products based on their previous actions towards products.

This can benefit both the consumer of the good/service and the seller. Moreover, machine learning through Chatbots that implement Natural Language Processing (NLP) from a Large Language Model (LLM) can use Neural Networks to communicate with users. Chatbots, hence, are a great implementation for all demographics engaging in commercial activity. The use of different algorithms, such as Transformer Neural Networks (TNN) or Recurrent Neural Networks (RNN), poses individual advantages and disadvantages to users who interact with natural language processing technologies.

Firstly, Recurrent Neural Networks benefit from their memory capabilities and sequential handling. However, their high training time and vanishing gradient as a result of backpropagation through time (BPTT) can cause incorrect calculations and updates of the network's weights and biases, therefore affecting the accuracy of learning. Long short-term memory RNNs (LSTM) can be employed to solve this issue through a 3-gate memory cell architecture. Next, Transformer Neural Networks benefit from high performance and parallelisation that is enabled through their Self-attention mechanism. However, these networks suffer from high costs and their inability to interpret common sense language. Such applications of Natural Language Processing are a common example of Machine Learning within the commerce industry and are seen via various real-world examples.

## **AR Application In Commerce**

AR in shopping has seen an immense rise in popularity amongst younger generations. Companies such as Nike have used AR features to specifically cater towards younger generations. Augmented Reality technologies have been able to use product visualisation features to allow consumers to try on products before they buy them. These have applications in various businesses, such as fashion and furniture. With GenZ and Millennial generations accounting for 62% of e-commerce users, the implementation of AR for e-commerce would increase the influence of AR amongst younger generations as well.

A popular example of an AR Application is the IKEA Studio feature on Snapchat. This application combines both gaming and commerce as users try to declutter a room that is displayed in their visual environment through the help of IKEA Products. Such examples show the growing innovation in the application of Augmented Reality.

## **Implementation**

When combined with Augmented Reality, the overall physical and digital world can integrate the game aspects within the physical world.

- **Claim 8:** Through Machine Learning via realistic NPCs and in the form of performance insights, video game experiences through AR (a popular mainstream application of the technology) can improve user engagement and experience.
- **Claim 9:** Augmented Reality is currently a leading implementation within the e-commerce industry to attract younger audiences. Hence, by integrating product recommendations via Machine Learning, businesses can improve their user experience while attracting new customers.

### **Case C: Media and Language**

The survey found that respondents gave a 7.06/10 score for implementing Machine Learning within immersive AR experiences. Therefore, general users feel that immersive media through AR can be improved with Machine Learning. Moreover, respondents gave an average score of 6.72/10 for the implementation of language translation algorithms with machine learning through AR technologies.

### **Language Translation**

Augmented Reality translators are currently available for consumers to download and use. They incorporate Machine Learning algorithms to identify text and convert it into a specific language. “AR Translator - Translate Photo” by BrainMount LTD is an example of such an application. Moreover, through machine translation approaches, such technologies can output translated versions of imputed text to the user. Machine translation has a variety of approaches, such as RBMT (Rule-Based Machine Translation), where existing linguistic rules and dictionaries are used as the basis for generating textual translations. Other methods include SBMT (Syntax Based Machine Translation) and Neural Machine Translation (NMT). NMT incorporates deep learning architectures that can use transformer models to execute machine translation.

- **Claim 10:** Through AR services, Machine Translation services can provide users with increased convenience.

### **Immersive Media**

Immersive media, in general, focuses more on increasing the user’s interactivity with the AR technology. As found before, through this survey, a majority of the respondents felt that adding Machine Learning would increase their user interaction with AR technologies. Therefore, ML can help support the main aims of an immersive media experience. Despite the increase in adoption of Machine Learning technologies by companies worldwide, their applications within the media industry have remained limited. There is still a need to develop more solutions within the media industry through AR technologies.

AR/VR technologies such as the Apple Vision Pro can allow users to watch visual media content through their Head Mounted Displays, therefore creating an immersive environment. The applications of Machine Learning through visual media content such as television and movies are seen through recommendation algorithms that exist amongst entertainment streaming platforms.

- **Claim 11:** Immersive media for entertainment can also provide user recommendations through recommendation algorithms that are supported by ML architectures.

#### **Case D: Healthcare**

Respondents gave the use of Machine Learning within healthcare a score of 7.33/10. Firstly, Augmented Reality is majorly changing the way healthcare services are provided to patients. Using Head Mounted displays, surgeons can view patient vitals during operations without having to look away to check them. Such usages of Augmented Reality show a positive benefit towards efficiency and ease of conduct amongst such areas of application.

Given the critical aspect of time that is connected to health-related activities, the improved convenience has a higher level of importance when viewed in the context of healthcare applications. Through informatic overlays, there are major advantages towards doctors who work with patients during periods that require critical attention. Moreover, it was found that AR technologies can help individuals going through rehabilitation. Lastly, the interactive learning approaches provided through the use of Augmented Reality technologies increase medical awareness and training to increase medical personnel worldwide.

Lastly, as Machine Learning can provide meaningful insights through user interactions via chatbots, etc, the aspect of training and education in the field of Healthcare can dramatically be improved. Moreover, Machine Learning technologies can also be used to identify trends amongst data regarding patients or diseases, which can help prompt action by the medical community. This can, therefore, help assist medical personnel in the research and development of new treatments.

Moreover, deep learning approaches can also be used to analyse visual media through the use of different architectures to identify abnormalities. Lastly, applications through super-resolution architectures can help increase the quality of images generated in the field of healthcare, which can also help provide more accurate diagnoses to patients.

- **Claim 12:** Machine Learning can help provide a more accurate diagnosis, provide insights into patient data trends and help improve convenience through the usage of Augmented Reality headsets.

#### **Limitations Of The Study**

This investigation has not considered specific challenges in areas of application. The challenges identified are specific to the technologies and not to certain sectors. As sector-specific

concerns are critical to the implementation of AR and ML-integrated technologies, they must be considered as well. The study does not look at the usage and implementation from a development point of view to understand how different countries can implement or are currently implementing either of the technologies. To achieve widespread accessibility, the same level of awareness must be achieved generally amongst all demographics around the world. This would ensure that the integrated solutions would have the greatest impact amongst global communities and in their specific sectors.

## **Conclusion**

### **Summary Of Findings**

The findings of this investigation found that Augmented Reality and Machine Learning are both growing areas of application in various sectors. Computer vision and personalisation are seen to be two major areas in which an integrated solution can majorly help in their respective fields. Moreover, this investigation found that an increase in the usage of Augmented Reality is still required.

Lastly, this investigation identified a few select areas of application that include Historical understanding, Consumer Interaction, Media and Language, and Healthcare. All major points found during these explorations are stated below in the form of 12 claims. Moreover, the 101-respondent survey helped support the claims and explorations made to better understand the integration of AR and ML in select industries.

### **Future Research Suggestions**

Future research is the foundation of development for new technologies. Future research could happen to understand the usage of these integrated technologies for NGOs and social care sectors. Similarly, research can be done on the application of AR and ML integrations for assistive technologies and in combating global issues such as global warming. Therefore, future research on the application of the technologies in global areas of concern could help us understand the implications in critical aspects apart from general employment or convenience-based applications sectors that are explored through this paper.

Future research could be conducted on how to combat the issues that arise from both technologies. While this has been explored subtly in this paper, in-depth research would help individuals and governments take the required steps to ensure a healthy integration of these technologies within our society.

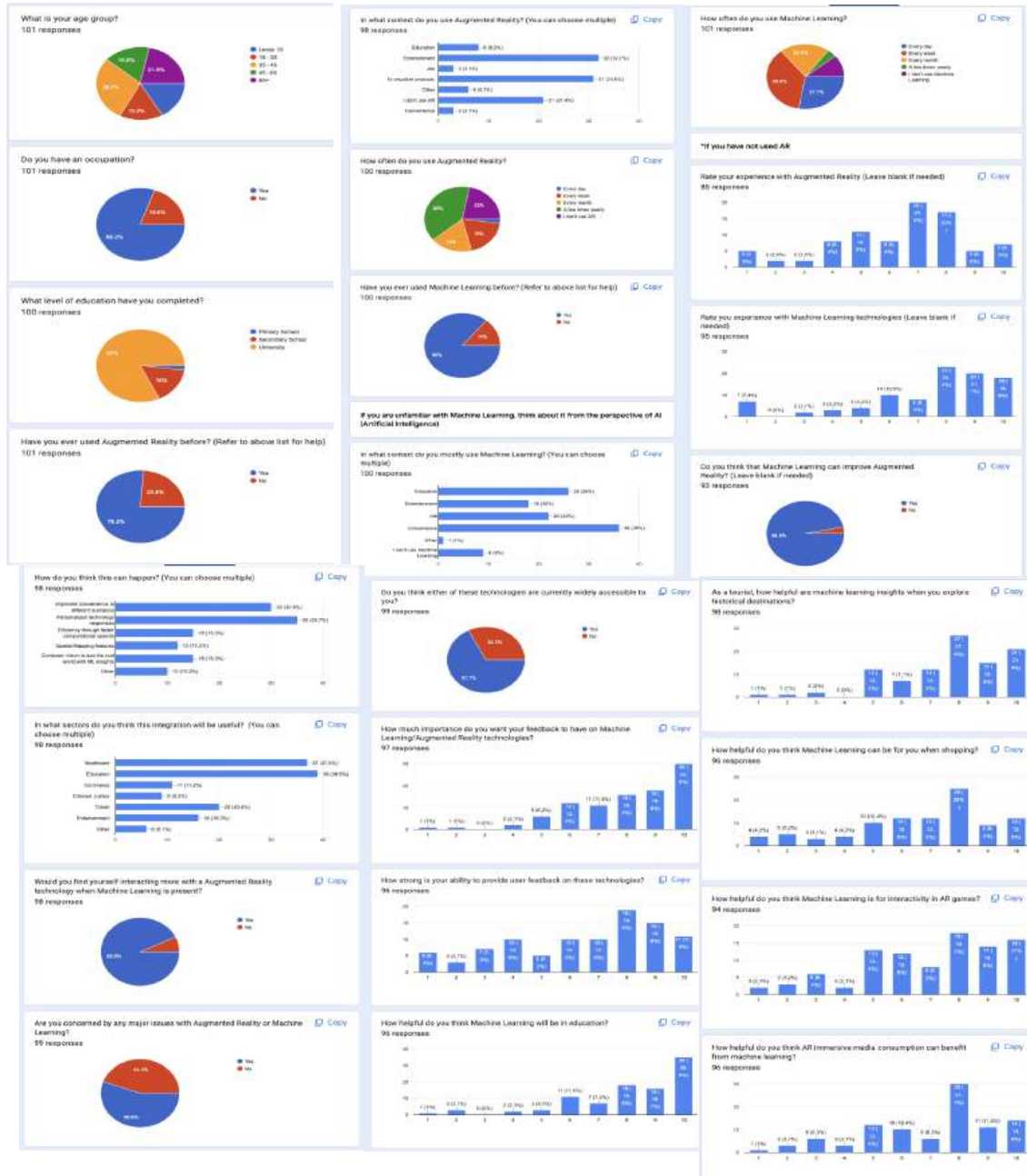
### **Final Thoughts**

Machine Learning is one of the most important technologies of the modern era. It is now revolutionising various industries and providing features that are changing the way we approach our jobs and our daily lives. This immense change can give way to various issues that could affect us negatively. Moreover, Augmented Reality is also a growing sector. The healthcare

concerns associated with the extensive use prompt that in the current environment, continued use of AR technologies would not be advisable to individuals.

Therefore, research and development are required to help build safer AR models through Head Mounted Displays to ensure user safety. Once this is achieved, combined with the integration of Machine Learning, Augmented Reality could become an incredibly powerful tool for all individuals that could change the way we use technology in our world.

### Survey Proof (24/27 Questions Shown)





## Works Cited

- Adamska, Iwona. "Navigating the Risks: The Hidden Dangers of Augmented Reality." Navigating the Risks: The Hidden Dangers of Augmented Reality, Nsflow, 29 Sept. 2023, nsflow.com/blog/navigating-the-risks-the-hidden-dangers-of-augmented-reality.
- Anthony, James. "98 Vital Augmented Reality Statistics: 2024 Market Share & Data Analysis." Financesonline.Com, FinancesOnline.com, 7 Aug. 2024, financesonline.com/augmented-reality-statistics/.
- "Ar and Ai: The Role of AI in Augmented Reality." GeeksforGeeks, GeeksforGeeks, 26 Dec. 2023, www.geeksforgeeks.org/ar-and-ai-the-role-of-ai-in-augmented-reality/.
- Bacciu, Davide, et al. Societal Issues in Machine Learning, www.esann.org/sites/default/files/proceedings/legacy/es2019-6.pdf. Accessed 16 Aug. 2024.
- Barakat, Basel, et al. "Integrating Machine Learning with Augmented Reality for Accessible Assistive Technologies: Universal Access in Human-Computer Interaction. User and Context Diversity." Guide Proceedings, 26 June 2022, dl.acm.org/doi/abs/10.1007/978-3-031-05039-8\_12.
- Baviskar, Lina, et al. HERITAGE IDENTIFICATION OF MONUMENTS USING DEEP LEARNING TECHNIQUES, 5 May 2024, www.irjmets.com/uploadedfiles/paper//issue\_5\_may\_2024/55481/final/fin\_irjmets1715786370.pdf.
- Danikovich, Darya. "Machine Learning in Retail." EffectiveSoft, 22 Mar. 2024, www.effectivesoft.com/blog/machine-learning-in-retail.html.
- Excel, Fore. "Is Augmented Reality Very Expensive?" LinkedIn, 24 Jan. 2024, www.linkedin.com/pulse/augmented-reality-very-expensive-fore-excel-0nycf#:~:text=Creating%20high%2Dquality%20AR%20content,on%20the%20industry%20and%20application.
- Hayes, Molly, and Amanda Downie. "What Is Augmented Reality?" IBM, 3 Apr. 2024, www.ibm.com/topics/augmented-reality.
- Hovsepyan, Tigran, et al. "Machine Learning in Gaming." Plat.AI, 18 Aug. 2022, plat.ai/blog/machine-learning-in-gaming/.
- Krawanski, Filip. "How Many People Play Pokemon Go? Pokemon Go Player Count." Dexerto, 4 May 2024, www.dexerto.com/pokemon/how-many-people-play-pokemon-go-pokemon-go-player-count-2132719/.
- Kulp, Patrick. "IKEA Created an Augmented Reality 'Escape Room' on Snapchat." Adweek, 26 Feb. 2021, www.adweek.com/commerce/ikea-augmented-reality-snapchat-ogilvy-buzzfeed/.
- McCluskey, Megan. "How Augmented Reality Shopping Could Go Mainstream." Time, Time, 28 Jan. 2022, time.com/6138147/augmented-reality-shopping/.
- Mekni, Mehdi, and André Lemieux. Augmented Reality: Applications, Challenges and Future ..., www.cs.ucf.edu/courses/cap6121/spr2020/readings/Mekni2014.pdf. Accessed 16 Aug. 2024.
- Mlynarek, Wendy. Deep Learning and Augmented Reality - Dassault Systèmes Blog, 8 June 2024, blog.3ds.com/brands/delmia/combining-deep-learning-algorithms-with-augmented-reality/.

- Parwal, Rahul. "How to Enhance Software Quality with User Feedback." How To Enhance Software Quality With User Feedback, 12 Aug. 2024, [muuktest.com/blog/software-quality-with-user-feedback#:~:text=User%20feedback%20provides%20a%20direct,marketing%20tool%20for%20your%20product.](https://muuktest.com/blog/software-quality-with-user-feedback#:~:text=User%20feedback%20provides%20a%20direct,marketing%20tool%20for%20your%20product.)
- Pelchen, Lexie, and Samantha Allen. "Internet Usage Statistics in 2024." Forbes, Forbes Magazine, 1 Mar. 2024, [www.forbes.com/home-improvement/internet/internet-statistics/#:~:text=On%20average%2C%20internet%20users%20spend,to%20the%20Global%20Web%20Index.](https://www.forbes.com/home-improvement/internet/internet-statistics/#:~:text=On%20average%2C%20internet%20users%20spend,to%20the%20Global%20Web%20Index.)
- Public Views of Machine Learning - Ipsos Mori, Apr. 2017, [royalsociety.org/-/media/policy/projects/machine-learning/publications/public-views-of-machine-learning-ipsos-mori.pdf](https://royalsociety.org/-/media/policy/projects/machine-learning/publications/public-views-of-machine-learning-ipsos-mori.pdf).
- Rothmann, Daniel. "Enhancing AR with Machine Learning." Medium, Towards Data Science, 4 Nov. 2019, [towardsdatascience.com/enhancing-ar-with-machine-learning-9214d2da75a6](https://towardsdatascience.com/enhancing-ar-with-machine-learning-9214d2da75a6).
- Saidin, Nor Farhah, et al. (PDF) a Review of Research on Augmented Reality in Education: Advantages and Applications, June 2015, [www.researchgate.net/publication/281336331\\_A\\_Review\\_of\\_Research\\_on\\_Augmented\\_Reality\\_in\\_Education\\_Advantages\\_and\\_Applications](https://www.researchgate.net/publication/281336331_A_Review_of_Research_on_Augmented_Reality_in_Education_Advantages_and_Applications).
- Sarker, Iqbal H. "Machine Learning: Algorithms, Real-World Applications and Research Directions - SN Computer Science." SpringerLink, Springer Singapore, 22 Mar. 2021, [link.springer.com/article/10.1007/s42979-021-00592-x](https://link.springer.com/article/10.1007/s42979-021-00592-x).
- Tabsharani, Fred. "What Is Machine Translation?: Definition from TechTarget." Enterprise AI, TechTarget, 29 July 2024, [www.techtarget.com/searchenterpriseai/definition/machine-translation](https://www.techtarget.com/searchenterpriseai/definition/machine-translation).
- "Top 6 Biggest Challenges to Implementing AR Technology." CGS, 2 June 2023, [www.cgsinc.com/blog/top-6-biggest-challenges-implementing-ar-technology-2021](https://www.cgsinc.com/blog/top-6-biggest-challenges-implementing-ar-technology-2021).
- "Types of Ar." Digital Promise, 13 May 2022, [digitalpromise.org/initiative/360-story-lab/360-production-guide/investigate/augmented-reality/getting-started-with-ar/types-of-ar/#:~:text=Marker%2Dbased%20AR%20requires%20a,exist%20in%20the%20real%20world.](https://digitalpromise.org/initiative/360-story-lab/360-production-guide/investigate/augmented-reality/getting-started-with-ar/types-of-ar/#:~:text=Marker%2Dbased%20AR%20requires%20a,exist%20in%20the%20real%20world.)
- "Watch Movies and TV in an Environment on Apple Vision Pro." Apple Support, [support.apple.com/en-in/guide/apple-vision-pro/tan7241583f5/visionos](https://support.apple.com/en-in/guide/apple-vision-pro/tan7241583f5/visionos). Accessed 16 Aug. 2024.
- Wright, Gavin. "What Is Augmented Reality Gaming (Ar Gaming)? – TechTarget Definition." WhatIs, TechTarget, 6 Feb. 2023, [www.techtarget.com/whatis/definition/augmented-reality-gaming-AR-gaming#:~:text=Augmented%20reality%20gaming%20\(AR%20gaming\)%20is%20the%20real%2Dtime,global%20positioning%20system%20\(GPS\).](https://www.techtarget.com/whatis/definition/augmented-reality-gaming-AR-gaming#:~:text=Augmented%20reality%20gaming%20(AR%20gaming)%20is%20the%20real%2Dtime,global%20positioning%20system%20(GPS).)
- Yin, Chenxi. "NPCs in Video Games: A Reflective Resource for Sports Coaches and Participant Engagement." Frontiers, Frontiers, 14 June 2024, [www.frontiersin.org/journals/sports-and-active-living/articles/10.3389/fspor.2024.1403829/full](https://www.frontiersin.org/journals/sports-and-active-living/articles/10.3389/fspor.2024.1403829/full).
- Zanni, Lorenzo, and Stan Moore. "Why Immersive Tech Need Artificial Intelligence." IABM, 12 Jan. 2018, [theiabm.org/immersive-tech-need-artificial-intelligence/](https://theiabm.org/immersive-tech-need-artificial-intelligence/).

# The Statistical Link Between ADD/ADHD and Substance Use, and Opportunities for Preventing Adolescent Overdose Fatalities within This Population

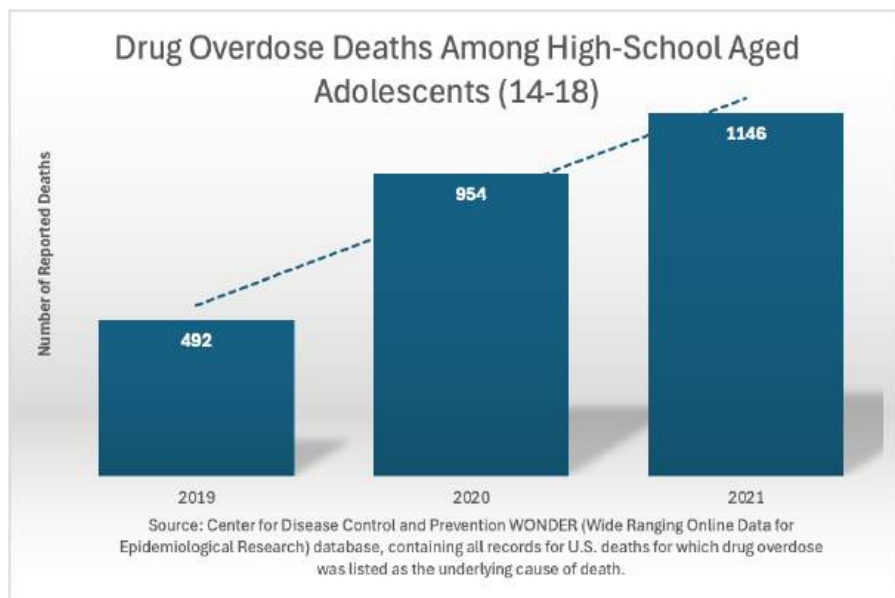
By Liv MacMillan

## Abstract

Identifying and treating adolescents with ADD/ADHD can reduce this population's risk of substance use and unintentional overdose. Increased awareness of this public health opportunity is critical to reducing adolescent overdose fatalities. Effective public health education efforts are needed to adequately inform mental health providers, pediatric care givers, school-based counselors, and parents of children with ADD/ADHD of these findings and their implications. Access to the mental health care services required to diagnose and treat ADHD/ADD varies significantly by geographic region, ethnicity, language spoken at home and access to health insurance. These disparities should be addressed through targeted public health education and increased access to care. Increasing access to treatment for adolescent ADD/ADHD should be seen as a critical component of public health policies and programs to reduce adolescent overdose fatalities.

## Introduction/Context

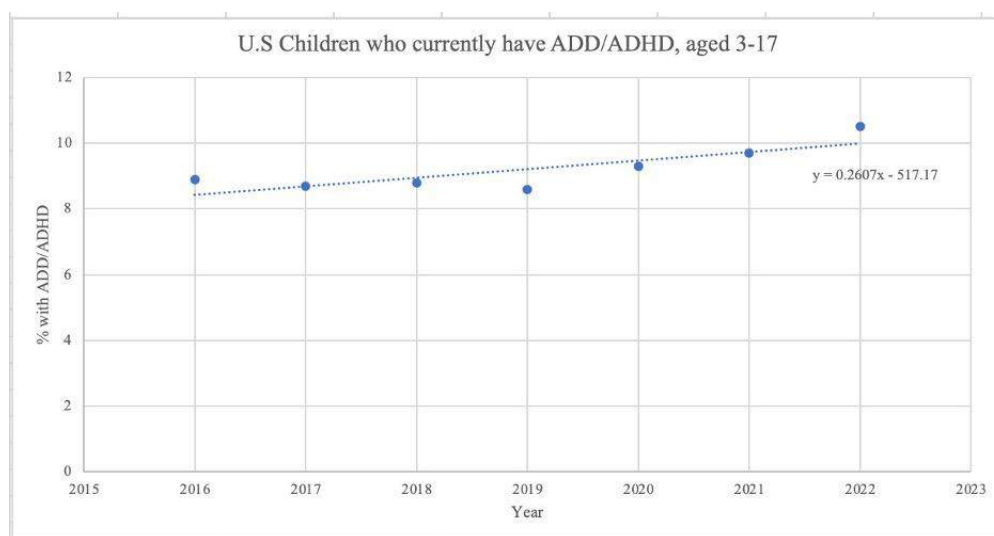
Drug overdose fatalities in the United States have become a national epidemic and a public health crisis. Drug overdoses have killed more than 1 million people in the United States since 1999 (Hoban and Kan). The majority of these fatalities have been driven by opioids, including the synthetic painkiller fentanyl (Hoban and Kan). Fentanyl is now the leading cause of death for Americans ages 18-49 (Jung and Leonard). Adolescent overdose fatalities have increased dramatically in recent years. Among adolescents ages 14-18, overdose deaths increased 94% from 2019 to 2020 and 20% from 2020 to 2021 (Danielson et al).



Within this context it is critical that we identify and implement effective strategies for reducing adolescent overdose deaths. Public health strategies and interventions that target specific at-risk populations can be an efficient and effective approach to reducing adolescent drug use. One such at-risk population is adolescents with ADD/ADHD.

### **Rates of Adolescent ADD/ADHD Diagnosis and Treatment**

Concurrent to the national overdose crisis, ADD/ADHD diagnoses in adolescents and children have been increasing. As of 2022, approximately 1 in 9 U.S. children ages 3-17 have ever received an ADHD diagnosis (11.4%, 7.1 million children) and 10.5% (6.5 million) had current ADHD (Danielson, Claussen et al). Approximately 1 million more children had received an ADHD diagnosis in 2022 than in 2016. (Danielson, Claussen et al).



Source: Danielson, Melissa L., Angelika H. Claussen, et al. “ADHD prevalence among U.S. children and adolescents in 2022: Diagnosis, severity, co-occurring disorders, and treatment.” *Journal of Clinical Child & Adolescent Psychology*, vol. 53, no. 3, 3 May 2024, pp. 343–360, <https://doi.org/10.1080/15374416.2024.2335625>.

### **The Link between ADD/ADHD and Substance Use**

ADD/ADHD is a neurodevelopmental disorder characterized by symptoms of inattention, hyperactivity and/or impulsivity. Research indicates that individuals with ADHD have a higher risk of developing substance use disorders (SUD). One meta-analysis reported that up to 30% of individuals with ADHD develop SUDs, a rate significantly higher than the 10-15% observed in the general population (Wilens). One factor that contributes to this increased risk is the impulsivity associated with ADHD. Impulsive behavior can increase risk-taking, including substance use (Biederman, Wilens, et al). In the age of fentanyl tainted drugs, this impulsivity and risk taking can be particularly dangerous. Adolescents who do not have substance use disorders but are instead experimenting with substances can be exposed to fatal doses at first

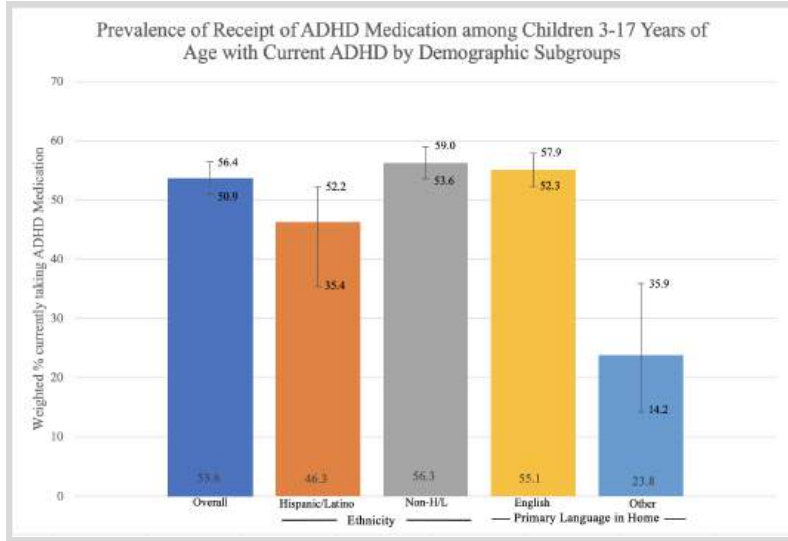
onset of substance use. Additionally, the comorbidity of ADHD with other psychiatric disorders, such as depression and anxiety, can further elevate the risk of substance abuse (Biederman, Wilens, et al). Also notable is that studies have found that some people with ADHD use substances as a form of self-medication to alleviate ADHD symptoms (McLoughlin, Murray, and Asherson).

### **Treatment of ADD/ADHD Can Reduce the Risk of Substance Use**

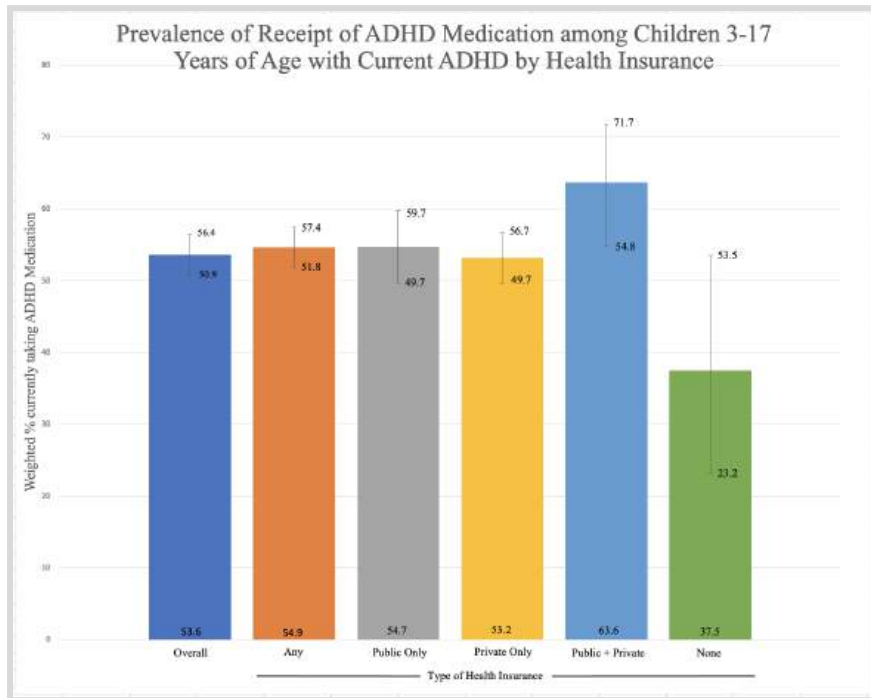
Pharmacological treatments for ADHD such as stimulant medications have been shown to reduce substance use risk (Hahn and Mannuzza). A meta-analytic review of long-term studies in which pharmacologically treated and untreated youths with ADHD were examined for later SUD outcomes found that the pooled estimate of the odds ratio indicated a 1.9 fold reduction in risk for SUD in youths that were treated with stimulants compared with youths who did not receive pharmacotherapy for ADHD (Wilens et al). Behavioral interventions and cognitive behavioral therapy (CBT) have also been found effective in addressing both ADHD symptoms and substance use. These approaches can help individuals strengthen executive function skills and reduce impulsivity which may, in turn, reduce the likelihood of substance use (Safren, Sprich and Fava).

### **Access to Treatment**

Many children and adolescents diagnosed with ADD/ADHD do not receive treatment. In 2022 nearly 2 million children with ADHD, or approximately 30%, did not receive ADHD specific treatment (Danielson et al). Additionally, rates of ADD/ADHD diagnosis and subsequent treatment vary by region. ADHD diagnosis estimates among children ages 3-17 vary from 6% to 16% across states (Danielson, Holbrook, et al). According to some studies, estimates of the number of these children that are receiving ADHD medication treatment vary from 38% to 81% and those receiving behavior therapy vary from 39% to 62% (Danielson, Holbrook, et al). The 2022 National Survey of Children's Health also found significant variation in rates of ADHD medication treatment by ethnicity, language spoken at home, and health insurance enrollment.



Source: National Survey of Children’s Health, 2016-2022, Child and Family Health Measures, Prevalence of current ADHD diagnosis among children aged 3-17 years, childhealth.org; Results shown with 95% confidence interval.



Source: National Survey of Children’s Health, 2016-2022, Child and Family Health Measures, Prevalence of current ADHD diagnosis among children aged 3-17 years, childhealth.org; Results shown with 95% confidence interval.

**Conclusions/Recommendations:**

Increasing rates of adolescent drug overdoses underscore the urgency for effective public health responses. Parents and adolescents need to be educated regarding the link between ADD/ADHD and substance use and the opportunity to reduce the probabilities of substance use through ADD/ADHD treatment. Increasing access to and rates of treatment for ADD/ADHD in adolescents could serve as an important strategy in the prevention of substance misuse and overdose in this vulnerable population. Clinical training as well as novel public health education and communication strategies need to be developed to more effectively reach this population.

## Works Cited

- Biederman, J., Wilens, T. E., Mick, E., Faraone, S. V., & Spencer, T. J. (2010). Current concepts in the treatment of ADHD and its comorbidities. *Journal of Clinical Psychiatry*, 71(9), 1163-1171.
- Dalsgaard S;Mortensen PB;Frydenberg M;Thomsen. “ADHD, Stimulant Treatment in Childhood and Subsequent Substance Abuse in Adulthood - a Naturalistic Long-Term Follow-up Study.” *Addictive Behaviors*, U.S. National Library of Medicine, pubmed.ncbi.nlm.nih.gov/24090624/. Accessed 25 Sept. 2024.
- Danielson, Melissa L., Angelika H. Claussen, et al. “ADHD prevalence among U.S. children and adolescents in 2022: Diagnosis, severity, co-occurring disorders, and treatment.” *Journal of Clinical Child & Adolescent Psychology*, vol. 53, no. 3, 3 May 2024, pp. 343–360, <https://doi.org/10.1080/15374416.2024.2335625>.
- Danielson, Melissa L., Joseph R. Holbrook, et al. “State-level estimates of the prevalence of parent-reported ADHD diagnosis and treatment among U.S. children and adolescents, 2016 to 2019.” *Journal of Attention Disorders*, vol. 26, no. 13, 22 May 2022, pp. 1685–1697, <https://doi.org/10.1177/10870547221099961>.
- Faraone, Stephen V., and Himanshu P. Upadhyaya. “The effect of stimulant treatment for ADHD on later substance abuse and the potential for medication misuse, abuse, and diversion.” *The Journal of Clinical Psychiatry*, vol. 68, no. 11, 15 Nov. 2007, <https://doi.org/10.4088/jcp.1107e28>.
- Friedman, J, et al. “Trends in Drug Overdose Deaths among US Adolescents, January 2010 to June 2021.” *JAMA*, U.S. National Library of Medicine, 2022, <pubmed.ncbi.nlm.nih.gov/35412573/>.
- Hahn, L. L., & Mannuzza, S. (2013). “Treatment of ADHD and the risk of substance use disorder.” *Journal of Clinical Psychiatry*, 74(7), 804-810.
- Hoban, Julie, and Courtney Kan. “Why Is Fentanyl so Dangerous?” *Washington Post*, 3 Nov. 2022, [www.washingtonpost.com/nation/2022/11/03/fentanyl-opioid-epidemic/](http://www.washingtonpost.com/nation/2022/11/03/fentanyl-opioid-epidemic/).
- Jung, Yoohyun, and Christian Leonard. “Drug Overdose Deaths: Tracking Devastating Toll of U.S. Crisis.” *The San Francisco Chronicle*, The San Francisco Chronicle, 5 Sept. 2024, [www.sfchronicle.com/projects/us-drug-overdose-deaths/](http://www.sfchronicle.com/projects/us-drug-overdose-deaths/).
- McLoughlin, G., Murray, R., & Asherson, P. (2009). “ADHD and substance misuse: a review.” *Journal of Attention Disorders*, 13(2), 149-157
- Molina, Brooke S., and William E. Pelham. “Childhood predictors of adolescent substance use in a longitudinal study of children with ADHD.” *Journal of Abnormal Psychology*, vol. 112, no. 3, Aug. 2003, pp. 497–507, <https://doi.org/10.1037/0021-843x.112.3.497>.
- “The National Survey of Children’s Health.” *National Survey of Children’s Health - Data Resource Center for Child and Adolescent Health*, [www.childhealthdata.org/learn-about-the-nsch/NSCH](http://www.childhealthdata.org/learn-about-the-nsch/NSCH). Accessed 25 Sept. 2024.
- Rohner, Henrik, et al. “Prevalence of attention deficit hyperactivity disorder (ADHD) among substance use disorder (SUD) populations: Meta-analysis.” *International Journal of*



- Environmental Research and Public Health*, vol. 20, no. 2, 10 Jan. 2023, p. 1275,  
<https://doi.org/10.3390/ijerph20021275>.
- Safren, S. A., Sprich, S., & Fava, M. (2005). "Cognitive-behavioral therapy for ADHD in medication-treated adults." *Journal of Clinical Psychiatry*, 66(2), 281-289.
- "U.S. Department of Health & Human Services - CDC Wonder: Mortality - Underlying Cause of Death." *Catalog*, Publisher Centers for Disease Control and Prevention, Department of Health & Human Services, 26 July 2023,  
[catalog.data.gov/dataset/cdc-wonder-mortality-underlying-cause-of-death](https://catalog.data.gov/dataset/cdc-wonder-mortality-underlying-cause-of-death).
- Wilens, Timothy E. "Attention deficit hyperactivity disorder and substance use disorders." *American Journal of Psychiatry*, vol. 163, no. 12, Dec. 2006, pp. 2059–2063,  
<https://doi.org/10.1176/ajp.2006.163.12.2059>.
- Wilens, Timothy E., et al. "Does stimulant therapy of attention-deficit/hyperactivity disorder beget later substance abuse? A meta-analytic review of the literature." *Pediatrics*, vol. 111, no. 1, 1 Jan. 2003, pp. 179–185, <https://doi.org/10.1542/peds.111.1.179>.
- Wilens, Timothy E. "The nature of the relationship between attention-deficit/hyperactivity disorder and substance use." *The Journal of clinical psychiatry* vol. 68 Suppl 11 (2007): 4-8.

# **Association Between Social Media and Forensic Psychoanalysis: Investigative Methodology**

## **By Vinícius Meireles Aun Carteiro**

### **Abstract**

In case studies, forensics has a fundamental role in identifying the source of the issue being investigated. This article delves into the psychological motives and factors that lead an individual to take a particular action. A wide range of evidence influencing the search can be obtained by investigating social media. Negative mental consequences generated in digital interactions, such as dissatisfaction, anxiety, and depression, lead to disastrous behavior that directly affects the hippocampus and prefrontal cortex of the brain. Social engineering techniques and their methods of operation are addressed, corresponding to current criminal schemes. Most of the results include that disorders condition the actions taken by a human being, and nowadays, forensics investigation must collect data associated with personality and social behavior through online information in addition to disseminating knowledge for the science.

**Keywords** investigation, mental consequences, social engineering, social media.

### **1. INTRODUCTION**

In contemporary neuroscience, the human brain is intricately divided into five distinct sections: the frontal lobe, parietal lobe, temporal lobe, occipital lobe, and insular cortex. Among these regions, the area primarily responsible for crucial functions such as decision-making, sociability, and the regulation of emotional control is the prefrontal cortex, which is situated within the frontal lobe. This highly specialized structure consists of a complex arrangement of neurons, glial cells, cortical layers, and extensive synaptic connections, all of which operate together to facilitate a wide range of cognitive processes and behaviors essential for everyday functioning. Understanding the intricacies of this brain region is vital for grasping how it influences our actions and interactions within various social contexts.

Despite the intricate and complex nature of neural connections within the brain, it is important to recognize that experiences significantly impacting feelings and emotions play a crucial role in shaping daily decision-making processes. Numerous studies have demonstrated this relationship, and the findings have remained unchallenged by the scientific community, at this time. For instance, Chapter 3 of "The Role of Affect in Decision Making" and the work, "Affect and Decision Making" provide compelling evidence illustrating how emotional experiences can sway our choices and judgments in everyday life. These studies highlight the undeniable influence that emotional states have on our cognitive functions and decision-making abilities.

Since more than 60% of the world's population uses social media, extracting significant information about a suspect is possible. While ordinary people adopt this type of profile, so-called "low profile" accounts are commonly used by individuals with malicious intentions.

They are characterized by not exposing themselves frequently on social media, leaving few details about their lives.

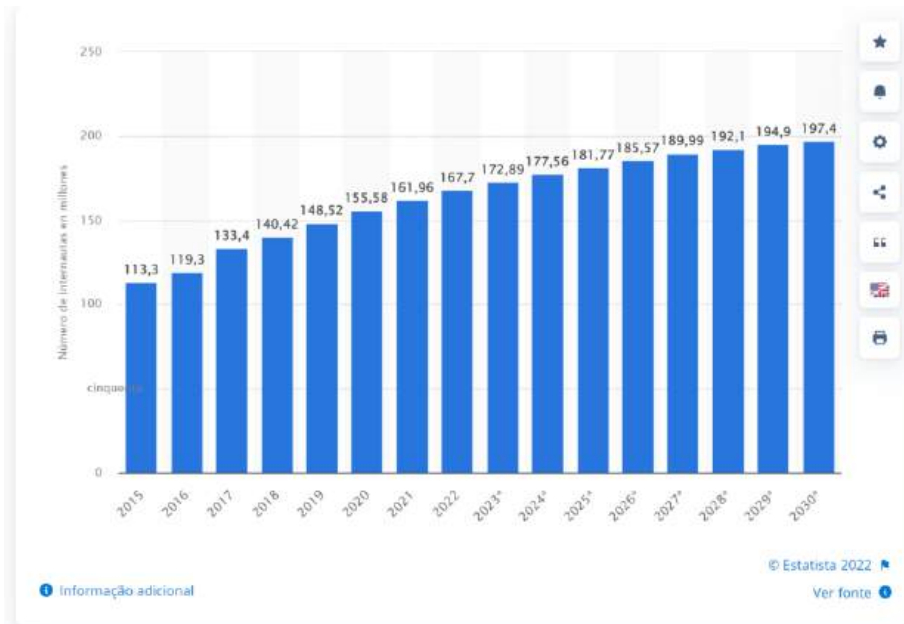


Fig 1:Graph of “alternative culture” on the number of internet users per year.

A large portion of the elderly population who use social media reports have been victims of some form of online scam. Among the most common are fake digital store products, which, through a highly scalable advertisement, lure individuals to click on the page and lead them to a malicious site that can steal their data.

Those who create such schemes are computer criminals disguised by a digital security system, often using a VPN ("Virtual Private Network"), which prevents scanning of the user's IP address. This issue is increasingly prevalent and harms the lives of law-abiding citizens. However, applying an investigative method that seeks a loophole in the lack of information is possible. This can be done through the profiles of people close to the subject who frequently post new content on their pages—this is where forensics comes into play. With this in mind, conducting a simple search or even assuming another identity is enough to achieve a greater good.

## 2. METHODOLOGY

Upon accessing the most popular social networks, one is met with a profile creation page, where one can add a name, photo, and description. With a profile created, a simple search in the platform is all needed. If not found, the second step is to search for relatives, friends, or related individuals. This increases the chances of a more accurate search.

It is worth noting that the following pages by a user are also important. If they are specific, the content discussed on these pages can be associated with part of the investigated person's personality, as the content consumed by the mind influences habitual actions.

The medium that may be the most important is the practice of Social Engineering. This area, which has been gaining significant traction, aims to obtain private data, whether financial

information or anything that may benefit the criminal. It is essential to understand the methodology they employ to combat it. In general, the main methods used by criminals involve software, enticing links, or phone scams.

The real problem is that this is part of an illicit practice called "Human Hacking," cybersecurity professionals can't prevent all these scams, because it involves a psychological game. In other words, it creates a connection between the criminal and the person on the other side, making the criminal act on the subconscious mind and often evoke emotions in the victim's mind.

Among the techniques used, persuasion and "Phishing" stand out. Classic persuasion is a communication technique that induces decision-making under the influence of another person, typically to benefit the influencer. Phishing, on the other hand, is a style of cyberattack that influences the victim to grant access to passwords, credit card numbers, or banking details. Criminals cast the bait, often through emails or SMS, which direct the person to a malicious website where they enter their private information.

Some forensic scientists perform extensive searches utilizing various statistical data points, including details such as a car's license plate or a residential address. While these pieces of information are crucial for enhancing the overall handling of a case and help to enrich the investigation and provide essential evidence in court, it is equally important to delve deeper into understanding the individual's identity, background, and behavioral nature. This deeper comprehension allows for a more nuanced analysis, which can ultimately lead to more effective investigative strategies and a clearer portrayal of the individual's motivations and actions.



Fig 2: Illustrative image from "AbcMed" on personality.

### 3. RESULTS

Dopaminergic stimuli caused by the seductive social engineering techniques activate various reward systems in the human brain, primarily the nucleus accumbens, which is

responsible for fundamental emotional perceptions in the forebrain. This structure plays a direct role in addiction-related issues and is capable of triggering neurotransmitter nerve impulses to demand new stimuli that can fulfill certain needs. This is directly related to purchasing a product or an online subscription, as the attractiveness chemically induces the brain to act.

After completing the fraud, the most common and recommended path of action is to alert the police and other law enforcement officials. Among them, two main groups are responsible for handling the case. Cybersecurity experts deal with the "hacking" aspect, searching for vulnerabilities in the suspect's system. Forensic investigators are responsible for both the psychological aspect of the incident and the physical and digital traces left by the criminal.

A comprehensive study is generated on three areas of investigation. The first encompasses as much information as possible from the internet about the suspect, from posts to followed pages. The second is a report of the victim, which is already done as standard practice, making it simple. The third focuses on the techniques and means used in the scam, such as websites, and emails. This way, cybersecurity professionals can stay updated on new methods and develop a strategy for future countermeasures.

As a result of the analysis, much evidence and an understanding of the investigated individual's personality are obtained. This provides a social dynamic about online interactions and behavioral understanding.

In addition to traditional methods of collecting digital evidence, the results indicate that the thorough analysis of online social interactions can uncover more nuanced behavioral patterns that might go undetected in standard investigative approaches. For instance, shifts in the frequency and tone of posts, recurrent thematic content, and even the deliberate choice of specific keywords can signal emotional fluctuations or the presence of malicious intent. These findings emphasize that by integrating advanced natural language processing techniques with continuous social media monitoring, investigators can predict and identify potentially risky behaviors with enhanced precision. Such a comprehensive approach allows for a faster and more adaptive response, especially in online fraud, where time-sensitive actions are essential in reducing the extent of harm and mitigating the overall impact.

It also allows for continuous monitoring of routine changes and the people with whom the individual interacts. This opens doors to a chance to predict the future of criminals and, creates a range of other involved groups. All this directly influences the legal decisions made in court and should be handled by forensic professionals or investigators.



Fig 3: Illustration of digital monitoring from “Personal Marketing Digital”.

#### 4. CONCLUSIONS

The world is in a state of constant transformation, and technology continues to advance at an increasingly rapid pace, populations around the globe are quick to adopt and integrate these new tools into their daily lives. These technological innovations reshape how people communicate, interact, and commit crimes. Consequently, security agents and law enforcement professionals must continually update their skills, knowledge, and tools to keep up with these changes and remain effective in their investigative practices. This need for constant adaptation is particularly pronounced in regions where technological access is limited or less developed. In such areas, the gap in investigative capabilities between these locations and those with more advanced technological infrastructures remains substantial, creating significant challenges in combating crime effectively. This disparity highlights the importance of ensuring equitable access to technological advancements to strengthen global investigative efforts.

Complementing the findings of this study, it is important to highlight that the use of social media for forensic purposes is not limited to the collection of public data. Forensic social network analysis can also provide valuable insights into interpersonal interactions, communication patterns, and emotional states through posts, comments, and language patterns. By cross-referencing this information with other forensic techniques, such as behavioral pattern analysis and digital background checks, investigators can gain a more holistic view of the personalities and motivations of those involved. This type of integrated analysis has proven particularly effective in cybercrime investigations, where the digital trail is often the only concrete evidence available. The continuous evolution of technology poses a challenge to forensic science, which must constantly innovate to keep up with new forms of manipulation and concealment of online information.

Moreover, the integration of social media as a tool for forensic analysis not only broadens investigative possibilities but also enables a more preventive approach. Through continuous monitoring of suspicious behaviors and digital interaction patterns, it is possible to detect early signs of criminal behavior or psychological vulnerabilities that criminals can exploit. This predictive approach contributes to the creation of more effective security systems. It can inform the development of public policies to protect online users, especially the most vulnerable, such as the elderly people. Thus, forensic science, combined with digital behavioral analysis, plays a crucial role in maintaining public safety and enhancing investigative practices.

Therefore, the dissemination of technical knowledge should also be equally addressed for the general population but, especially the elderly, need to be aware of different types of scams to avoid future problems. Initiatives aimed at digital inclusion for the elderly are extremely welcome in the current scenario.

It is concluded that leveraging social media as a forensic tool is essential for modern investigations, especially in cybercrimes. The combination of psychoanalytic techniques with the collection of digital data allows for a deeper understanding of psychological motivations and behavioral patterns. Future research should focus on exploring the application of artificial intelligence to automate the identification of risk profiles, further expanding the reach of this methodology.

Finally, educational institutions and professionals in the forensic field need to promote ongoing training in digital and analytical skills, preparing new generations for the challenges that the age of information poses. The interdisciplinary approach between psychology, technology, and law should be encouraged, creating a body of knowledge that goes beyond criminal investigation to include and developing prevention strategies. Furthermore, collaboration between security agencies, researchers, and the community is vital for creating a safer and more informed environment, where protecting citizens from cybercrime becomes a shared priority. Forensic science, by embracing and integrating these innovative methods, will be more prepared to tackle upcoming challenges and uphold justice in a world that is becoming increasingly digital.

### Works cited

- Lerner, J. S., Li, Y., Valdesolo, P., & Sherman, D. K. "The Role of Affect in Decision Making." *Journal of Behavioral Decision Making* 28, no. 2 (2015): 125-137. Accessed August 15, 2024. <https://doi.org/10.1002/bdm.1898>.
- Van Lange, P. A. M., de Lange, J. M. S. J. R., & van der Meer, K. "Affect and Decision Making." *Annual Review of Psychology* 65 (2014): 123-145. Accessed August 17, 2024. <https://doi.org/10.1146/annurev-psych-010213-115234>.
- Wang, P., & Johnson, S. "Human Factors in Security: Social Engineering Attacks." *Journal of Cybersecurity Research* 15, no. 3 (2019): 204-219. Accessed August 23, 2024. <https://doi.org/10.1016/j.jcsr.2019.07.012>.
- Gupta, A., & Sharman, T. "Human Factors in Security: Social Engineering Attacks." *International Journal of Information Security* 45, no. 4 (2020): 389-405. Accessed September 04, 2024. <https://doi.org/10.1007/s10207-020-00567-3>.
- Wilson, R. P. "Digital Forensics and Social Media Investigation." *Arabian Journal for Science and Engineering* 46, no. 6 (2021): 457-470. Accessed September 10, 2024. <https://doi.org/10.1007/s13369-021-05924-6>.



## The Role of Anti-PD-1 as A Treatment for Cancer By Siena Lin

### Abstract

Immunotherapy represents a therapeutic strategy to treat various cancers, involving harnessing one's immune system to fight cancer. One common type of immunotherapy is anti-PD-1 treatment, which is commonly used to treat melanoma patients and patients with different cancers. In melanoma, patients typically respond well to anti-PD-1. In contrast, patients with other types of cancers such as glioblastoma do not respond as well to the treatment. Anti-PD-1 may still have a role in tumors such as glioblastoma, but not as a single-agent therapy. Identifying more efficient combinations of therapies will be critical to successfully treating these patients.

### Introduction

According to the American Cancer Society, nearly 2 million people in the United States were diagnosed with cancer in 2022 and of these, nearly  $\frac{1}{3}$  of those patients passed away due to cancer (*Cancer Facts & Figures 2022*, n.d.). These statistics demonstrate that more durable, efficacious treatments are needed to address cancer. Currently, FDA-approved cancer treatments include chemotherapy, hormone therapy, hyperthermia, photodynamic therapy, targeted therapy, surgery, stem cell transplant, radiation therapy, and immunotherapy. Depending on the type of cancer, different treatments have varying levels of efficacy.

One of the newest cancer treatments is immunotherapy, which was first approved by the FDA in 1986 for hairy cell leukemia (Eno, 2017). Immunotherapy has become mainstream only in the last 15 years. The goal of immunotherapy is to harness the immune system to attack cancer cells. Some cancers respond to this therapy more effectively than others. For example, two aggressive forms of cancer, melanoma (skin cancer) and glioblastoma (brain cancer), respond very differently to immunotherapies, both in terms of efficacy and side effects. Immunotherapy is very effective in melanoma, but not in glioblastoma.

Normally, cancer cells use a variety of mechanisms to avoid the immune system. Immunotherapy works by boosting or changing how the cells in the immune system respond to tumor cells. Within immunotherapy, there are many types of treatments that are often used individually or in combination. These treatments include the following: cellular therapies, anti-tumor vaccines, and immune checkpoint blockade.

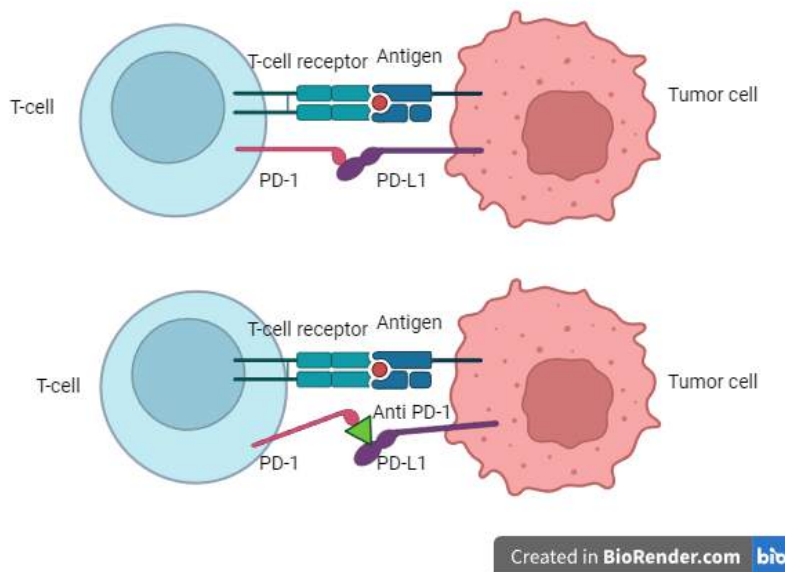
Immune checkpoint blockade is one of the most commonly used immunotherapy treatments. Checkpoints are cell-surface receptor proteins that control the effector function of cells. In the absence of checkpoint blockade, the ligand binds with the receptor which shuts down the T-cell's ability to fight off the tumor. Checkpoint blockade blocks the connection between a receptor on the tumor cell and a T-cell receptor. By administering ICB, T-cells can resume their typical effector function and kill the tumor cells. The two most commonly researched and used checkpoint blockades are anti-PD-1 and anti-CTLA-4. This paper will

review immune checkpoint blockade, specifically anti-PD-1. In particular, this review will discuss the mechanism of action of anti-PD-1, as well as its effectiveness in other cancers.

### Mechanism of Anti-PD-1

Anti-PD-1 inhibits the PD-1 receptor on T-cells. In the context of a tumor, PD-1 binds PD-L1 on a tumor cell. When the T-cell and tumor cell signal through these receptors, the T-cell undergoes dephosphorylation, where its normal functions are modified. This reduces proliferation, activation, and cytokine production, and leads to altered metabolism and cytotoxic T lymphocytes (CTLs) killer functions, and eventually apoptosis (Egger et al., 2016). Overall, signaling through the PD-1 receptor shuts down the T cell's ability to fight off the tumor. The role of anti-PD-1 is to intercept the connection between PD-1 on the T cell and PD-L1 on the tumor cell receptors. Blocking PD-1 stops the interaction with PD-L1 from affecting the T-cell, preventing the "shut down" of that T-cell. The T-cell is then able to function normally and kill the tumor cell. The end goal is that by inhibiting these receptors, the T-cells are able to properly target the tumor cells, resulting in tumor cell death and ultimately remission.

For some cancers, such as melanoma, checkpoint blockade is an extremely effective treatment. Checkpoint inhibitors, specifically against CTLA-4 and PD-1 have shown to be much more efficient compared to older treatments like chemotherapy (Raedler, 2019). To increase the number of patients who can benefit from checkpoint inhibition, researchers are looking at combining multiple checkpoint inhibitors with other therapies for maximum benefit. However, for some cancers, such as glioblastoma, the lack of T-cells that traffic into the tumor presents an additional challenge to anti-PD-1 therapy.



**Figure 1. Mechanism of Anti-PD-1** Anti-PD-1 inhibits the binding of PD-1 to PD-L1.  
Created with BioRender

### **Anti-PD-1 and Melanoma**

Melanoma is the deadliest form of skin cancer. Patients with metastatic melanoma, cancer that spreads beyond the origin (also called stage IV cancer) (National Cancer Institute, 2020), have a 5-year survival rate of only 16%. Unfortunately, based on an analysis from 1970 to 2009, patients with metastatic melanoma between the ages of 18 and 39 have increased significantly (women - 8 fold, men - 4 fold). Since tumor cells can spread quickly, surgery and radiation therapy are often ineffective. The first cancer to receive approval for the use of immune checkpoint blockade was metastatic melanoma (Alexander, 2016). Three types of checkpoint blockade have been approved, CTLA-4, PD-L1, and PD-1 (Raedler, 2019). Of these anti-PD-1 drugs, there are two approved in melanoma, nivolumab and pembrolizumab (Raedler, 2019).

In the KEYNOTE-001 clinical trial, pembrolizumab (PD-L1 inhibitor) was shown to be very effective in melanoma patients, including treatment-naive patients. The estimated 5-year overall survival rate increased from 16% to 34% in all melanoma patients. This was further increased to 41% in treatment-naive patients (Hamid et al., 2019). According to the analysis of the data from KEYNOTE-001, “The 5-year estimated PFS rate was 21% and 29%, respectively. Median PFS was 8.3 months (95% CI, 5.8–11.1) in all patients and 16.9 months (95% CI, 9.3–35.5) in treatment-naive patients” (Hamid et al., 2019). This shows that the progression-free survival (PFS) for patients is relatively high, especially in treatment-naive patients due to the lack of exposure and therefore lack of immunity towards pembrolizumab. This analysis of the 5-year study “confirms the durable antitumor activity” in advanced melanoma and how well one can tolerate the treatment (Hamid et al., 2019). The success rate of pembrolizumab alone is 30-40%.

Though pembrolizumab was shown to be a very durable treatment for melanoma, 86% of patients faced treatment-related adverse events (AEs), which included any side effects experienced by the patients. There are multiple grades of AEs, the higher the grade, the more severe the AE is. In the KEYNOTE-001 study, 17% of patients experienced grade 3 or 4 treatment-related AEs (TRAE), and no grade 5 was reported (Hamid et al., 2019). 36 of the 411 patients discontinued the treatment due to TRAEs.

### **Anti-PD-1 and Glioblastoma**

Glioblastoma (GBM) is an aggressive form of brain cancer with a 5-year OS of 4.7% (Jaoude et al., 2019). Often, the course of treatment followed by patients with GBM, if not spread too much, would be surgical resection, followed by chemotherapy and radiation. Given the 5-year OS with all these treatments, the efficacy of current standard-of-care therapies for patients with GBM still remains poor.

In the brain, inflammation is limited because there is a finite amount of space available in the human skull. As a result, immune cells are suppressed when entering the brain. Immune infiltration is when immune cells enter the tumor. Immune infiltration typically helps fight cancer, however, in the setting of GBM, the tumor changes the infiltrating macrophages into tumor-associated macrophages, which create an immunosuppressive tumor microenvironment.

PD-1 signaling is just one of the many challenges in the tumor microenvironment. In addition, patients with GBM also suffer from a compromised immune system because GBM affects the immune system outside of the brain (Lorrey et al., 2023).

Due to the many genetic variations in the tumor of GBM patients (Rocha Pinheiro et al., 2023), it is difficult to treat with anti-PD1 or any type of immune checkpoint inhibitor. In CheckMate 143, nivolumab, an anti-PD-1 inhibitor, is compared to bevacizumab, a chemotherapy, in GBM patients. Unfortunately, the median overall survival was similar in both trials, 9.8 months for nivolumab and 10.0 months for bevacizumab. Both treatments resulted in a 12-month overall survival of 42%. These results show the low efficacy of nivolumab in GBM patients.

## **Discussion**

Immunotherapy has significantly fewer risks and costs when compared to surgery, radiotherapy, and chemotherapy, allowing cancer treatments to become more accessible (Pinheiro et al., 2023). However, there still are risks associated with immunotherapy, particularly in the AEs that may occur. Most people overlook how much AEs can affect one's recovery from cancer. Due to the high toxicity levels of therapies, specifically anti-PD-1, which is necessary for the therapy to be effective, patients often face AEs, sometimes so severe that it may lead to patients discounting treatment or mortality. Due to these harsh consequences of the therapies, more research must be focused on decreasing the severity of AEs.

Anti-PD-1 will likely be a vital ingredient to the cocktail of drugs, but not the only ingredient. Researchers expect combination therapy to be the most effective treatment for patients with GBM (Pinheiro et al., 2023). Trials have shown that combining pembrolizumab and bevacizumab can prolong survival time and benefit tumor control (Yang et al., 2020). Though the results can be discouraging, immunotherapy is viewed as a fundamental factor in increasing patient life expectancy and quality of life.

As most therapies, anti-PD-1 is not a one-size-fits-all solution to every and all cancers. Though it provides a durable (18% increased survival rate (Hamid et al., 2019)) response to melanoma, patients can still experience severe AEs. These AEs can downgrade a patient's quality of life significantly, something possibly as important as OS at times. This is even more true for glioblastoma and other cancers. Researchers will need to continue to persevere and be thoughtful about different drug combinations and rationally design interventions with the specific challenges for each cancer in mind. This will take collaboration from many different fields, but once successful, could save millions of lives.

## Works Cited

- Alexander, W. (2016, March). The Checkpoint Immunotherapy Revolution - PMC. NCBI. Retrieved March 5, 2024, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4771089>
- Egger, J. V., Lane, M. V., Antonucci, L. A., Dedi, B., & Krucher, N. A. (2016, September 19). Dephosphorylation of the Retinoblastoma protein (Rb) inhibits cancer cell EMT via Zeb. NCBI. Retrieved March 5, 2024, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5137485/>
- Luoyan, A., Antao, X., & Jie, X. (2020). Roles of PD-1/PD-L1 Pathway: Signaling, Cancer, and Beyond. *Advances in Experimental Medicine and Biology*, 1248, 33-59. 10.1007/978-981-15-3266-5\_3.
- Raedler, L. A. (2019, June 24). Keytruda (Pembrolizumab): First PD-1 Inhibitor Approved for Previously Treated Unresectable or Metastatic Melanoma. National Library of Medicine. Retrieved December 19, 2023, from [https://dpl6hyzg28thp.cloudfront.net/media/anti-PD1\\_in\\_Melanoma.pdf](https://dpl6hyzg28thp.cloudfront.net/media/anti-PD1_in_Melanoma.pdf)
- Hamid, O., Robert, C., Daud, A., Hodi, F. S., Hwu, W. J., Kefford, R., Wolchok, J. D., Hersey, P., Joseph, R., Weber, J. S., Dronca, R., Mitchell, T. C., Patnaik, A., Zarour, H. M., Joshua, A. M., Zhao, Q., Jensen, E., Ahsan, S., Ibrahim, N., & Ribas, A. (2019). Five-year survival outcomes for patients with advanced melanoma treated with pembrolizumab in KEYNOTE-001. *Annals of oncology : official journal of the European Society for Medical Oncology*, 30(4), 582–588. <https://doi.org/10.1093/annonc/mdz011>
- Moser, J. C., Wei, G., Colonna, S. V., Grossmann, K. F., Patel, S., & Hyngstrom, J. R. (2020). Comparative-effectiveness of pembrolizumab vs. nivolumab for patients with metastatic melanoma. *Acta oncologica (Stockholm, Sweden)*, 59(4), 434–437. <https://doi.org/10.1080/0284186X.2020.1712473>
- Jaoude, D. A., Moore, J. A., Moore, M. B., Twumasi-Ankrah, P., Ablah, E., & Moore, D. F., Jr (2019). Glioblastoma and Increased Survival with Longer Chemotherapy Duration. *Kansas journal of medicine*, 12(3), 65–69.
- Pinheiro, S. L. R., Lemos, F. F. B., Marques, H. S., Luz, M. S., De Oliveira Silva, L. G., Santos, C. F. S. M. D., Da Costa Evangelista, K., Calmon, M. S., Loureiro, M. S., & De Melo, F. F. (2023). Immunotherapy in glioblastoma treatment: Current state and future prospects. *World Journal of Clinical Oncology*, 14(4), 138–159. <https://doi.org/10.5306/wjco.v14.i4.138>
- Lorrey, S., Polania, J. W., Wachsmuth, L. P., Hoyt-Miggelbrink, A., Tritz, Z., Edwards, R., Wolf, D. M., Johnson, A. J., Fecci, P. E., & Ayasoufi, K. (2023). Systemic immune derangements are shared across various CNS pathologies and reflect novel mechanisms of immune privilege. *Neuro-oncology Advances*, 5(1). <https://doi.org/10.1093/noajnl/vdad035>
- Yang, T., Kong, Z., & Ma, W. (2020). PD-1/PD-L1 immune checkpoint inhibitors in glioblastoma: clinical studies, challenges and potential. *Human Vaccines & Immunotherapeutics*, 17(2), 546–553. <https://doi.org/10.1080/21645515.2020.1782692>
- Cancer Facts & Figures 2022. (n.d.). <https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2022.html>
- Eno, J. (2017, December 1). Immunotherapy through the years. PubMed Central (PMC). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6188092/>

# **Genetic and Parental Behavior's Effect on the Development of Childhood Anxiety**

## **By Aon Huang**

### **Introduction**

The factors that contribute to anxiety disorders are not only diverse, but complex. The development of anxiety disorders in children under 18 years of age is affected by environmental and genetic factors. These factors can also interact; for example, the heritability of anxiety disorders can impact how the influence of environmental factors on the development of anxiety symptoms. While anxiety on its own is a perfectly normal experience, anxiety disorders (heretofore, Anxiety) are diagnosed due to persistent symptoms and stress, and clinically significant impairment in someone's life. There are many experiences that can develop childhood anxiety disorders, but the focus of this review will mainly focus on specific parenting styles. We will cover authoritarian versus permissive parenting styles and how either style may be connected to Anxiety development in children. We will also cover how interactions between genetic risk for Anxiety may make parents feel inclined to use one of the more problematic parenting styles, furthering childhood risk for Anxiety.

Heritability refers to the genetic transmissibility of a trait between a parent and their child, in this case, the transmissibility of anxiety disorders. When it comes to the heritability of Anxiety, an analysis of many twin studies determines the heritability as 31.6%, meaning that 31.6% percent of a child's Anxiety development is due to the genetic risk they inherit from their parents (Gottschalk & Domschke, 2017). Different types of anxiety disorders, however, vary in heritability. For example, social anxiety disorder ranges from 13-76% (Moreno et al., 2016) compared to Generalized anxiety disorder, which is 25-35% heritable (Hettema et al., 2005). In other words, about 70% of the development of Anxiety is due to environmental factors. While heritability and genetic influence do play a role, environmental factors are by far more impactful on the development of anxiety disorders.

The focus of this paper is one specific environmental factor: how parents behave toward their children, called parenting styles. It is commonly held that the emotional development of children is based on the support that parents give, which is typically stagnant over time. Compared to everyday life events - called discrete life events - chronic stressors including abuse, neglect, or household dysfunction more significantly contribute to the development of anxiety (Allen et al. 2008). Since parents are a part of a child's everyday life, it is important to analyze how we parent children in order to be aware of what preventative measures we can take to diminish the development of anxiety in children. Specifically, this paper addresses the following questions: How do parental behaviors develop anxiety in children? Do specific cultures or comorbidity contribute to this development? What do we do with this information?

### **Parenting Styles & Anxiety Disorders**

There are many different parenting styles that can be implemented in a parenting relationship. Researchers often classify and describe parenting styles by the frequency and

intensity of the way parents exert control, warmth, and/or discipline of their children. This paper will mainly focus on Authoritarian and Permissive parenting and how each style can be detrimental towards an eventual development of anxiety. Authoritarian parents are parents who have high expectations and strict rules that are paired with severe punishments for children. Permissive parents let their children have a lot of autonomy in their decision making, creating a more hands-off environment (Wei & Kendall, 2014). These styles are often compared to Authoritative parenting, which gives a more positive balance of both these styles. The children of parents who use authoritarian and/or permissive parenting behaviors are more likely to develop anxiety disorder compared to other parenting styles (Kadoglou et al., 2024).

Authoritarian parents tend to be more demanding of their children; and due to high demands and expectations that children fear they may not be able to meet, creating greater symptoms of anxiety (Warner & Strawn, 2023). Since authoritarian parents tend to give consequences to their children to make the children keep up with their demands, threats are more apparent to children, and they become more sensitive to cues that lead to threats (Warner & Strawn, 2023), a key feature of anxiety disorders. The situations that authoritarian parents put their children in cause them to be more anxious of possible punishments that come their way if they mess up. This heavily contributes to the development of anxiety disorders because of the hostile environments that authoritarian parents surround their children in. This causes children to function worse in social environments like school, where they may make fewer friends and participate less in class due to social anxiety (Warned & Strawn, 2023). In contrast, when parental demandingness is low, the association between executive functioning deficits and anxiety is weaker (Warner & Strawn, 2023). Thus, low parental demandingness may protect children from developing anxiety disorders and parental warmth during adolescence is associated with less anxiety in adulthood (Warner & Strawn, 2023).

On the other end of the control spectrum, the lack of control from permissive parents inhibits the development of competence and success in children (Pinquart & Gerke, 2019). This can happen because permissive parents don't push and challenge their children as much as authoritative parents; therefore, children don't develop personal resilience and growth as they develop. Parents being more hands-off can also lead to children developing a sense of helplessness. Children can also perceive permissive parents as more distant, and tend to develop attachment anxiety. Overall, permissive parenting leads to the development of more internalized problems for children, which is a risk factor for anxiety (Ciuhan, 2024).

### **Interaction between Parenting Styles and Anxiety Disorders**

Genetics and parental behaviors are heavily intertwined when it comes to causing anxiety disorders. For example, parental anxiety disorders increase the risk of their children developing anxiety disorders. We can see this through anxious parents' exhibition of more controlling behaviors, as well as being more negative and critical of their children (Eley et. al, 2015). This shows that the parents pass on anxiety disorders not only through genetics, but also through their behaviors. Since parents can both pass on genetic information and express environmental

behaviors, the existence of anxiety disorders in parents can heavily increase the chances of anxiety disorders in children.

Anxiety disorders also have high comorbidity with other psychiatric illnesses (Koskinen & Hovatta, 2023). In fact, 11% to 40% of children have comorbid anxiety and ADHD (Doyle, 2022). The existence of other disorders may cause a shift in parenting behaviors causing further development of anxiety disorders.

Culture's impact on the development of anxiety disorders is relatively uninvestigated, but we can see that across cultures, fathers tend to spend less time with their children, being less sensitive and available to their children (Lawrence et. al, 2019).

Overall, the genetic input that contributes to anxiety disorders is exacerbated by negative parent behaviors that children may face. The background of the children and the type of parenting is integral to the development of anxiety. It is important that parents understand how they should parent their children in order to avoid their children developing anxiety disorders. It's also important to understand how anxiety can be comorbid as anxiety often ends up showing up along with more genetically detectable illnesses. For example, 11% to 40% of children have comorbid anxiety and ADHD (Doyle, 2022). Moreover, parenting behaviors can affect different genetics differently; as anxiety is comorbid with other mental illnesses, it is important to understand that different parenting behaviors may be more harmful in some situations than others. It is important for us to understand that the authoritarian parenting style is harmful as authoritarian parents provide low acceptance/warmth, place high demands as well as high control/restriction of autonomy, and exert punitive measures (Wei and Kendall, 2014) which breeds fear, which can contribute to the development of anxiety disorders.

From this paper, we can make the conclusion that authoritarian parenting and permissive parenting have negative effects on children, increasing the likelihood of childhood anxiety. However, we can see that the actions presented should be avoided, as not being over controlling like authoritarian parents, but not being too hands off like permissive parents. Avoiding these negative parenting behaviors lead to authoritative parenting, a mix of both that contains more positive behaviors, like engagement with children. The more the world adopts authoritative parenting, the less childhood anxiety there will be.

## **Methods**

To conduct this literature review, a search using keywords and Boolean operators was performed across Pubmed and Google Scholar [add any other search operators you used]. Inclusion criteria for study review were: empirical (e.g., an experiment was conducted or two groups with different levels of anxiety were compared across measures), and the main focus of the study was anxiety and heritability or parenting behaviors. Many of the studies identified were control studies, twin studies, or survey studies, which yielded a variety of sources and data to review. Relevant descriptive statistics (e.g., mean, correlation coefficient) and effect sizes (e.g., Beta values, Cohen's d, t-statistics) were extracted from the empirical studies to qualitatively



describe and review the effect of different parenting behaviors on anxiety development under different genetic risk.

## Works Cited

- Allen, Jennifer L., and Ronald M. Rapee. "Are reported differences in life events for anxious children and controls due to comorbid disorders?" *Journal of Anxiety Disorders*, vol. 23, no. 4, May 2009, pp. 511–518, <https://doi.org/10.1016/j.janxdis.2008.10.005>.
- Cucu Ciuhan, Geanina. "Relationship between permissive parenting style and atypical behaviour in preschool children, with generalized anxiety as mediator." *Early Child Development and Care*, vol. 194, no. 4, 28 Dec. 2021, pp. 624–632, <https://doi.org/10.1080/03004430.2021.2017908>.
- Doyle, Melissa M. "Anxiety disorders in children." *Pediatrics In Review*, vol. 43, no. 11, 1 Nov. 2022, pp. 618–630, <https://doi.org/10.1542/pir.2020-001198>.
- Eley, Thalia C., et al. "The intergenerational transmission of anxiety: A children-of-twins study." *American Journal of Psychiatry*, vol. 172, no. 7, 1 July 2015, pp. 630–637, <https://doi.org/10.1176/appi.ajp.2015.14070818>.
- Gottschalk, Michael G., and Katharina Domschke. "Genetics of generalized anxiety disorder and related traits." *Dialogues in Clinical Neuroscience*, vol. 19, no. 2, 30 June 2017, pp. 159–168, <https://doi.org/10.31887/dcns.2017.19.2/kdomschke>.
- Hettema, John M., et al. "The structure of genetic and environmental risk factors for anxiety disorders in men and women." *Archives of General Psychiatry*, vol. 62, no. 2, 1 Feb. 2005, p. 182, <https://doi.org/10.1001/archpsyc.62.2.182>.
- Kadoglou, Maria, et al. "Preschoolers and anxiety: The effect of parental characteristics." *Journal of Child and Adolescent Psychiatric Nursing*, vol. 37, no. 1, 10 Dec. 2023, <https://doi.org/10.1111/jcap.12445>.
- Koskinen, Maija-Kreetta, and Iris Hovatta. "Genetic insights into the neurobiology of anxiety." *Trends in Neurosciences*, vol. 46, no. 4, Apr. 2023, pp. 318–331, <https://doi.org/10.1016/j.tins.2023.01.007>.
- Lawrence, Peter J., et al. "Environmental factors in the development and maintenance of Anxiety Disorders." *Pediatric Anxiety Disorders*, 2019, pp. 101–124, <https://doi.org/10.1016/b978-0-12-813004-9.00006-2>.
- Moreno, André Luiz, et al. "Heritability of social anxiety disorder: A systematic review of methodological designs." *Archives of Clinical Psychiatry (São Paulo)*, vol. 43, no. 4, Aug. 2016, pp. 83–92, <https://doi.org/10.1590/0101-60830000000090>.
- Pinquart, Martin, and Dana-Christina Gerke. "Associations of parenting styles with self-esteem in children and adolescents: A meta-analysis." *Journal of Child and Family Studies*, vol. 28, no. 8, 20 Apr. 2019, pp. 2017–2035, <https://doi.org/10.1007/s10826-019-01417-5>.
- Warner, Emily N., and Jeffrey R. Strawn. "Risk factors for pediatric anxiety disorders." *Child and Adolescent Psychiatric Clinics of North America*, vol. 32, no. 3, July 2023, pp. 485–510, <https://doi.org/10.1016/j.chc.2022.10.001>.
- Wei, Chiaying, and Philip C. Kendall. "Parental involvement: Contribution to childhood anxiety and its treatment." *Clinical Child and Family Psychology Review*, vol. 17, no. 4, 15 July 2014, pp. 319–339, <https://doi.org/10.1007/s10567-014-0170-6>.

## **Athletes and Mathletes: How Non-contact Sports Improve Math Learning By Improving Students' Spatial Skills By William McAfee**

### **Abstract**

In the 2012 hit movie “21 Jump Street,” Channing Tatum and Jonah Hill portrayed the stereotypical dichotomy of jocks and nerds. However, going against usual expectations, they were not working against each other but rather working together. In art as in life, the connection between sports and academics can pull together rather than pulling apart. Students who participate in sports reap a variety of benefits in regards to their academic performance. But what sports should they play, and how will it help them? What are the skills that sports might reinforce that improve academic performance?

Existing research debates the extent to which sports participation, in general, improves academic performance. Studies find that there is a correlation (Darling et al. 2005), while others find no correlation when controlling for certain factors (Rees and Sabia 759). However, little literature dives into the fundamentals that sports teach that connect to mathematical abilities. This paper analyzes both what sports improve math performance as well as why this effect is present.

By comparing contact and non-contact sports, this paper highlights how non-contact sports benefit mathematical learning. Sports participation of any type can inspire kids to succeed academically so they can go to college and continue playing the game they love; sports can boost their confidence and sports can push them to succeed. However, I argue that beyond these other benefits, non-contact sports improve math education because they emphasize overlapping skills.

Because non-contact sports focus on spatial awareness over potential threats to physical safety, they help students improve core math abilities in spatial orientation, mental rotation, and spatial visualization. Spatial orientation is the ability to perceive the position of opponents and objects while playing sports and improves geometric abilities. Mental rotation, the ability to recognize how spin and trajectory affect an object's motion, exercises geometric abilities and logical reasoning skills. Spatial visualization, the ability to mentally imagine situations and analyze outcomes, has the greatest overall translation to mathematical success. Athletes perform higher than non-athletes on tests of all three of these abilities, demonstrating how participation in sports improves fundamental math skills due to heightened spatial abilities.

This paper first defines the core concepts of sports and mathematical skills that are necessary in understanding the argument being presented. Then, it discusses alternative explanations including family background, location, college aspirations, confidence, and gender. Finally, I present my argument that non-contact sports are superior to contact sports in improving math skills because of their heightened focus on spatial awareness by placing the core concepts into conversation with alternative explanations.

This paper speaks to scholars interested in the connection between sports participation, spatial awareness, and mathematical performance. It is also relevant to activists working on

improving mathematical achievement, especially among inner-city youth, and public officials who are concerned with funding for organizations focused on access to sports and academics.

## **Introduction**

In the 2012 hit movie “21 Jump Street,” Channing Tatum and Jonah Hill portray the stereotypical dichotomy of jocks and nerds. However, they were not going against each other but actually working together. The connection between sports and academics in real life has a similar union. Students who participate in sports reap a variety of benefits in regards to their academic performance. But what sports should they play, and how will it help them? What are the skills that sports might reinforce that improve academic performance?

Existing research debates the extent to which sports participation in general improves academic performance. Some studies find that there is a correlation, while others find no correlation when controlling for certain factors, such as socioeconomic status or location. However, little literature dives into the fundamentals that sports teach that connect to mathematical abilities. This paper attempts to analyze both the type of sports that helps improve academic performance the most as well as why this effect occurs.

The two broad categories of sports being examined are contact and non-contact sports, and the academic benefits from sports participation that are focused on are mathematical benefits. In general, sports participation can inspire kids to succeed academically so they can go to college to continue playing the game they love, can boost their confidence, and can push them to succeed. However, I argue that beyond these other factors, non-contact sports improve math education because they emphasize overlapping skills. Because non-contact sports focus on spatial awareness over potential threats to physical safety, they help students improve core math abilities in spatial orientation, mental rotation, and spatial visualization.

This essay will first describe the core concepts that are necessary for understanding the argument being presented. Then, alternative explanations for a student’s math success, both related and unrelated to sports participation, will be analyzed as well. Finally, I present my argument that non-contact sports are superior to contact sports in improving math skills because of their heightened focus on spatial awareness by synthesizing the context of the core concepts with responses to alternative explanations.

## **Core Concepts**

It is important to understand the way that sports and math learning will be defined for the purposes of this paper. While the line between what is and is not a sport is not completely solid, and the ways by which students learn mathematical skills are not either, the next section attempts to define those lines in regards to how the terms will be used throughout the paper.

## **Sports**

While there are many potential definitions of what makes a sport a sport, the critical distinction for the impact of sports on academic learning is that a sport is an activity that requires

technical skill beyond simply a mental capacity. The sports and physical activities that are being observed in this paper involve “permanent adaptation of movement to a dynamic environment” and “movement in a relatively stable environment” (Gentile 2000 cited in Jansen and Pietsch 7133). These definitions encompass a wide variety of sports, from football to tennis to esports, as all of them fit this definition and require these technical abilities in some form.

One group of sports that push the boundaries of the definition are esports because they require a player's ability to use their hands on either a keyboard and mouse or a controller to direct their character's body. Esports players must have strong fine motor skills in order to precisely use their computers or consoles in a way that helps them outplay their opponent (Jenny et al. 2017 cited in Chikish et al. 478). While the esports athletes are doing a small amount of movement themselves, they are also moving the character they are controlling around a changing environment and noting where the opponents are, requiring them to be spatially aware of what is going on around them, not necessarily in the physical world but very much so in the virtual world.

A line between what is and what is not a sport has to be drawn somewhere, and a good example that defines what is not a sport is chess. Chess, despite requiring a massive amount of mental effort, strategy, and involving competition, does not require any physical activity besides moving one's hand and the piece that hand is moving. All sports depend on some sort of movement or physical skill, not just a physical action (Gentile 2000 cited in Jansen and Pietsch 7133). Since chess does not require any significant physical movement, exercise, or skill, it should not be classified as a sport despite some of the similar characteristics it possesses. Because of this, games like chess or competitions such as Academic Decathlon are not being reviewed in this paper because they do not fit the definition of a sport.

Participation in sports leads to improved spatial awareness, which is a crucial aspect of mathematical achievement. Jansen and Pietsch report the correlation between participation in sports and spatial awareness: “In studies with athletes, it has been demonstrated that sports students performed better in spatial tasks than physically inactive people” (Pietsch and Jansen 2012 cited in Jansen and Pietsch 7134). Simply being an athlete and participating in a sport can have substantial benefits regarding heightened spatial abilities. These abilities can then be used in the classroom, where students can have a better sense of mathematical concepts, such as how to translate objects in a coordinate plane, because they are familiar with spatial visualization from playing their sport.

Contact sports are defined as sports in which athletes are required to interact in a shared space, such as basketball or football, while non-contact sports give each individual or team their own space that the other team is not allowed to enter, such as tennis or volleyball. Physical contact in some form will most likely occur in contact sports, whereas in non-contact sports, it is much more unlikely or impossible (Campo et al. 2016 cited in Castro-Sánchez et al. 3). The distinction between contact and non-contact sports must be made because it is important to recognize how the differences in playing environment can affect the skills that are translated to the classroom.

Contact sports emphasize spatial awareness in regards to how an athlete's own body should be used based on where their opponents are, such as where to go to tackle a running back in football. Contact sports require an athlete to focus on a large number of moving parts all at one time (Castro-Sánchez et al. 3). Contact sport athletes benefit from having a large number of obstacles in their playing environment, forcing them to be extremely spatially aware. The obstacles, however, also impede on an athlete's ability to solely focus on technique.

Athletes that play contact sports are forced to sacrifice focus on technical skills because of the constantly changing environment around them that may pose a threat to their safety. Because contact sport athletes have to focus on additional stressors in their environment, they cannot fully focus on the skills required to play their sport at all times (Giza et al. 2017 cited in Castro-Sánchez 12). They are subconsciously in a state of anxiety, and their brain is helping them avoid competitors in their space, which helps spatial awareness passively but not actively. Because they cannot consciously use their brain power on technical skills and how to use the space around them, contact sport athletes are hindered by their environment in terms of actively improving their spatial awareness skills compared to non-contact sports.

Non-contact sports, on the other hand, involve less physical contact naturally, and therefore, athletes can focus on other spatial skills based on factors they can control. One skill requires athletes to keep themselves and the equipment that they are handling in a steady position (Jansen and Pietsch 7133). This forces athletes to focus on how they are balancing their body and flexing their muscles, trying different combinations until one method works, in order to remain stable. This methodical approach translates to being able to handle math problems methodically as well.

Athletes in non-contact sports also have more time and less distractions that allow them to focus on spatial skills relative to equipment. Players are forced to recognize their position and how it relates to a ball, racket, or other piece of equipment, making sure they are neither too close nor too far to achieve optimal play (Jansen and Pietsch 7133). This heavily emphasizes an athlete's ability to consciously focus on the space between themselves and sports gear and then thinking about the best way to approach the task at hand. This translates to trial-and-error skills in mathematics and problem-solving based on previous outcomes because, just as in sports, the student-athletes are trying to find the best approach to a problem.

### **Academic Skills**

There are four core abilities which pedagogy scholars use to define math learning. Numerical skills incorporate counting, and the ability to manipulate numbers; arithmetic abilities involve solving for variables and understanding step-by-step procedures; geometric skills concern the ability to deal with shapes; and logical reasoning involves solving word problems and coming up with ways to solve problems (Xie et al. 2019 cited in Jansen and Pietsch 7133). The four broad categories of types of mathematical knowledge demonstrate the different skills that students need to work towards mastery of in order to have success in their math classes and assessments.

Spatial awareness is a skill that can be learned from participating in sports and translated to the classroom. In a study connecting sports, spatial awareness, and math learning, spatial awareness helps math learning transitively as “sports students performed better in spatial tasks,” and “spatial abilities predict performance in STEM” (Jansen and Pietsch 7134). Because spatial awareness is reinforced by sports and useful in the classroom, the component factors of spatial awareness are helpful to understand.

Spatial awareness is composed of three component factors: orientation, visualization, and mental rotation. These can be defined as “perception of the position of various objects in space,” “imagining the rotation of objects,” and “perception of complex spatial patterns,” respectively (Jansen and Pietsch 7132). Because spatial awareness is relevant to both sports and math learning, a deeper dive into these factors is relevant.

Sports use spatial orientation because athletes need to perceive the position of their opponents, a ball, and many other aspects of various sports. Spatial orientation is a key part of geometric success, one of the components of math learning, and geometric success correlated with heightened spatial abilities in the most recently mentioned study (Jansen and Pietsch 7138). Because students participated in sports, their spatial orientation abilities improved and, with that, their geometric performance.

Spatial visualization can be seen in sports in the form of drawing up plays, such as deciding what route to run in football based on what the quarterback or coach imagines the opposing team will do. The study reported that spatial visualization contains the most similarity to mathematical skills (Jansen and Pietsch 7134). This means that participation in sports, which translates to spatial visualization abilities, can reinforce skills that help a student towards mathematical success due to the noticeable overlap between the two.

Another aspect of spatial awareness is mental rotation, such as recognizing the spin on a baseball to decipher what pitch is coming. Interestingly, the study found that performance on a mental rotation test was not improved by participation in sports, specifically jumping (Jansen and Pietsch 7135). However, this may be due to the fact that jumping itself does not require an athlete to perceive and react to an object rotating, which may have caused these results. Examining a sport where the rotation of an object is a major focus of the game proves that, in these types of sports, such as baseball, mental rotation skills are increased. A study on professional baseball players found that they were able to complete tests of their mental rotation skills much faster than peers who did not play sports (Kang et al. 9). The improved mental rotation abilities from sports improve geometric skills as well as logical reasoning due to the thought process required to analyze an object's flight based on spin.

Spatial awareness that can be learned from participation in sports translates strongly into classroom performance, especially in mathematics. Studies show that spatial abilities correlate with a student's success in STEM subjects (Mix and Cheng 2012 cited in Jansen and Pietsch 7134). Because sports help student-athletes improve their spatial awareness, and spatial awareness is a significant factor in mathematical performance, it can be reasonably inferred that

sports participation can help mathematical performance, specifically in aspects of math related to spatial awareness, such as geometry.

Some high-school students, especially those with parents who did not attend college or are from a family without the means to afford college, do not plan on attending college due to their circumstances. However, playing a sport may counteract this idea. A study on student-athletes reports that high-school athletes are found to have more chances of college aspirations than students who are not athletes (Mahoney et al. 2003 cited in DeMeulenaere 132). Attending college is, for many kids, a way to boost their human capital, which will help them secure higher-paying jobs out of college than their peers without diplomas. While some students from low-income families may not have the means to attend college, going on an athletic scholarship because they were a high school athlete provides one solution to that problem. While some students do not plan on attending college, being a high-school athlete creates an environment in which a desire to attend college is higher than students who do not play sports.

### **Other Explanations**

There are a number of explanations which can improve students' performance in math. Factors beyond spatial abilities, such as family background, a student's location, and other intangible effects noticeably influence a student's mathematical performance when they participate in athletic activities. These factors should be taken into account when analyzing the role of sports participation on academic success.

### **Family Background**

Having a parent in a STEM field of work, especially a girl with a mother in STEM, may artificially inflate math scores as they are conditioned, whether intentionally or unintentionally, through nurture to be more apt towards STEM-related subjects. In a study on the effect of sports on girls, girls with mothers that work in STEM or got a STEM-related degree were excluded: "We end up with a sample of girls whose educated mothers have reading-related skills...so that we reduce the possibility that our results depend on the fact that mother education shapes math skills more than reading" (Rose et al. 2371). Girls who play sports but have mothers in STEM were excluded from the study because their mothers may affect their math scores in a way that may demonstrate improvement because of sports but they are actually solely attributable to their mother's influence.

When examining boys' or girls' mathematical performance due to athletic participation, it should be noted that controlling for socioeconomic status is important because of the ways boys and girls of different statuses perform compared to one another. It has been found that in wealthier families, boys are traditionally stronger at mathematics than their peers, and conversely, girls are better at reading than their peers, both due to more opportunity and focus on education from a young age (Reardon et al. 2018 cited in Rose et al. 2371). Because wealthy families typically have boys who achieve in mathematics and sciences at a higher rate than boys from a lower socioeconomic status, and girls perform better overall academically as well, it is



important to account for this difference when examining the impact of sports on these two groups.

### **Urban vs. Rural**

Students from low socioeconomic classes and urban environments show significant benefits from increased physical activity in school. In a study on students from low socioeconomic students in Australia, they found that “students in the AMPED intervention group significantly improved their performance in mathematics, in comparison with students in the control schools” (Lubans et al. 14). The students that received high-quality physical education in the AMPED study were students with a low socioeconomic status from urban environments in Australia, demonstrating how these demographics are positively affected due to physical education being incorporated into their academic life.

Conversely, students in rural environments experience less benefit from sports and their effect on academic performance. A study on students in the rural Appalachian region of the United States found little change in academic performance depending on whether or not a student participated in regular physical activity (Din 9). Because of this, it is important to note the location in which a study on athletic participation and its effect on academic performance because an identical study conducted in both a rural environment and an urban environment might come up with different results independent from spatial skills.

### **Intangible Effects**

College aspirations are increased among student-athletes due to the opportunities available to continue playing sports at the collegiate level. In order for athletes to be able to continue to seriously play the sport they love beyond high school, they most likely must attend college. To get offers and scholarships for their sport at different colleges, they must meet academic baselines (DeMeulenaere 132). Therefore, participating in athletics at the high-school level inspires students to work hard in school in order to be able to compete at the next level. This motivation to improve, however, has little to do with spatial abilities and therefore may lead to academic improvement that is unrelated to improved spatial skills.

Students who participated in sports also had increased confidence in themselves, which can translate into academic ability. If a student is struggling academically, they have another outlet where they may find success: the field or court (DeMeulenaere 131). This gives a student a place to not worry about their success in the classroom and come back to their studies mentally refreshed with a boost of confidence from believing that if they can compete athletically they can compete academically too. This, too, may lead to boosts in academic performance that are independent from increased spatial abilities.

Participation in sports can help girls go against the status quo, motivating them to do so as well academically and participate in STEM. A study found that, in schools, a larger percentage of the female population that participates in athletics correlates with higher female participation in AP classes, notably in AP science classes (Veliz and Shakib 12). Girls playing

sports naturally puts them against the grain athletically, giving them the inspiration to go against the grain academically as well, as demonstrated by the higher overall AP enrollment, specifically in STEM-related courses such as AP science. This is another example of a way that students, specifically girls, may achieve math related improvement because of participation in sports that does not pertain to heightened spatial skills.

For some boys who already engage in a baseline level of athletics, both in and out of the school contexts, increased participation in sports begins to take away from time to study and do homework, leading to a decrease in academic performance. Because boys, who stereotypically and statistically have a higher baseline for hours of athletic participation per week, generally play more sports than girls do, an increase in hours of sports may result in a decrease in mathematical performance (Rose et al. 2373). It is important to note that increasing a boy's level of athletic participation would definitely improve his spatial abilities, but they may lose out on the most essential element of academic performance: practicing academics. Therefore, they could display a negative relationship between sports and mathematical performance despite heightened spatial awareness skills.

On a related note, girls may show a more prominent relationship between sports participation and mathematical performance because, for many girls, the baseline level of athletic competition is very low. A study analyzing the effect of sports on mathematical performance noted that many of the girls they chose spent little or no time playing sports at the beginning of the study, and, when the frequency of sports increased, the increase in math scores was more prominent than it was for boys with a similar increase in time spent playing sports (Rose et al. 2373). Hence, an increase in the amount of time playing sports can help increase girls' spatial abilities without impeding on their time to study.

### **My Argument: Skill Overlap**

In the previous section, I discussed how family background, a student's location, college aspirations, confidence, and gender can improve math scores in students. I argue that beyond these other factors, non-contact sports improve math education because they emphasize overlapping skills. Because non-contact sports focus on spatial awareness over potential threats to physical safety, they help students improve core math abilities in spatial orientation, mental rotation, and spatial visualization. To explicate this argument, I will use the definitions of sports and mathematical skills I have developed above and place them in conversation with the alternative explanations from the prior section in order to demonstrate how spatial awareness from non-contact sports improves mathematical performance.

Non-contact sports emphasize spatial awareness while reducing threats to physical safety. It has been established that both contact and non-contact sport athletes achieve improved spatial awareness skills due to physical activity. Contact sport athletes are required to share a space in which their bodies can collide with their opponents, unlike in non-contact sports. This means that non-contact sport athletes can consciously focus with their full attention on spatial orientation, mental rotation, and spatial visualization because they do not have their subconscious mind

working to prevent collisions with opponents, unlike the minds of contact sport athletes. Because of this subtle difference in brain activity, non-contact sport athletes are able to actively learn the skills that translate to mathematical performance in a more economical way.

Math learning skills overlap with spatial abilities, leading to an improvement in math performance from sports. Sports improve spatial orientation, which relates to how athletes perceive the position of objects and people around them and translates to geometric success in the classroom. They also improve spatial visualization, which is the ability to imagine scenarios in one's mind and which also connects most directly to mathematical skills. Additionally, they improve mental rotation, which is the ability to imagine how objects spin. Mental rotation translates to enhanced geometric and logical reasoning skills. Math learning involves skills that can be heightened by spatial awareness skills, thus proving that sports, which improve spatial awareness, can improve math learning.

Family background is an alternative explanation for mathematic achievement, but it does not relate strongly with sports participation and is, therefore, not as significant of an influence as spatial abilities. Students with a parent in a STEM-related field, especially a girl and with a mother in STEM-related professions, achieve math scores on average compared to peers without parents in STEM. Another background-related variable is a student's socioeconomic status. Students from high socioeconomic statuses, on average, do better than students from low statuses. Both of these explanations, though, are significantly less consistent at improving scores and depend on a variety of factors and influences, with many outliers, while the argument for spatial awareness improving mathematical reasoning is much more sound.

The urban versus rural explanation of some students experiencing increased math scores and some not contains the same issue. Students from urban areas overall show more benefit from sports participation on their academic performance than students from rural areas. However, the findings are again less consistent as they contain many outliers, showing a weaker correlation between location and math skills than spatial awareness and math skills.

The other academic effects from students participating in sports make a slightly stronger case as they link sports and math performance more directly. However, none of them emphasize the actual skills needed to succeed mathematically; rather, they emphasize getting a good grade. Students want to play sports in college, so they try to boost their math grades, not prioritizing their understanding of fundamental mathematical principles. Increased confidence gives students the gumption to work hard, but it does not directly improve math skills. Girls in sports going against the status quo, and therefore, doing it academically too drives them to attempt math more earnestly, but it does not help boost any mathematical abilities. All of these effects of sports participation may improve a student's performance in math, but there is a less direct connection than spatial abilities, which improve the fundamentals needed to succeed mathematically.

While there are a number of factors that influence mathematical ability, spatial awareness skills learned through non-contact do so most effectively as a student-athlete uses their full brain power to practice spatial skills that directly improve a student's understanding of fundamental math principles. Contact sports also directly improve a student's understanding, but to a lesser

extent because part of the athlete's brain is focused on avoiding collision and injury. The other factors that influence a student's mathematical abilities, whether connected to sports participation or not, play a lesser role in math achievement because of their less direct connection with the core skills at the foundation of mathematical achievement.

## **Conclusion**

In the sections above, I have argued that non-contact sports, more significantly than contact sports and beyond other factors such as family background, improve math education because they directly emphasize overlapping skills. Because non-contact sports focus on spatial awareness over potential threats to physical safety, they help students improve core math abilities in spatial orientation, mental rotation, and spatial visualization. This paper makes an original contribution by going beyond analyzing whether or not sports in general affect academics in general. It analyzes the category of sports that benefit academic achievement as well as the underlying reasons behind this improvement.

In the core concepts section, I defined sports, the difference between contact and non-contact sports, the academic skills related to math learning, and spatial awareness abilities that are a byproduct of sports participation. This demonstrated that non-contact sports allow a participant to focus more of their attention on the skills involved in playing a sport and not have to worry about avoiding collision with other athletes. Additionally, it showed that various spatial abilities can be learned from participating in sports and also translated into mathematical abilities. Therefore, the core concepts section proved that participation in sports, specifically non-contact sports, improved mathematical performance because of the spatial skills learned from athletics.

In the other explanations section, I explained how family background, location, college aspirations, confidence, and gender can play a role in mathematical performance despite having no connection with spatial abilities. However, I also showed that despite this reality, there is much more variation in the role these factors play in developing math skills. Therefore, the other explanations section conceded that spatial abilities are not the only factor influencing math scores, but they are a more consistent and directly related factor.

I believe more research should be done where students take benchmark tests on the three spatial awareness skills as well as mathematical ability in the four categories of math learning, then are tasked with playing sports that strongly emphasize one or more of the spatial awareness skills, and finally retake the tests after focusing on that sport for a number of weeks. This type of experiment would more concretely show the link, if any, between sports, spatial skills, and mathematical ability.

## Works Cited

- Campo, M.; Laborde, S.; Mosley, E. Emotional Intelligence Training in Team Sports: The Influence of a Season Long Intervention Program on Trait Emotional Intelligence. *J. Individ. Differ.* 2016, 37, 152–158.
- Castro-Sánchez, Manuel, et al. “Motivation, Anxiety, and Emotional Intelligence Are Associated with the Practice of Contact and Non-Contact Sports: An Explanatory Model.” *Sustainability*, vol. 11, no. 16, 2019, pp. 1-17.
- Chikish, Yulia, et al. “eSports: A New Era For The Sports Industry And A New Impulse To Research In Sports (And) Economics?” *FUNCAS Social and Economic Studies*, vol. 7, pp. 477-508.
- Darling, N., Caldwell, L. L., & Smith, R. (2005). Participation in school-based extracurricular activities and adolescent adjustment. *Journal of Leisure Research*, 37(1), 51–56.
- DeMeulenaere, Eric. “Playing The Game: Sports As A Force For Promoting Improved Academic Performance For Urban Youth.” *Journal of Cultural Diversity*, vol. 17, no. 4, 2010, pp. 127-135.
- Din, Feng S. “Sport Activities Versus Academic Achievement For Rural High School Students.” *National Forum of Applied Educational Research Journal-Electronic*, vol. 19, no. 3E, 2005-2006, pp. 1-11.
- Gentile, A. M. (2000). Skill acquisition: Action, Movement, and neuromotor processes. In J. H. Carr, R. B. Shepherd, J. Gordon, A. M. Gentile, & J. M. Held (Eds.), *Movement science: Foundations for physical therapy in rehabilitation* (pp 111–180). Gaithersburg, MD: Aspen.
- Giza, C.C.; Prins, M.L.; Hovda, D.A. It’s not all fun and games: Sports, concussions, and neuroscience. *Neuron* **2017**, 94, 1051–1055.
- Jansen, Petra, and Stefanie Pietsch. “Sports and mathematical abilities in primary school-aged children: How important are spatial abilities? An explorative study.” *Current Psychology*, vol. 41, 2020, pp. 7132-7141.
- Jenny, S.E., Manning, R.D., Keiper, M. C. and Olrich, T. W. (2017). Virtual(ly) Athletes: Where eSports Fit Within the Definition of “Sport”. *Quest*, 69(1), pp. 1-18.
- Kang, Jin Oh, et al. “Comparison of Psychological and Cognitive Characteristics between Professional Internet Game Players and Professional Baseball Players.” *International Journal of Environmental Research and Public Health*, vol. 17, no. 13, 2020, pp. 1-12.
- Lo, Chueseng. “Defining Esports Student-Athletes and the Behaviors that Affect Academic Performance.” *University of Wisconsin-Stout Journal of Student Research*, vol. 20, 2022, pp. 31-42.
- Lubans, David R., et al. “School Physical Activity Intervention Effect on Adolescents’ Performance in Mathematics.” *Medicine & Science in Sports & Exercise*, vol. 50, no. 12, 2018, pp. 2442-2450.
- Mahoney, J. L., & Cairns, R. B. (1997). Do extracurricular activities protect against early school dropout? *Developmental Psychology*, 33, 241-253.

- Mix, K. S., & Cheng, Y.-L. (2012). The relation between space and math: Developmental and educational implications. *Advances in Child Development Behavior*, 42, 197–243. <https://doi.org/10.1016/B978-0-12-394388-0.00006-X>.
- Pietsch, S., & Jansen, P. (2012). Different mental rotation performance in students of music, sport and education science. *Learning and Individual Differences*, 22(1), 159–163. <https://doi.org/10.1016/j.lindif.2011.11.0127>.
- Reardon, S. F., Fahle, E. M., Kalogrides, D., Podolsky, A., & Zarate, R. C. (2018). Gender Achievement Gaps in U.S. School Districts. Center for Education Policy Analysis, Stanford University.
- Rees, Daniel I., and Joseph J. Sabia. “Sports participation and academic performance: Evidence from the National Longitudinal Study of Adolescent Health.” *Economics of Education Review*, vol. 29, no. 5, 2010, pp. 751-759.
- Rose, Giuseppe, et al. “Train to Reason: Sports and the Girls’ Math-Reading Gap.” *Creative Education*, vol. 13, no. 7, 2022, pp. 2354-2384.
- Seises, Kristina L. “Examining racial differences in the effect of popular sports participation on academic achievement.” *Social Science Research*, vol. 40, no. 4, 2011, pp. 1142-1169.
- Trudeau, François, and Roy J. Shephard. “Physical education, school physical activity, school sports and academic performance.” *International Journal of Behavioral Nutrition and Physical Activity*, vol. 5, 2008, pp. 1-12.
- Veliz, Philip, and Sohaila Shakib. “Gender, Academics, and Interscholastic Sports Participation at the School Level: A Gender-specific Analysis of the Relationship between Interscholastic Sports Participation and AP Enrollment.” *Sociological Focus*, vol. 47, no. 2, 2014, pp. 101-120.
- Xie, F., Zhang, L., Chen, X., & Xin, Z. (2019). Is spatial ability related to mathematical ability: A meta-analysis. *Educational Psychology Review*. published ahead of print, 32, 113–155. <https://doi.org/10.1007/s10648-019-09496-y>.

# **Sustainability of Bioplastics in Aquatic Environments: A Comparative Analysis into the Biodegradability and Environmental Impacts of PLA, PHA, Starch-based, and Cellulose-based Bioplastics**

By Sean Tey Yi-Shuen

## **Abstract**

In the 21<sup>st</sup> century, sustainability is an omnipresent trend that industries are turning to, and bioplastics have been touted as biodegradable and environmentally safe options to traditional plastics. Examples of these bioplastics include polylactic acid (PLA), polyhydroxyalkanoates (PHA), cellulose-based and starch-based bioplastics. This paper investigates the environmental impacts of these bioplastics in an oceanic environment, as aquatic plastic pollution in the form of oceanic garbage patches is one of the largest problems surrounding traditional plastics. Using literature data from scientific journals, an investigation was conducted specifically into the biodegradability of bioplastics in an aquatic environment and the environmental impact of these bioplastics being in these aquatic environments. It was concluded that in terms of biodegradation, PHA was the best, followed by cellulose-based and PLA which was non-biodegradable, while there was insufficient data for starch-based bioplastics. All four bioplastics also showed levels of baseline toxicity that were harmful to humans at certain concentrations. It was evaluated that there was insufficient data to form concrete conclusions; further experiments were recommended to be conducted *in situ* for greater replicability of aquatic environments.

## **Introduction**

In the 21<sup>st</sup> century, plastics are ubiquitous as they are the cornerstone of the industrial sectors, with varying applications in sectors such as pharmaceutical instruments, packaging, household applications and automotives [1].

According to [2], the problems with plastics can be categorized into four largely identifiable environmental problems – the landfill problem, oceanic garbage patches, plastic incineration and non-biodegradable plastics. Thankfully, in a world where sustainability is becoming increasingly prevalent, there has been a significant push for bioplastics to be proposed as an alternative to petroleum-based plastics in industrial use [2]. According to McGill University's definition of sustainability, which defines sustainability to be: "The ability to meet the needs of the present without compromising the ability of future generations to meet their own needs." [3] The use of a renewable replacement to conventional plastics, such as bioplastics, aligns with this principle by reducing dependency on finite fossil resources.

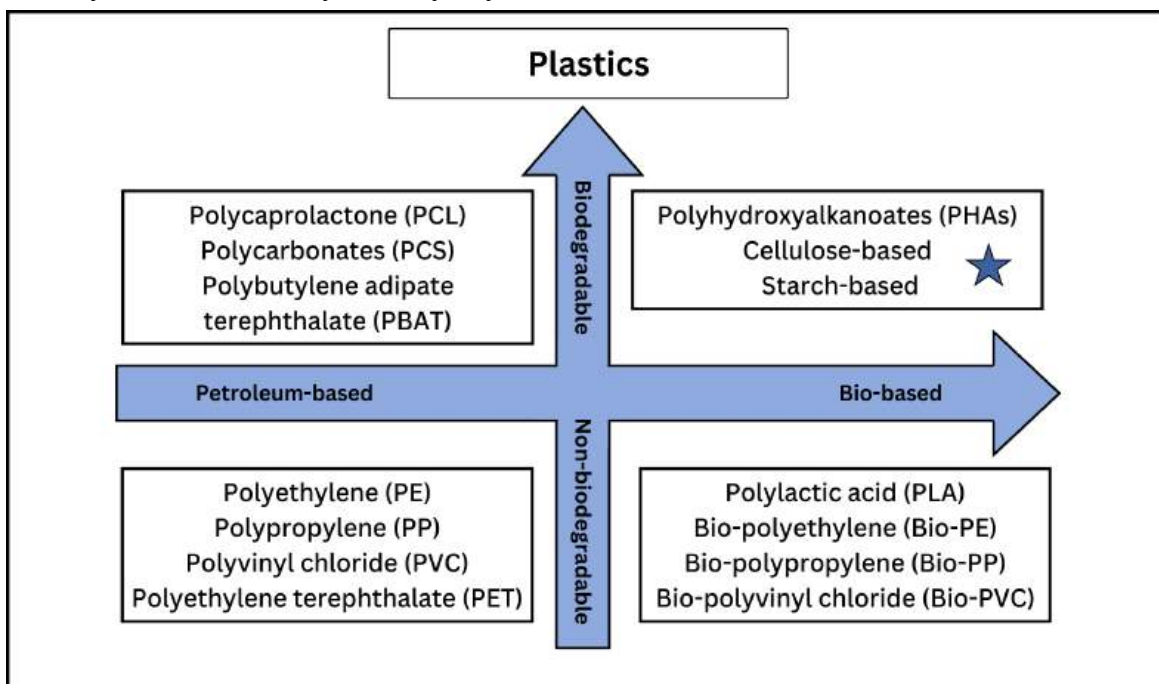
This paper will focus particularly on the problem of oceanic garbage patches. It is the opinion of the paper that the landfill problem and plastic incineration are largely social issues that cannot be easily mitigated through the adoption of bioplastics. Hence, out of the four environmental problems listed above, efforts should be best spent on analysing whether bioplastics are suited to solve the issue of oceanic pollution, as well as biodegradability in aquatic environments. The bioplastics to be analysed are chosen based on popularity, as plant-based polymers (like cellulose), thermoplastic starch, polyhydroxyalkanoates (PHAs), and

polylactic acid (PLA) are the most reported amongst bioplastics [4].

### What are Bioplastics?

As bioplastics are often misunderstood by the general public, in this paper, bioplastics are defined according to the International Union of Pure and Applied Chemistry (IUPAC)'s definition of bioplastics: "bio-based polymers shaped by flow as they are processed, that are either made from biomass directly or monomers which themselves are derived from a biomass source." [5]

In *Fig 1*, all bioplastics are bio-based and can be further categorized as being biodegradable or non-biodegradable. Thus, it is important to note that bioplastics are not necessarily "environmentally friendly" by definition.



*Fig 1: Dichotomy of Plastics Mindmap [6]  
(Blue star represents the category of bioplastics discussed in this paper)*

Generally, we define biodegradability to mean the complete breakdown of materials through biological activity, such as via the action of microorganisms like bacteria, fungi or algae [7, 8]. Biodegradation can take place under aerobic conditions (requiring O<sub>2</sub>) to form biomass, CO<sub>2</sub> and H<sub>2</sub>O, or under anaerobic conditions (absence of O<sub>2</sub>) to form biomass, CO<sub>2</sub>, CH<sub>4</sub> and H<sub>2</sub>O [7, 9].

It is important to note that for plastic to biodegrade within a desirable timeframe, there are specific conditions that must be present, such as light level, temperature, oxygen and humidity levels. However, these specific conditions are impossible to be replicated in natural environments due to the array of uncontrollable physical and biological factors that change locally. Many materials that are touted as "biodegradability" are in truth only biodegradable



when tested in special composting facilities that have the most optimal conditions – which are better described as “compostable”. As suggested by [9], biodegradability should instead be viewed as a “system property”, by considering whether a material can biodegrade depending solely on the interaction between the material and its surrounding environment. Different environments will have different potential for biodegrading various types of plastics; we must standardise an environment for evaluation – in our case being the aquatic environment.

### Types of Bioplastics

Fig 2 illustrates the distribution of global bioplastic production in 2024, and below is a brief explanation of the production of the different types of bioplastics in this paper.

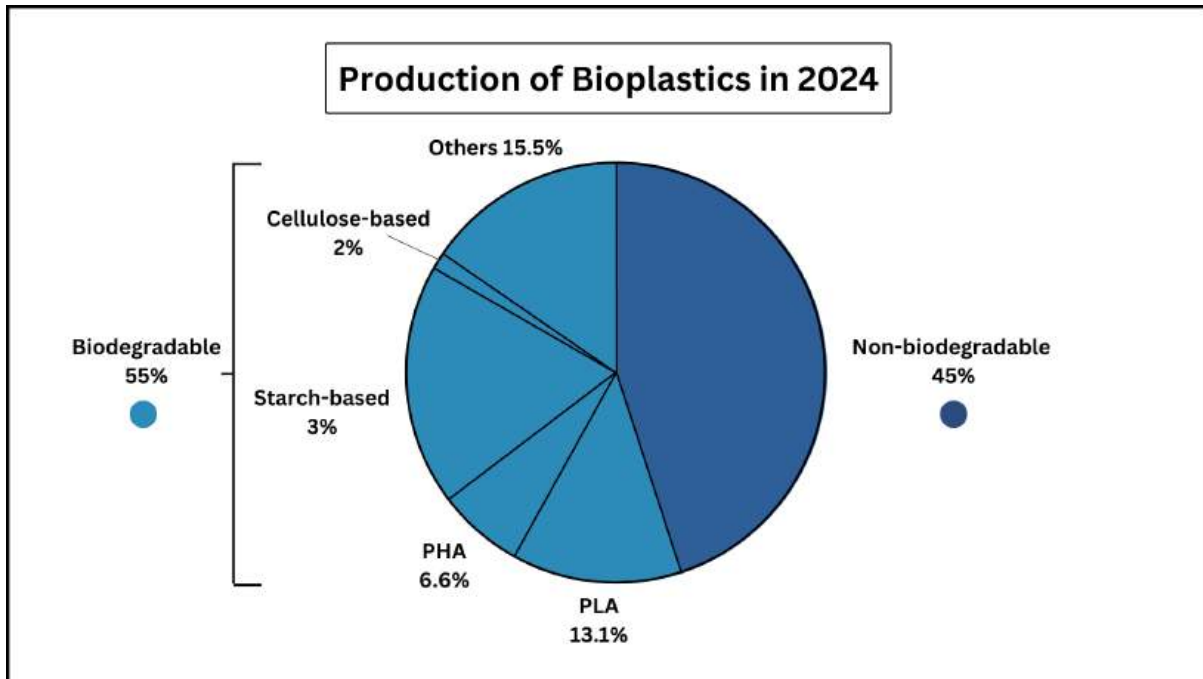


Fig 2: Distribution of global production capacities of bioplastics in 2024 [10]

#### 1. PLA-based Bioplastics

Poly(lactic acid) (PLA) is a commercial bioplastic synthesized mainly from natural sources such as corn starch, sugarcane or cassava roots, that represents 13.1% of global bioplastic production in 2024 [10,11]. The biomass first undergoes microbial fermentation to create lactic acid with two possible enantiomers [11]. By undergoing polycondensation, water molecules are released to form short chains of lactic acid known as oligolactic acid. These chains are further broken down via depolymerisation into PLA monomers of varying spatial configuration [11]. Finally, these monomers undergo a ring-opening reaction that “opens” the cyclic structure, allowing monomers to form PLA of varying spatial configurations [11].

#### 2. PHA-based Bioplastics

Polyhydroxyalkanoates (PHAs) are a family of bioplastics synthesised in various bacteria cell walls through microbial fermentation, with a 6.6% representation in 2024 global bioplastic production [10]. This occurs as a way of storing carbon in the form of PHAs as an energy source as a survival response in nutrient-scarce environments [4, 12, 13]. Once fermentation is complete, the PHAs are extracted by breaking down the cells to create PHAs of different lengths, with varying properties [4, 12].

### 3. Cellulose-based Bioplastics

Cellulose is a naturally abundant polymer produced by a variety of biomass (e.g. wood, fibres, grass, marine animals, fungi and bacteria [4]) in their cell walls. Cellulose occupies 1.3% of global bioplastic production [10]. Cellulose is a natural homopolymer composed of thousands of glucose monomers, which are bonded in a way that results in a rigid and straight polymer chain that contribute to the polymer's properties [4, 12]. Through complex processes like esterification and etherification, cellulose is converted into water-soluble semi-synthetic polymer [14]. The efficacy and properties of cellulose-based polymers are often improved using the addition of additives like plasticizers, like elongation and their moulding capabilities [14].

### 4. Starch-based Bioplastics

Starch represents about 18.5% of global bioplastic production [10] and is largely produced by food crops such as corn, cassava, wheat, rice or potatoes [12, 15] and is synthesised via photosynthesis [12]. Starch is composed mainly of two polymers: amylose and amylopectin, which contributes elasticity and tensile strength respectively [11]. However, natural starch is limiting, as it is hydrophilic and has strong bonds that makes it difficult to process [12]. Hence, starch is often mixed with plasticizers such as glycerol or glycerine.

## **Oceanic Garbage Patches**

Oceanic garbage patches are undoubtedly caused by marine plastic pollution, given 80% of all oceanic pollution is plastic based [16]. Most consumer plastics are petroleum-based, such as polypropylene (PP) or polyethylene (PE), and are hence designed to be single-use, while waste management systems are ill-equipped for recycling [17]. Hence, these petroleum-based plastics take millennia to naturally biodegrade due to their chemical structures naturally consisting of long and stable polymer chains, [2, 18] which only break down in the presence of UV radiation and heat [19]. This is particularly exacerbated in aquatic environments, as the extremely low temperatures in the ocean and minimal UV exposure slow the degradation of plastics when submerged [19]. Ultimately, this results in an estimated eight million metric tons of macroplastics and 1.5 million metric tons of microplastics entering the oceans yearly, with projections of pollutions to double by year 2050 [16, 17], threatening the marine ecosystems in the water through ingestion, entanglement and habitat destruction [20]. This paper will use the following metrics to compare each bioplastic's sustainability in an aquatic environment:

1. Biodegradability of bioplastics in an aquatic environment
2. Environmental Impact of bioplastics in an aquatic environment

Compared to the more widely investigated land-based environments, there have not been many studies conducted specifically into the biodegradation of bioplastics in aquatic environments [22]. A possible reason for this is due to the multiple uncontrollable factors that affect marine degradation of bioplastics. Firstly, biotic factors such as the composition and density of microbial communities vary depending on geographical location and water depth – certain species of bacteria biodegrade different materials at different rates [4, 21]. Additionally, there are abiotic factors, such as salinity, temperature and UV light which affect temperature, that can depend on different liquid environments (e.g. pelagic, eulittoral, sandy sublittoral, sandy supralittoral, seabed, sediments) which depends on geography [21, 22].

### Metric 1: Biodegradability of Bioplastics in aquatic environments

There are three methods used to determine biodegradation, which have been summarised below in *Table 1*.

Method Name	Explanation	Utilisation
Mass Loss	Any significant decrease in mass suggests bioplastic biodegradation, calculated as $Biodegradation = \frac{M_{final} - M_{initial}}{M_{initial}} \cdot 100\%$	Gel Permeation Chromatography (GPC) separates molecules by size using specific pore sizes as a filter. After degradation, polymers degrade, leading to decrease in molecule weight distribution indicating mass loss [23].
Evolution of CO <sub>2</sub> / CH <sub>4</sub>	CO <sub>2</sub> is produced in aerobic biodegradation and CH <sub>4</sub> is produced in anaerobic biodegradation, which can be used to track how organic material is converted to carbon.	Produced CO <sub>2</sub> is trapped by KOH solution during the test.
Biological Oxygen Demand (BOD)	As microorganisms consume oxygen in the water when breaking down bioplastics, BOD measures the amount of biodegradable matter in the water.	The WTW OxiTop system records pressure difference over time from the absorption of CO <sub>2</sub> evolved, which is proportional to BOD [21].

*Table 1 – Summary of methods to determine biodegradability*

Table 2 summarises experiments that have used these methods to calculate the biodegradability of bioplastics in aquatic environments.

Type of Bioplastic	Material Tested	Biodegradation (%)	Experimental Duration (days)	Aquatic environment	Experimental Conditions	Experimental Scale	Biodegradation Indicator	Source
PHAs	Mirel PHA - 2200	38.22	180	<i>Half Moon Bay, California</i>	1g Sample 30.C 400ml Ocean Water 100g Ocean Soil Sample	Laboratory	CO <sub>2</sub> Release	[24]
PHAs	Mirel PHA - 4100	45.08						
PLA	NatureWorks PLA bottle	4.45						
PLA	NatureWorks PLA Frito Lay bag	3.11						
Cellulose	Avicell Microcellulose powder	33.00						
PLA	PLA Films (320 ± 20 µm thickness)	< 1.00 (Freshwater & Seawater)	365	-	25.C 16 h light and 8 h dark Fluorescence Light	Laboratory (Thermostatic Chamber)	Mass Loss (Gel Permeation Chromatography)	[23]
PHAs	PHB Films (320 ± 20 µm thickness)	8.50						
PLA	ME331050/1 poly-L-lactic acid (PLLA) (0.05mm thickness)	0.00	28	Brackish  <i>Northern Baltic Sea</i>  <i>coastal location off the southwest coast of Finland</i>	15.C Dark conditions	Laboratory	BOD	[21]
PHAs	BV301025 (PHB/HV) (0.025 mm thickness)	70.00						
Cellulose	AC311051 cellulose acetate (CA) (0.05 mm thickness)	15.00						
PLA	-	0.00	28	<i>La Jolla, California</i>	30.C Conducted at Sea Surface and Seafloor	<i>In situ</i>	CO <sub>2</sub> Release	[9]
Cellulose	1. Lyocell (CLY) 2. Modal (CMD) 3. Viscose (CV) 4. Organic Virgin Cotton (OCO)	CLY - 76.00 CMD - 81.40 CV - 82.20 OCO - 81.20						
Cellulose	Cellulose Films	39.00	30	<i>Malaga (Spain)</i>	200 mg sample 164 ml seawater 20.C	Laboratory	BOD	[25]

	(70 $\mu\text{m}$ thickness)							
PHAs	PHB Films	9.00	70	<i>Pacific Ocean</i>		Laboratory	Gravimetric Analysis	[26]
Starch	Mater-Bi plastic bags	1.50	90	Freshwater Marsh & Seawater	25.C	Laboratory / <i>In situ</i>	Mass Loss	[27]
Starch	Mater-Bi partially bio-based copolyester and corn starch	68.90	236	Simulated Sandy Sublittoral Zone Marine Water with Sediment	Room Temp	Laboratory	BOD	[28]
Starch	Cassava starch/glycerol (3/1 mixture)	11.46	10	-	30.C Dip-hanging method, Aerobic Biodegradation,	Laboratory	Mass Loss	[29]

*Table 2 - Table of studies carried out on bioplastics biodegradation in aquatic environments*

Generally, *Table 2* shows that PLA-based bioplastics consistently have the lowest biodegradability rates among the four bioplastics. [9, 21, 23] all report less than 1% biodegradation, despite the experiments being conducted at different sea levels (sea surface and sea floor), different environments (freshwater and seawater), light presence and temperature. This can be attributed to the specific conditions that PLA requires for biodegradation – temperatures of 140.C, and exposure to special digestive microbes – hence, requiring a special composting facility [11, 30]. These findings are consistent with Raman Spectroscopy results from study [9], in which PLA’s Raman spectra were unchanged over 400+ days, indicating no degradation [31]. This study concluded that the ocean’s low temperature and low bacterial density makes it unsuitable for the biodegradation of PLA.

Both PHA and PHB studies were included in *Table 2*. [24] tested PHA of differing crystallinity, yet both having substantial degradation, that passes the ASTM D-7081 criteria of 30% aerobic degradation [24]. However, PHB results are more varied with [23] and [26] showing low biodegradability (~9%) yet [21] has a large 70% biodegradability. While there is no definite answer, this result might be a result of the geography, as [21] was conducted in the brackish Baltic Sea, in which microbial communities are a combination of marine and freshwater species and hence not directly comparable to marine bacterial communities [21].

*Table 2* showed that starch-based bioplastics show high variability in biodegradability, as [27] reports only a 1.5% degradation in 90 days under both freshwater & seawater environments at 25.C while [28] records a 69% degradation within 236 days at similar temperature [27, 28]. These discrepancies are likely due to the bioplastic blend. As mentioned earlier, starch-based bioplastics are often mixed with plasticizer, and [28] is mixed with partially bio-based copolyester, which are renewable monomers made from waste vegetable oil and corn starch [28]. The sample in [27] is mainly composed of starch with a minor fraction of natural plasticizer [27].

Cellulose-based bioplastics are shown to have high variability, but high consistency with different types of bioplastics. [9] tested with four different types of cellulose-based bioplastics, achieving >75% biodegradability across the board, at 30°C both at sea surface and seafloor. This study explored the differences in biodegradability between natural and regenerated cellulose-based bioplastics. Raman spectroscopy results show that both natural and regenerated cellulose exhibit degradation signs in the aquatic environment, but the results show a higher degradation rate in the regenerated variant [9]. Regenerated cellulose fibre have similar chemical structure but are processed and have lower crystallinity (degree of structural order) [9]. It is reasoned that lower crystallinity results in greater ease of moisture regain which contributes to faster biodegradation [9]. This trend is repeatedly shown in land-based biodegradation experiments conducted in [9] too. Using data from [21-24] which compare two or more bioplastics under the same conditions, the general trend of biodegradation is likely to be PHA being the best followed by cellulose-based and PLA being the worst. Starch is more biodegradable than PLA but not enough data is available to rank.

## **Metric 2: Environmental Impact of bioplastics in aquatic environments**

PLA's environmental impact is the most well studied and shows that PLA has prominent negative effects on marine life in environments. Many researchers report that PLA exhibits a low or negligible biodegradation rate in aquatic environments, even under various conditions [9, 11, 30]. Hence, most PLA that end up in the water undergoes mechanical weathering and fragments into small pieces of less than 5 mm in size termed as microplastics [11]. These microplastics are known to accumulate in phytoplankton and zooplankton, the primary producers at the bottom of the marine food chain. According to researchers, microplastics resemble the prey of zooplankton and consumption can have negative impacts [11]. A study concluded that PLA microplastics causes the mortality of the *Daphnia magna* species of zooplankton to rise from 5% to 60% after exposure to just 500 mg/l of microplastic, as well as decreased reproduction rates and body length [11, 33]. Additionally, another study found that PLA was found in the gelatinous tissue of *Aurelia* species of jellyfish which was tied to a change in swimming behaviour due to direct toxicity from microplastics [11, 32].

Zooplankton's role as primary producers at a base of a food chain means that any impact on them has the largest compounding domino effect across all trophic levels. With population sizes decreasing due to increased mortality and decreasing reproduction, as per study [33], this might lead to a decrease in food availability higher up in the food chain, destabilising an entire ecosystem. More concerningly, from the result of [32], consumed zooplankton may result in bioaccumulation – when the intake of a contaminant is greater than the ability of the organism to purge it – which creates an increase in concentration of contaminants at higher trophic levels, known as biomagnification [34]. The problem can compound till it reaches the apex predators – humans who consume seafood, posing severe health risks ranging from oxidative stress, cytotoxicity to even necrosis and impairment of immune systems [35].

Some bioplastics are known to include additives in their plastics, such as plasticizers, antioxidants and stabilisers [36]. Plasticizers are used to make bioplastics flexible; antioxidants prevent oxidation, and stabilisers enhance plastic's durability against UV radiation or heat. These additives are most prominent in cellulose and starch-based bioplastics, given their limited physical properties [36]. A study by [36] analysed samples of bioplastic products via infrared spectroscopy to analyse the toxicity of 4 products. By using a bacterium that emits light and changes luminosity in the presence of toxicity, researchers recorded the concentration of plastic needed to inhibit 20% inhibition of luminescence, the criteria met to induce a general toxic effect on living organisms [36]. Out of the 43 varied types of daily plastic products analysed, ranging from plastic cups to bottles and food tray, 67% of the plastics indicated toxicity, particularly cellulose-based, starch-based and PHA plastics. These plastics had a low threshold for the criteria, meaning that a lower concentration of these plastics can induce toxicity. This study concludes that observed toxicities are primarily due to chemicals added, rather than microplastics [37].

However, cellulose-based bioplastics do have some environmental upside depending on its source material. As cellulose is one of the most naturally abundant materials in the world, found in a variety of biomass, research has been conducted into using environmentally damaging biomass/organisms to create it. For example, *Cladophora sp. algae* is a type of algae known for uncontrollable growth in freshwater environments [37]. This excessive growth forms dense mats on the water's surface which blocks sunlight and reduces available oxygen to aquatic life, harming aquatic environments [37]. Research has shown that the algae contain 20-30% cellulose with up to 95% crystallinity, making it a good candidate for a bioplastic with strong mechanical properties (e.g. strength and durability) [37]. While more research must be done, using environmentally harmful biomass to create biodegradable bioplastics to replace environmentally harmful plastics is a promising solution that kills two birds with one stone.

## **Conclusion**

It is clear that although bioplastics are often marketed as being the “biodegradable option” or the “most environmentally safe”, neither slogan is accurate. This paper investigated scientific publications on four of the most common bioplastics – PLA, PHA, cellulose-based, starch-based. Our investigation has revealed that in the oceanic environment, not all bioplastics have a 100% biodegradability rate, especially PLA which has negligible biodegradation in the oceans. Using these studies, we conclude that in terms of oceanic biodegradability, PLA is not biodegradable, PHA is the best followed by cellulose-based, and there is not enough information available to evaluate starch-based bioplastics. This paper also concludes that bioplastics all have substantial risk to human health and are not always environmentally safe due to the toxicity of chemicals included in additives in bioplastics.

However, this investigation into the first metric of comparison had major limitations. Firstly, there was a scarcity of experimental data to review due to the sheer difficulty in conducting aquatic-based experiments. Hence, many experiments default to using experimental setups in laboratories that simulate environmental conditions. The difficulty is that these

simulations are not entirely accurate, as a natural environment like the ocean is too complex to be simulated accurately given the vast numbers of biotic factors in a natural environment that must be accounted for, such as: temperature, location in water column, presence of UV light, different native microbial species among others.

Secondly, even among these simulated experiments, it is difficult to draw meaningful comparisons about the bioplastics' biodegradability, due to the inconsistent experimental conditions. Each experiment has its own duration, experimental setup, method to track biodegradation and aquatic environment (each with their own biotic factors). To circumvent this, we prioritised the inclusion of studies/data that test multiple types of bioplastics so that we obtain results from multiple types of bioplastics under the same conditions for comparison. Unfortunately, there were no experiments found that included all four bioplastics discussed in this study.

Ultimately, there is no doubt that bioplastics are revolutionary in the fight for environmental sustainability. In fact, reports show that manufacturers are already meeting sustainability goals and that bioplastics have the potential to reduce CO<sub>2</sub> emission by a minimum of 30%, lower carbon footprint by 42% and consume 65% less energy than the production of traditional plastics [38]. However, these flashy statistics and false advertisements about bioplastics overshadows their negative impacts on the environment. Further research should aim to be conducted *in situ* instead of relying on lab simulations, as well as standardising the same experimental conditions across the board. It is only with these changes will scientists be able to collaborate to determine the true environmental impact of bioplastics in aquatic environments.



## Works Cited

- [1] Lim, C., et al. "Bioplastic Made from Seaweed Polysaccharides with Green Production Methods." *Journal of Environmental Chemical Engineering*, vol. 9, no. 5, Oct. 2021, p. 105895, <https://doi.org/10.1016/j.jece.2021.105895>. ScienceDirect. Accessed 28 Sept. 2024.
- [2] Bezirhan Arikian, Ezgi, and Havva Duygu Ozsoy. "A Review: Investigation of Bioplastics." *Journal of Civil Engineering and Architecture*, vol. 9, no. 2, Feb. 2015, pp. 188–92, <https://doi.org/10.17265/1934-7359/2015.02.007>. David Publishing. Accessed 28 Sept. 2024.
- [3] McGill University. "What Is Sustainability?" *McGill*, University of Alberta, 2013, [www.mcgill.ca/sustainability/files/sustainability/what-is-sustainability.pdf](http://www.mcgill.ca/sustainability/files/sustainability/what-is-sustainability.pdf). Accessed 10 Aug. 2024.
- [4] Ahsan, Wazir Aitizaz, et al. "Biodegradation of Different Types of Bioplastics through Composting—a Recent Trend in Green Recycling." *Catalysts*, vol. 13, no. 2, Jan. 2023, p. 294, <https://doi.org/10.3390/catal13020294>. MDPI. Accessed 28 Sept. 2024.
- [5] Samsalee, Namfon, et al. "Rice Husk Nanocellulose: Extraction by High-Pressure Homogenization, Chemical Treatments and Characterization." *Carbohydrate Polymer Technologies and Applications*, vol. 6, Dec. 2023, p. 100353, <https://doi.org/10.1016/j.carpta.2023.100353>. Elsevier. Accessed 28 Sept. 2024.
- [6] The Sustainable Alternative. "Sustainable Alternative to PVC - the Sustainable Alternative." *Sustainable Alternative to PVC*, 26 Jan. 2023, [sustainable-alternative.com/sustainable-alternative-to-pvc/](http://sustainable-alternative.com/sustainable-alternative-to-pvc/). Accessed 28 Sept. 2024.
- [7] Shafqat, Arifa, et al. "Synthesis and Characterization of Starch Based Bioplastics Using Varying Plant-Based Ingredients, Plasticizers and Natural Fillers." *Saudi Journal of Biological Sciences*, vol. 28, no. 3, Mar. 2021, pp. 1739–1749, [www.sciencedirect.com/science/article/pii/S1319562X20306707](http://www.sciencedirect.com/science/article/pii/S1319562X20306707), <https://doi.org/10.1016/j.sjbs.2020.12.015>. Accessed 28 Sept. 2024.
- [8] Gadhave, Ravindra V., et al. "Starch Based Bio-Plastics: The Future of Sustainable Packaging." *Open Journal of Polymer Chemistry*, vol. 08, no. 02, 2018, pp. 21–33, <https://doi.org/10.4236/ojpcem.2018.82003>. Accessed 28 Sept. 2024.
- [9] Royer, Sarah-Jeanne, et al. "Not so Biodegradable: Polylactic Acid and Cellulose/Plastic Blend Textiles Lack Fast Biodegradation in Marine Waters." *PLOS ONE*, vol. 18, no. 5, May 2023, pp. e0284681–81, <https://doi.org/10.1371/journal.pone.0284681>. PubMed. Accessed 28 Sept. 2024.
- [10] Bioplastics, E. "Bioplastics market development update 2023." European Bioplastics eV: Berlin, Germany (2023).
- [11] Ali, Wajid, et al. "Polylactic Acid Synthesis, Biodegradability, Conversion to Microplastics and Toxicity: A Review." *Environmental Chemistry Letters*, vol. 21, Jan. 2023, pp. 1761–86, <https://doi.org/10.1007/s10311-023-01564-8>. Springer Link. Accessed 28 Sept. 2024.

- [12] Narancic, Tanja, et al. “Recent Advances in Bioplastics: Application and Biodegradation.” *Polymers*, vol. 12, no. 4, Apr. 2020, p. 920, <https://doi.org/10.3390/polym12040920>. National Library of Medicine. Accessed 28 Sept. 2024.
- [13] Surendran, Arthy, et al. “Can Polyhydroxyalkanoates Be Produced Efficiently from Waste Plant and Animal Oils?” *Frontiers in Bioengineering and Biotechnology*, vol. 8, no. 169, Mar. 2020, <https://doi.org/10.3389/fbioe.2020.00169>. National Library of Medicine.
- [14] Sirviö, Juho Antti, et al. “Effect of Plasticizers on the Mechanical and Thermomechanical Properties of Cellulose-Based Biocomposite Films.” *Industrial Crops and Products*, vol. 122, 15 Oct. 2018, pp. 513–521. Elsevier, [www.sciencedirect.com/science/article/pii/S0926669018305491](http://www.sciencedirect.com/science/article/pii/S0926669018305491), <https://doi.org/10.1016/j.indcrop.2018.06.039>. Accessed 28 Sept. 2024.
- [15] Nanda, Sonil, et al. “Innovations in Applications and Prospects of Bioplastics and Biopolymers: A Review.” *Environmental Chemistry Letters*, vol. 20, no. 1, Nov. 2021, pp. 379–95, <https://doi.org/10.1007/s10311-021-01334-4>. Springer Link. Accessed 28 Sept. 2024.
- [16] Fava, Marta. “Plastic Pollution in the Ocean: Data, Facts, Consequences.” *Ocean Literacy Portal*, 9 May 2022, [oceanliteracy.unesco.org/plastic-pollution-ocean/#:~:text=Plastic%20waste%20makes%20up%2080](http://oceanliteracy.unesco.org/plastic-pollution-ocean/#:~:text=Plastic%20waste%20makes%20up%2080). Accessed 28 Sept. 2024.
- [17] Lau, Winnie W. Y., et al. “Evaluating Scenarios toward Zero Plastic Pollution.” *Science*, vol. 369, no. 6510, July 2020, pp. 1455–61, <https://doi.org/10.1126/science.aba9475>. Science. Accessed 28 Sept. 2024.
- [18] Zhao, Xianhui, et al. “Sustainable Bioplastics Derived from Renewable Natural Resources for Food Packaging.” *Matter*, vol. 6, no. 1, Jan. 2023, pp. 97–127, <https://doi.org/10.1016/j.matt.2022.11.006>. ScienceDirect. Accessed 28 Sept. 2024.
- [19] Andrady, Anthony L. “Persistence of Plastic Litter in the Oceans.” *Marine Anthropogenic Litter*, edited by Melanie Bergmann et al., Springer International Publishing, 2015, pp. 57–72, [https://doi.org/10.1007/9783319165103\\_3](https://doi.org/10.1007/9783319165103_3). Accessed 28 Sept. 2024.
- [20] Dilkes-Hoffman, Leela Sarena, et al. “The Rate of Biodegradation of PHA Bioplastics in the Marine Environment: A Metastudy.” *Marine Pollution Bulletin*, vol. 142, 2019, pp. 15–24, <https://doi.org/10.1016/j.marpolbul.2019.03.020>. ScienceDirect. Accessed 28 Sept. 2024.
- [21] L. Eronen-Rasimus, Eeva, et al. “Degradation Rates and Bacterial Community Compositions Vary among Commonly Used Bioplastic Materials in a Brackish Marine Environment.” *Environmental Science Technology*, vol. 56, no. 22, American Chemical Society, 2022, pp. 15760–69, <https://doi.org/10.1021/acs.est.2c06280>. Accessed 28 Sept. 2024.
- [22] Folino, Adele, et al. “Biodegradation of Wasted Bioplastics in Natural and Industrial Environments: A Review.” *Sustainability*, vol. 12, no. 15, July 2020, p. 6030, <https://doi.org/10.3390/su12156030>. MDPI. Accessed 28 Sept. 2024.

- [23] Bagheri, Amir Reza, et al. "Fate of So-Called Biodegradable Polymers in Seawater and Freshwater." *Global Challenges*, vol. 1, no. 4, June 2017, <https://doi.org/10.1002/gch2.201700048>. Wiley. Accessed 28 Sept. 2024.
- [24] Greene, Joseph. *Marine Biodegradation of PLA, PHA, and Bio-Additive Polyethylene Based on ASTM D7081*. [www.academia.edu/9639333/Marine\\_Biodegradation\\_of\\_PLA\\_PHA\\_and\\_Bio\\_additive\\_Polyethylene\\_Based\\_on\\_ASTM\\_D7081](http://www.academia.edu/9639333/Marine_Biodegradation_of_PLA_PHA_and_Bio_additive_Polyethylene_Based_on_ASTM_D7081). Accessed 28 Sept. 2024.
- [25] Guzman-Puyol, Susana, et al. "Transparent, UV-Blocking, and High Barrier Cellulose-Based Bioplastics with Naringin as Active Food Packaging Materials." *International Journal of Biological Macromolecules*, vol. 209, no. B, June 2022, pp. 1985–94, <https://doi.org/10.1016/j.ijbiomac.2022.04.177>. Elsevier. Accessed 28 Sept. 2024.
- [26] Tsuji, Hideto, and Kaori Suzuyoshi. "Environmental Degradation of Biodegradable Polyesters 1. Poly( $\epsilon$ -Caprolactone), Poly[(R)-3-Hydroxybutyrate], and Poly(L-Lactide) Films in Controlled Static Seawater." *Polymer Degradation and Stability*, vol. 75, no. 2, Jan. 2002, pp. 347–55, [https://doi.org/10.1016/s0141-3910\(01\)00240-3](https://doi.org/10.1016/s0141-3910(01)00240-3). Elsevier. Accessed 28 Sept. 2024.
- [27] Accinelli, Cesare, et al. "Deterioration of Bioplastic Carrier Bags in the Environment and Assessment of a New Recycling Alternative." *Chemosphere*, vol. 89, no. 2, Sept. 2012, pp. 136–43, <https://doi.org/10.1016/j.chemosphere.2012.05.028>. ScienceDirect. Accessed 12 Dec. 2019.
- [28] Tosin, Maurizio, et al. "Laboratory Test Methods to Determine the Degradation of Plastics in Marine Environmental Conditions." *Frontiers in Microbiology*, vol. 3, 2012, <https://doi.org/10.3389/fmicb.2012.00225>. Accessed 9 Nov. 2019.
- [29] Nissa, R. C., et al. "Preliminary Study of Biodegradability of Starch-Based Bioplastics Using ASTM G21-70, Dip-Hanging, and Soil Burial Test Methods." *IOP Conference Series: Earth and Environmental Science*, vol. 277, June 2019, p. 012007, <https://doi.org/10.1088/1755-1315/277/1/012007>. IOPScience. Accessed 28 Sept. 2024.
- [30] SwiftPak. "Why Is PLA Not Sustainable?" *Swiftpak*, 24 Oct. 2023, [www.swiftpak.co.uk/insights/why-is-pla-not-sustainable#:~:text=Analysts%20have%20e,ven%20estimated%20that](http://www.swiftpak.co.uk/insights/why-is-pla-not-sustainable#:~:text=Analysts%20have%20e,ven%20estimated%20that). Accessed 28 Sept. 2024.
- [31] Kliem, Silvia, et al. "Review on the Biological Degradation of Polymers in Various Environments." *Materials (Basel, Switzerland)*, vol. 13, no. 20, Oct. 2020, <https://doi.org/10.3390/ma13204586>. MDPI. Accessed 28 Sept. 2024.
- [32] Giannantonio, Michela Di., et al. "Ecotoxicity of Polyvinylidene Difluoride (PVDF) and Polylactic Acid (PLA) Microplastics in Marine Zooplankton." *Toxics*, vol. 10, no. 8, 2022, <https://doi.org/10.3390/toxics10080479>. MDPI. Accessed 28 Sept. 2024.
- [33] Zimmermann, Lisa, et al. "What Are the Drivers of Microplastic Toxicity? Comparing the Toxicity of Plastic Chemicals and Particles to *Daphnia Magna*." *Environmental Pollution*,

- vol. 267, Dec. 2020, <https://doi.org/10.1016/j.envpol.2020.115392>. Elsevier. Accessed 28 Sept. 2024.
- [34] Miller, Michaela E., et al. "Bioaccumulation and Biomagnification of Microplastics in Marine Organisms: A Review and Meta-Analysis of Current Data." *PLOS ONE*, edited by Amitava Mukherjee, vol. 15, no. 10, Oct. 2020, <https://doi.org/10.1371/journal.pone.0240792>. PubMed. Accessed 28 Sept. 2024.
- [35] Alberghini, Leonardo, et al. "Microplastics in Fish and Fishery Products and Risks for Human Health: A Review." *International Journal of Environmental Research and Public Health*, vol. 20, no. 1, Dec. 2022, <https://doi.org/10.3390/ijerph20010789>. PubMed. Accessed 28 Sept. 2024.
- [36] Zimmermann, Lisa, et al. "Are Bioplastics and Plant-Based Materials Safer than Conventional Plastics? In Vitro Toxicity and Chemical Composition." *Environment International*, vol. 145, Sept. 2020, p. 106066, <https://doi.org/10.1016/j.envint.2020.106066>. ScienceDirect. Accessed 28 Sept. 2024.
- [37] Steven, Steven, et al. "Facile Preparation of Cellulose Bioplastic from Cladophora Sp. Algae via Hydrogel Method." *Polymers*, vol. 14, no. 21, Nov. 2022, p. 4699, <https://doi.org/10.3390/polym14214699>. MDPI. Accessed 28 Sept. 2024.
- [38] Daphne, Tian. "Bioplastics: Sustainability in the Plastics Industry." *Circularise.com*, 31 Dec. 2023, [www.circularise.com/blogs/bioplastics-sustainability-in-the-plastics-industry#:~:text=The%20progress%20is%20encouraging%20so](http://www.circularise.com/blogs/bioplastics-sustainability-in-the-plastics-industry#:~:text=The%20progress%20is%20encouraging%20so). Accessed 28 Sept. 2024.

# **Bridging Cultures: Analyzing Phonological Challenges in Korean Language Acquisition for Filipino Learners By Dominic Hahm**

## **Abstract**

This study investigates the linguistic challenges faced by Filipino learners of Korean, rooted in the significant phonological differences between Tagalog and Korean. As the Korean Wave (*Hallyu*) continues to influence Filipino interest in Korean culture and language, understanding these challenges becomes increasingly pertinent. Employing theories of adult second language acquisition (SLA), the research begins by comparing the sound inventories and phonological structures of both languages. It highlights phonetic differences that may hinder Tagalog speakers in perceiving and producing Korean sounds accurately. The analysis ends in a discussion of the implications for language acquisition, detailing how these challenges surface in the Korean language classroom. This research has implications for pedagogical approaches and curriculum development, providing insights for educators working with Filipino students pursuing Korean language proficiency.

## **Introduction**

The rise of globalization has fostered unprecedented cultural and linguistic exchanges between nations, exemplified by the dynamic relationship between South Korea and the Philippines. Thanks to the geographical proximity and political alliance, the two countries maintain a positive relationship, and Filipino citizens have been further drawn to Korean language and culture because of *Hallyu*, the “Korean Wave.” The term refers to the rise in the popularity of Korean pop culture around the globe. It began in the mid 1990s and early 2000s when the Korean TV dramas *What Is Love?* and *Winter Sonata* became megahits in China and Japan respectively; soon after, K-pop groups like *Super Junior* and *Girls Generation* also took global music charts by storm (“Hallyu (Korean Wave)”). The impact of *Hallyu* can be quantified through increased consumer spending on Korean products, tourism to Korea, and global viewership of K(Korean)-pop, K-drama, and Korean films. In the Philippines, K-dramas have consistently aired on TV for the past 20 years and K-pop has become background music for everyday life (Igno & Cenidoza). The obsession with Korean media has led to an increase in Filipinos visiting Korea for tourism, school, or work. Thus, the number of Filipinos learning Korean, whether in Korea or in the Philippines, is higher than ever.

Korean and Tagalog are markedly different, however, so Filipino learners who speak Tagalog as their native language (L1) may find it difficult to learn Korean. Using theories of adult language acquisition as a framework, this research compares the characteristics of Korean and Tagalog to predict the challenges that Filipino learners of Korean will encounter while acquiring Korean as their second language (L2). Section 1 describes the sound inventories and phonological systems of each language, focusing on points of difference. Section 2 places those phonetic and phonological differences within a larger context of L1 and L2 acquisition. Finally, Section 3 identifies the challenges Filipino learners of Korean will face while perceiving and

speaking Korean based on differences between their L1 and L2. This research has implications for our understanding of the perceptual system, language acquisition, and language education. Effective curricula and teaching strategies may be developed to suit those who are planning to learn Korean as a native speaker of Tagalog.

## 1. Sound Inventory & Phonology

### 1.1 Tagalog Vowels

Tagalog is an Austronesian language and the national language of the Philippines. It is spoken by over 28 million people as a first language, and millions more as a second language, across the nation.

The Tagalog language consists of 5 vowel phonemes: a, i, e, o, u. They are divided into front, central, and back vowels, classified by the position of the end of tongue when pronounced. They are also divided into close, mid, and open, classified by the relative state of the jaw when pronounced. The vowels and their positioning can be seen in Figure 1 below. As indicated in the figure, Tagalog does not have central vowels, and most are closed. Additionally, Tagalog also has semivowels /j/ and /w/, which sound like vowels but act as consonants.

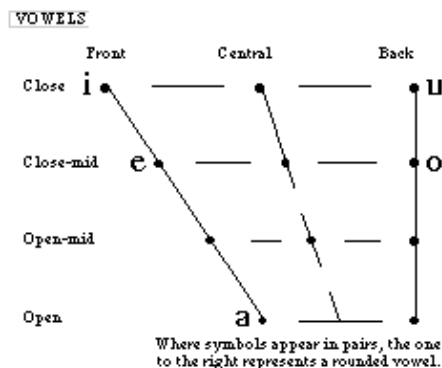


Figure 1. Tagalog Vowel Chart

Source: George Mason University, Linguistics Department

### 1.2 Tagalog Consonants & Syllable Structure

	<i>Labial</i>	<i>Dental</i>	<i>Alveolar</i>	<i>Palatal</i>	<i>Velar</i>	<i>Glottal</i>
Voiceless stop	p	t			k	ʔ
Voiced stop	b	d			g	
Nasal	m	n			ŋ	
Fricative			s			h
Lateral			l			
Tap or trill			r			
Glide	w			y		

Figure 2. Tagalog Consonant Phonemes

Source: University of Hawaii at Manoa

Tagalog has 16 consonant phonemes as shown in Figure 2 above. Consonants are classified according to their place (row) and method (column) of articulation. Tagalog follows the syllable system of consonant-vowel-consonant (CVC), where the most complex arrangement of alphabets a syllable can have is two consonants surrounding a vowel. In Tagalog, consonant clusters with two or more consonants together in the same syllable are avoided, though borrowed foreign words may contain them. For instance, the borrowed word *traysikel*, from the English word “tricycle,” contains the consonant cluster “tr,” which would typically be avoided in native Tagalog. Additionally, there are only a limited amount of consonants that appear in the syllable-final position, and those include p, t, k, and ng among others.

### 1.3 Korean Vowel System

Korean is the official language of South Korea. It is spoken by approximately 77 million people around the world as their first language. Though its classification in relation to other languages is still debated, it is generally believed to be an isolate in its own Koreanic language family.

Korean has 10 vowel phonemes. Figure 3 below shows the different classifications of the Korean vowels, with the added category of unround/round; this signifies the shape of the lips when the vowel is pronounced. High/mid/low correlates to the close/mid/open category discussed above.

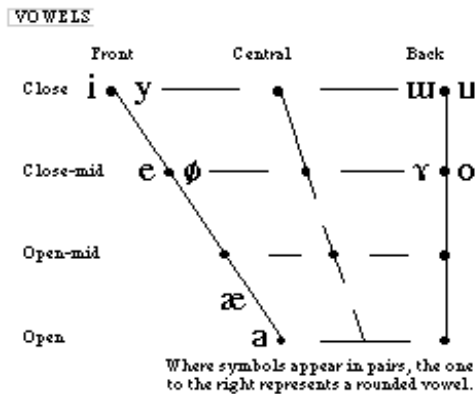


Figure 3. Korean Vowel Chart  
Source: Speech Accent Archive

### 1.4 Korean Consonant/Syllable System

		Bilabial	Alveolar	Post-alveolar	Velar	Glottal
<b>Nasal</b>		ㅁ /m/	ㄴ /n/		ㅇ /ŋ/[20]	
<b>Plosive and Affricate</b>	plain	ㅂ /p/	ㄷ /t/	ㅈ /t͡ʃ/	ㄱ /k/	
	tense	ㅃ /p̚/	ㄸ /t̚/	ㅉ /t͡ʃ̚/	ㄲ /k̚/	
	aspirated	ㅍ /pʰ/	ㅌ /tʰ/	ㅊ /t͡ʃʰ/	ㅋ /kʰ/	
<b>Fricative</b>	plain		ㅅ /sʰ/			ㅎ /h/
	tense		ㅆ /s̚/			
<b>Liquid</b>		ㄹ/w <sup>1</sup> /	ㄹ /l/	ㄹ/j <sup>1</sup> /		

Figure 4. Korean Consonant Phonemes  
Source: Hanguk Babble

Korean has 21 consonant phonemes as shown in Figure 4 above. There are a few characteristics which make Korean consonants distinct. First, many of Korean consonants are classified by the presence of aspiration, which refers to a burst of breath from the mouth, when pronouncing the consonant. Thus, the /p/ phoneme is the unaspirated variant (no puff of breath) and the /pʰ/ indicates the aspirated variant (with the puff of breath) in Figure 4 above. Additionally, Korean has many “strong consonants,” officially known as tense or fortis consonants. These consonants, including ㄲ (*kk*), ㄸ (*tt*), ㅃ (*bb*), and more, are produced with increased tension in the vocal cords and unaspirated.

Korean allows several syllable structures including consonant-vowel (CV), consonant-vowel-consonant (CVC), and CGVC. A most complex syllable (CGVC) begins with a consonant, followed by a glide and a vowel, and ends in another final consonant. Korean does not allow a consonant cluster at the beginning of a syllable. Also, similarly to Tagalog, it has limitations on the syllable-final consonants; these are ㄱ /g/, ㄴ /n/, ㄷ /d/, ㄹ /l/, ㅁ /m/, ㅂ /b/, ㅇ /ŋ/. Finally, there also cannot be multiple consonants in the final position.

### 1.5 Comparison

As explored above, the two languages differ vastly in terms of their phonetic inventories. First, the number of Korean vowel phonemes is double the number of phonemes in Tagalog, meaning Korean has many vowel categories that do not exist in Tagalog. For instance, [i̯] exists in Korean but not Tagalog. Additionally, Korean also has an extensive inventory of diphthongs and glides in addition to the pure vowels mentioned in Section 1.3. This means that Tagalog native speakers must acquire at least 10 new vowel phonemes to speak Korean proficiently. Similarly, Korean language’s unique usage of aspirated, unaspirated, and strong consonants will also be unfamiliar to the Filipino ear, so Filipino learners will have to train their ear to detect aspiration and strong vowels. This might be a challenge if their perception of vowels and consonants is restrained by the phonemes in their own native language of Tagalog. The theoretical underpinning behind this will be discussed in the next section.



## 2. Theoretical Background: The Relationship Between L1 and L2

### 2.1. *Second Language Acquisition (SLA)*

To begin to shed light on how the phonemic and phonological differences between Tagalog and Korean affect Filipinos learning Korean, it is necessary to first understand how a person's native language (L1) can affect acquisition of the target language (L2). Second language acquisition (SLA) is a process in which individuals learn a language other than their mother-tongue. SLA encompasses the study of the cognitive, psychological, and linguistic processes that are involved when learning a second language. For decades, Linguistics scholars have been exploring how a person's L1 directly influences the process of second language acquisition. Exploring this relationship reveals why some learners struggle with learning a specific second language, or develop specific strategies to mend the challenges they face.

### 2.2. *Abstract & Surface Transfer*

When learning a new language, adults do not start from scratch. Previous knowledge impacts the acquisition process through abstract and surface transfer. Abstract transfer, also known as conceptual transfer, is the application of principles, rules, and patterns from L1 to L2 at a more advanced level (Kartushina). When participating in abstract transfer, learners not only follow the simple linguistic elements, but also more complex structures and problem-solving strategies when learning their L2. For instance, a Spanish speaker learning English could transfer his or her problem-solving skills, realizing that all sentence structures of English questions utilize inversion (e.g. "Where are you?"), just like the sentence structures of Spanish questions (e.g. "¿Dónde estás?). On the other hand, surface transfer refers to the transfer of previous knowledge related to basic phonetic properties and categories when learning L2. For instance, native Japanese speakers often have a harder time acquiring the English /l/ sound than the English /r/ sound. This is because the English /r/ sound is more distinct from the Japanese /r/ sound, making it relatively easier to differentiate and understand (Kartushina). However, the similarity of the English /l/ sound and the Japanese /r/ sound have offers a possibility for confusion. In this way, the L1 phonemic categories present in a speaker's mind directly impact how quickly they are able to distinguish new sounds from each other. Furthermore, as with Japanese learners of the English language, language learners can also find it challenging when a phoneme in L2 is so similar to an L1 phoneme because their brain maps the new sound onto the existing mental category instead of creating a new one.

### 2.3. *Priming Effects*

As learners begin to acquire simple sounds and the basic phonology of their L2, they begin acquiring words and sentences. Low proficiency learners often struggle with both word tasks and sentence context tasks, due to the early reversed priming effects. Priming is a cognitive function where the brain allows you to quickly sort and relate certain concepts with another, in order to facilitate a response. On the other hand, reversed priming effects are the opposite;

previous, conflicting knowledge (of the L1 in this case) inhibits the learners from thinking quickly, instead slowing them down and leading to a repressed response. This often occurs due to the different linguistic nature of learners' L1 and L2; if the two languages are extremely dissimilar, then the expectations made from their L1 will often get conflicted and rebutted during their L2 acquisition (Elston-Güttler). Therefore, because they are not much familiar with the L2 language structure yet, low proficiency learners will have a difficult time eliminating the conflicting linguistic ideas.

Despite the incorrect nature of the combination of the two languages, learners often do integrate it into their learning processes, because of how pervasive the interference is during their cognitive procedures. It is proven that when learners do not know the correct L2 syntax, they will end up simply utilizing the L1 syntax instead, which demonstrates how much L1 affects L2 production. Unlike high proficiency learners who are able to “think in the second language,” low proficiency learners often believe in the word-for-word translation from L1 to L2, which results in incorrect results most of the time (Bhela).

#### *2.4. Implications for Filipino Learners of Korean*

This information suggests that native speakers of Tagalog will experience difficulties in mapping on Korean sounds due to the difference in their sound inventories. Specifically, the Korean “ㄹ” can be pronounced either as a flap [r] or an [l], while in Tagalog there is only one /r/ sound. This limitation will present barriers for Filipinos in differentiating the “ㄹ” in “로” (ro) and “방울” (bang-ul). Another example is the alphabet “ㅅ” /s/. In Korean, this alphabet diverges into two different sounds depending on the situation – typically an unaspirated “s,” but similar to “sh” if placed before high vowels. Having a singular “s” sound in their phonetic system, Filipino speakers will face difficulty when mapping and producing sounds like ㅅㅣ (shi), failing to adjust based on the linguistic context and pronouncing it with a consistent /s/ sound instead.

In addition, it will be possible to notice reversed priming effects extensively affecting Filipinos due to the significant differences in their phonology. Let's take final consonants as an example. In Tagalog, final consonants are fully pronounced, as seen in words such as “sarap.” In Korean, however, they are unreleased, meaning that they are not completely pronounced. As such, if Filipino students allow reversed priming effects to interfere with their learning, they will struggle with producing sounds like “ㅃ” (bap), pronouncing it as “bapuh.” The effect happens because of the instant connection learners make to their L1 as soon as they see final consonants; the consistent full pronunciation in their L1 will most definitely interfere with the accurate production of the L2 sound.

### **3. Implications for Language Production in the Classroom**

Due to the significant differences in the phonological systems of Tagalog and Korean – mainly the vowel, consonant, and syllable systems – Filipino learners will struggle mapping on Korean sounds during their learning, which often leads to incorrect production of sounds and sentences. These mistakes are already documented in existing SLA literature. For example, an

extensive error analysis from University of the Philippines Diliman describes the main type of errors Filipinos make when producing Korean sentences. In the study, Filipino learners of Korean as a foreign language were divided into 3 groups—beginner, intermediate, and advanced groups. From 16 students, 644 sentences were produced and analyzed, with 383 of them containing at least one error. The most common types of errors varied greatly: omission, addition, selection, and ordering (Manrique).

Omission errors refer to when an item that must be present in a well-formed utterance is absent. For instance, omission of space, particles like the Korean *eul* (을) and *ga* (가), complementizers like the Korean *rago* (라고), and adverbs were present throughout the analysis. Addition errors include unnecessary space, failure to omit particles, unnecessary pronouns, and redundancy. Selection errors include misspelling, misselection of particles, misselection of allomorphs like the Korean *gwa* (과) and *wa* (와), etc. Ordering errors, like misplacement of particles and word order, were the least common out of all. Finally, counterintuitively, it was researched that advanced learners made more errors than intermediate learners, possibly due to their tendency to create more complex sentences, experiment more. Inversely, beginners and intermediate learners tend to answer more conservatively, leading to fewer mistakes.

As aforementioned, many reasons could factor into the production of incorrect sentences. Firstly, the differences between the sound inventory of the two languages, like the unsimilar syllable structure and quantity of vowels may account for most errors. The difficulty in mapping the other language's sounds onto the learners' brains will require lots of mental effort, and therefore lead to a longer time fully acquiring the sounds. Additionally, the glides or semivowels that exist in both languages will definitely be an inhibitory factor when learning. The two types of glides are different: /w/ and /j/ in Korean and /w/ and /y/ in Tagalog. This will affect the language acquisition even more negatively than learners whose native languages do not have glides at all, because the previous (but inaccurate, in terms of Korean) knowledge of glides will most definitely attempt to intervene while learning.

Additionally, it is also possible that there are problems regarding non-grammar related issues, such as the learning environment and method. For instance, listening to L2 languages at a fast tempo will evoke difficulties for beginning to intermediate students. According to the International Society for Humanities and Social Studies, some difficulties that students argue include fast-speaking professors and speakers who do not open their mouth widely and clearly (Andres). Applying the concept of abstract and surface transfer to Korean and Tagalog, it is undeniable that most learners will suffer difficulties when attempting to map sounds onto the other language. Due to the great difference in the sound systems, challenges in assimilating the two together will inevitably surface.

So, given the fact that Filipino learners will encounter the above difficulties when acquiring Korean, what can be done to improve the quality of Korean language education for Filipino learners? First, it may be beneficial for the Korean language instructor to be familiar with Tagalog. For example, in a qualitative study about non-native teachers of Korean at the University of the Philippines, Filipino teachers identified that their knowledge of Tagalog was

beneficial in the classroom because they could “predict what type of mistakes and errors Filipino students would make” and check comprehension of the class material in the students’ native language (Bae & Igno). This kind of pre-emptive action taken by Filipino teachers shortens the time of confusion for Filipino students learning Korean. In situations where the Korean language instructor is a native Korean speaker, it would be beneficial for them to learn the linguistic differences between the two languages if the majority of their students are Filipino. This is a reasonable recommendation for teachers based in the Philippines, but perhaps less so for teachers based in South Korea who are likely leading diverse classes. In that case, Korean language instructors in Korea can slow down the pronunciation of words that contain several consonant clusters, glides, and strong or aspirated consonants, as these are some of the most distinctive features of the Korean language that students are unfamiliar with. The pace of instruction and whether additional practice is necessary should be left at the discretion of the teacher whenever possible, as it is impossible to recommend a suitable pace for all language learners.

### **Conclusion**

In conclusion, due to globalization, migration, and *Hallyu*, the number of Filipinos learning the Korean language has drastically increased in the Philippines and South Korea. However, linguistic differences between their native language and Korean poses challenges for those who wish to fully master the language. Some of these linguistic differences include different vowel and consonant phonemes as well as a different syllable structure. Because Filipino students first learning Korean will perceive and produce Korean based on the sounds in their native language, they are likely to make a lot of mistakes. Luckily, these mistakes are predictable if one understands the sounds and rules of Tagalog. Therefore, in a Korean class with predominantly Filipino students, it is recommended that the teacher is familiar with the basics of the Tagalog language. In mixed classrooms, such as at language schools in Korea, it would be beneficial for the program to create additional handouts that address the common types of errors made by Filipino learners. Given the trend over the past two decades, the Filipino interest in Korea is not likely to wane soon; it is thus imperative for language classrooms to pivot and meet the current and future demands of language students.

## Works Cited

- Ancho, Inero V. "Exploring Motivation, Satisfaction, and Preparedness of Teacher Candidates in the Philippines." *Bicol University Research and Development Journal*, vol. 19, no. 1, 2022, pp. 9-18. [https://journal.bicol-u.edu.ph/assets/journal\\_pdf/2%20Ancho\\_9-18.pdf](https://journal.bicol-u.edu.ph/assets/journal_pdf/2%20Ancho_9-18.pdf).
- Andres, Henry, and Jong Oh Ra. "Problems, Strategies, and Developments of Filipino KFL Learners in Learning the Korean Language." ResearchGate, 2023. [https://www.researchgate.net/profile/Henry-Andres-4/publication/372393076\\_Problems\\_Strategies\\_and\\_Developments\\_of\\_Filipino\\_KFL\\_Learners\\_in\\_Learning\\_the\\_Korean\\_Language/links/64b34de68de7ed28ba9fa432/Problems-Strategies-and-Developments-of-Filipino-KFL-Learners-in-Learning-the-Korean-Language.pdf](https://www.researchgate.net/profile/Henry-Andres-4/publication/372393076_Problems_Strategies_and_Developments_of_Filipino_KFL_Learners_in_Learning_the_Korean_Language/links/64b34de68de7ed28ba9fa432/Problems-Strategies-and-Developments-of-Filipino-KFL-Learners-in-Learning-the-Korean-Language.pdf).
- Bae, Kyungmin, and Jay-Ar Igno. "A Study on Non-native Teachers of Korean as a Foreign Language (KFL): Focused on Filipino KFL Teachers." *Strengthening Korean Studies in the Philippines: 2012 Philippine Korean Studies Symposium Proceedings*. 2012.
- De Leon, Jocelyn T. Acoustic Characteristics of the Filipino Vowel Space. Academia.edu. [https://d1wqtxts1xzle7.cloudfront.net/35291931/Acoustic\\_Characteristics\\_of\\_the\\_Filipino\\_Vowel\\_Space-libre.pdf](https://d1wqtxts1xzle7.cloudfront.net/35291931/Acoustic_Characteristics_of_the_Filipino_Vowel_Space-libre.pdf).
- "Error Analysis of Written Compositions of Filipino Learners of Korean as a Foreign Language." *Linguistics UP Diliman*, 2022. <https://linguistics.upd.edu.ph/wp-content/uploads/2022/08/14-Error-Analysis-of-Written-Compositions-of-Filipino-Learners-of-Korean-as-a-Foreign-Language.pdf>.
- "Hallyu: The Korean Wave." *Korea.net*, Korean Culture and Information Service, [www.korea.net/AboutKorea/Culture-and-the-Arts/Hallyu](http://www.korea.net/AboutKorea/Culture-and-the-Arts/Hallyu).
- Klemhofer, Petra, et al. Native Language Interference in Learning a Second Language: Exploratory Case Studies. University of Essex, 2005. <https://repository.essex.ac.uk/12921/1/089892905774597245.pdf>.
- Lemhöfer, Kristin, et al. LexTALE: A Test to Assess Dutch and English Word Knowledge in Non-Native Speakers. ResearchGate, 2008. [https://www.lextale.com/pdf/Lemhofer\\_et\\_al\\_2008.pdf](https://www.lextale.com/pdf/Lemhofer_et_al_2008.pdf).
- Leon, Lance Isaac. Notes on Filipino-Hallyu Pop Culture: Exploration of a Korean Proto-Lexicon Framework on Tagalog Speakers. ResearchGate, 2022. [https://www.researchgate.net/profile/Lance\\_Isaac\\_Leon/publication/362239722\\_Notes\\_on\\_Filipino-Hallyu-Pop\\_Culture\\_Exploration\\_of\\_a\\_Korean\\_Proto-Lexicon\\_Framework\\_on\\_Tagalog\\_Speakers/links/62fcd262eb7b135a0e3f85cd/Notes-on-Filipino-Hallyu-Pop-Culture-Exploration-of-a-Korean-Proto-Lexicon-Framework-on-Tagalog-Speakers.pdf](https://www.researchgate.net/profile/Lance_Isaac_Leon/publication/362239722_Notes_on_Filipino-Hallyu-Pop_Culture_Exploration_of_a_Korean_Proto-Lexicon_Framework_on_Tagalog_Speakers/links/62fcd262eb7b135a0e3f85cd/Notes-on-Filipino-Hallyu-Pop-Culture-Exploration-of-a-Korean-Proto-Lexicon-Framework-on-Tagalog-Speakers.pdf).
- "Oral Communication Skills and Experience in Learning the Korean Language by Filipino Learners." *IOER International Multidisciplinary Research Journal*, 2022. <https://www.ioer-imrj.com/wp-content/uploads/2022/08/Oral-Communication-Skills-and-Experience-in-Learning-the-Korean-Language-by-Filipino-Learners.pdf>.
- Radford, Andrew, et al. "The Neural Bases of Learning and the Lexicon: A Report." *Frontiers in Psychology*, vol. 5, 2014, Article 1246. <https://doi.org/10.3389/fpsyg.2014.01246>.

Scholar Google. "Native Language Interference in Learning a Second Language."

[https://scholar.google.com/scholar?hl=ko&as\\_sdt=0%2C5&q=Native+language+interference+in+learning+a+second+language%3A+Exploratory+case+studies+of+native+language+interference+with+target+language+usage&btnG=](https://scholar.google.com/scholar?hl=ko&as_sdt=0%2C5&q=Native+language+interference+in+learning+a+second+language%3A+Exploratory+case+studies+of+native+language+interference+with+target+language+usage&btnG=).

ScholarSpace Manoa. The Article from ScholarSpace at the University of Hawaii.

<https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/8e7f819f-68be-4cb3-95a3-18a057f436fe/content>.

So, Kyunghee, and Mina Noraisikin. "The Role of Hallyu 1.0 and Hallyu 2.0 in the Korean Wave: A Case Study of Taiwan." *International Journal of Social Science and Humanity*, vol. 6, no. 9, 2016, pp. 691-695, [www.ijssh.org/vol6/740-SH013.pdf](http://www.ijssh.org/vol6/740-SH013.pdf).

"Tagalog IPA Vowels." The Speech Accent Archive. George Mason University.

[https://accent.gmu.edu/nl\\_inventories/tagalogipavowels.gif](https://accent.gmu.edu/nl_inventories/tagalogipavowels.gif).

# Hyperparameter Tuning Improves Computer Vision Tools for Wildlife Conservation

## By Achyut Venkatesh

### Abstract

Camera traps are an excellent way to collect a large amount and variety of wildlife data, however, the volume of images poses a serious challenge for manual analysis. Through the use of machine learning and computer vision, this process can be automated, enabling rapid and accurate classification and identification, thus aiding conservation efforts. This study explores the effectiveness of EfficientNet models in classifying eight different classes from camera trap images, focusing on how hyperparameters such as learning rate and batch size affect model performance. Using a dataset of 16,487 images from Tai National Park, we experimented with different hyperparameter values and found that a lower learning rate and a moderate batch size yielded the highest accuracy. The `efficientnet_b1` model was the most effective model, achieving 86.27% accuracy with 20 epochs. It identified leopards, hogs, birds, and genet civets with over 90% accuracy but struggled with the blank class. Training was completed in under eight hours on a single laptop, showcasing the efficiency of lightweight models. Our findings underscore the potential of computer vision in conservation, enabling rapid and accurate analysis of large datasets that would take large amounts of time and workers to go through manually. This work highlights the importance of hyperparameter tuning in enhancing model performance, paving the way for more effective automated wildlife monitoring tools. Future work will focus on improving accuracy for challenging classes and testing model adaptability in different environments.

### Introduction

Camera traps are invaluable tools for conservation, providing non-invasive methods to monitor wildlife populations and behaviors over large geographic areas and extended periods. These devices can capture images of animals in their natural habitats without human interference, offering crucial data for ecological research and conservation planning. However, the sheer volume of images generated by camera traps presents a significant challenge for manual analysis.

Machine learning (ML) and computer vision present a solution to this challenge. These technologies allow for the data analysis to be automated, significantly reducing the time and effort required to comb through the data. They allow for tasks such as classification and identification to be completed at astonishing speeds. By using ML models, conservationists can process large datasets quickly and accurately, gaining insights that inform conservation strategies and policy decisions.

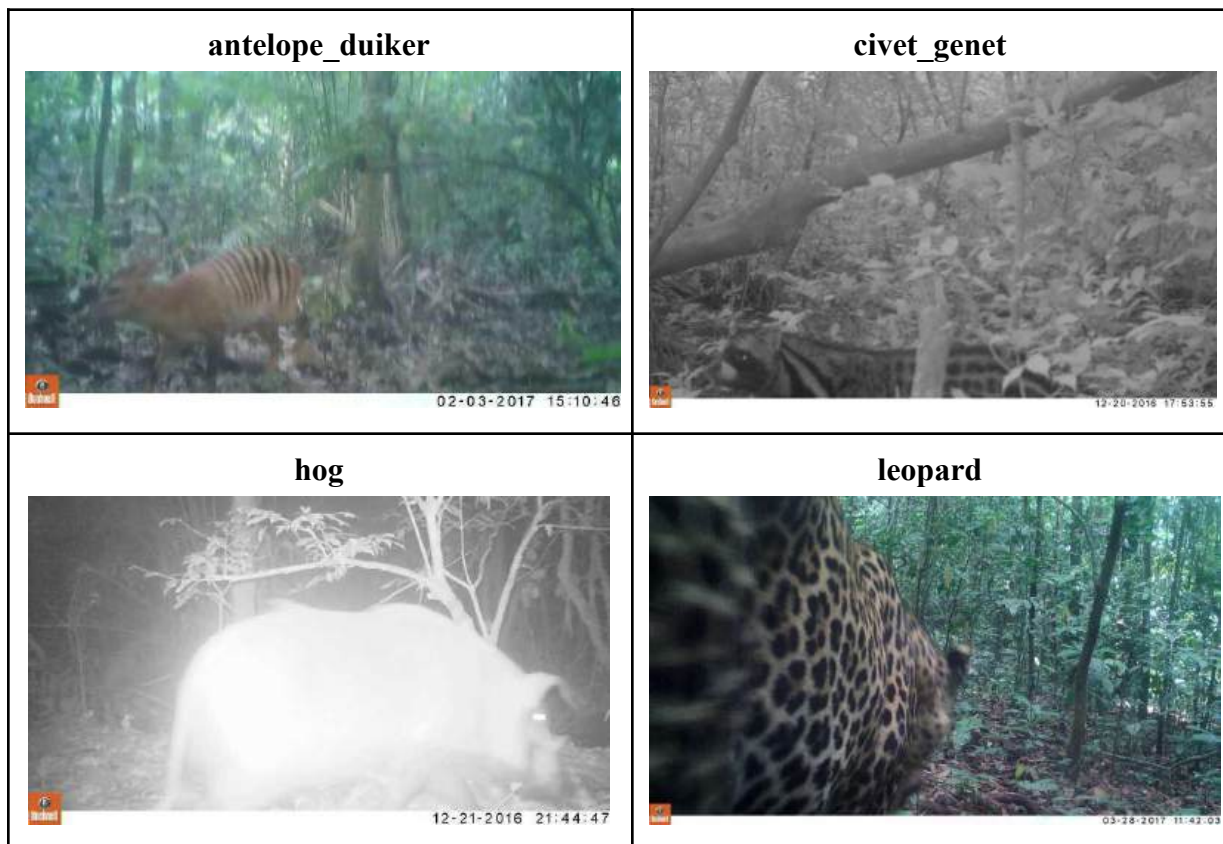
Several studies have already demonstrated the effectiveness of ML in conservation efforts. A study conducted by Mohammad Sadegh Norouzzadeh and his colleagues utilized deep learning to classify 48 species from camera trap images, achieving high accuracy and demonstrating the potential for large-scale automated wildlife monitoring (Norouzzadeh et al.,

2020). Another study conducted by Marco Willi and his team developed a deep learning model to identify animal species in camera trap images across diverse ecosystems, highlighting the generalizability of these methods (Willi et al., 2018).

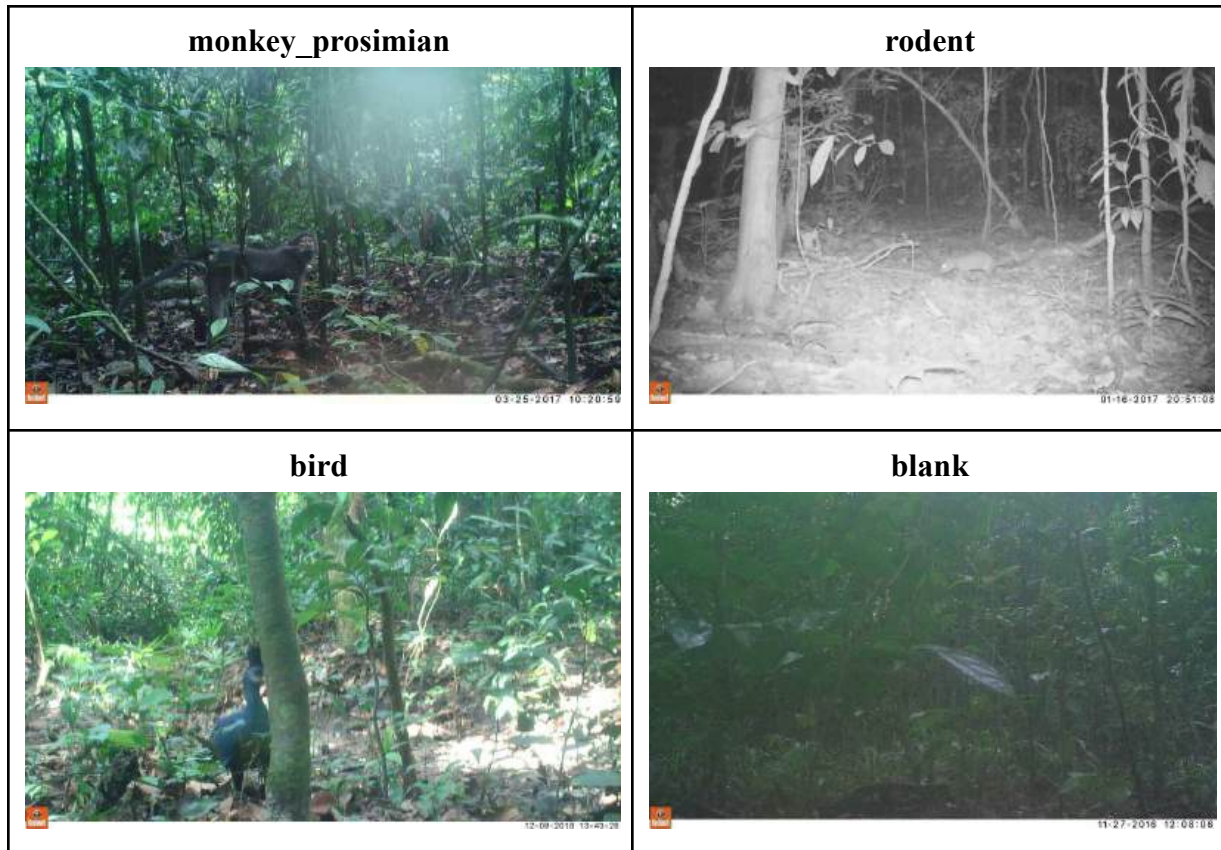
In this paper, we aim to evaluate how different hyperparameters affect the performance of a computer vision model trained to classify camera trap images. We focus on the EfficientNet model architecture and assess the impact of varying the batch size and learning rate values on the model's training efficiency and accuracy. By tuning the hyperparameters and identifying the optimal values, we aim to maximize the model's performance in identifying animal species from a camera trap dataset. This work contributes to the broader goal of enhancing automated wildlife monitoring tools, making them more effective and accessible for conservation efforts worldwide.

### Data

The data used for this project were camera trap images taken from various locations across the Tai National Park in Côte d'Ivoire. These images were collected by the Wild Chimpanzee Foundation and the Max Planck Institute for Evolutionary Anthropology from 2016 to 2018. The dataset used had a total of 16487 images split into 8 unique classes, each containing a different species. The 8 classes were the Duiker Antelope, Genet Civet, Hog, Leopard, Prosimian Monkey, Rodent, Bird, and a blank class. Table 1 below shows a sample image from each class.







**Table 1.** Sample image of each class

Each class had 2,000+ images except for the hog which only had 979 images. This discrepancy indicates that the hog class is underrepresented compared to the others. This dataset presents several challenges while using lightweight models. First of all, the image isn't always centered and doesn't contain the full animal. Often times it will only contain a limb such as a leg or a tail. Secondly, the animals aren't in the same spot in each image. Some images will have the animal in the center while others have it in a corner. There is also foliage and other disruptors that are also captured in the camera trap image. Finally, the lightweight models used in this study only operate with low-resolution images. Since the camera trap images are of a higher resolution, they need to be coarsened to fit the model's requirements.

## Methods

The models used were the EfficientNet machine learning models. The EfficientNet architecture was first introduced by researchers at Google in 2019 and now consists of numerous iterations. The EfficientNet model is a computer vision model based on convolutional neural networks that allows for rapid training and quick results, however, it can only handle lower-resolution images. For this project, only EfficientNet models with parameters under 250 pixels by 250 pixels were used and compared to ensure that the models had equal standing.

The primary model experiments we explored involved altering the EfficientNet hyperparameters. Tweaking hyperparameters is crucial in machine learning to optimize model performance and ensure it generalizes well to new data. Additionally, proper parameter tuning enhances training efficiency and stability, adapting the model to the specific characteristics of the dataset. By finding the optimal parameters, the model can function at full capacity and produce the best results.

The parameters that were experimented with in this project were the learning rate, the batch size, and the number of epochs. The learning rate is a hyperparameter that determines the size of the steps the model takes during optimization when updating the weights based on the gradient of the loss function. A properly set learning rate ensures that the model converges efficiently to a minimum of the loss function; if it's too high, the model might overshoot the minimum, and if it's too low, the training process can become excessively slow or get stuck in local minima. A higher learning rate would be expected to result in faster training time but less accurate identification. In contrast, a lower learning rate would be expected to result in a slower training time but more accurate identification.

The next main parameter experimented with in this project is batch size. Batch size is a hyperparameter in machine learning that determines the number of training images processed together before the model's internal parameters are updated. A smaller batch size provides more frequent updates and can lead to more precise adjustments, while a larger batch size results in faster computation per epoch and smoother gradient estimates but requires more memory.

The final parameter in the efficientnet architecture manipulated here is the number of epochs. The number of epochs is a hyperparameter in machine learning that defines the number of complete passes through the entire training dataset during the training process. Each epoch allows the model to learn and update its parameters based on the entire dataset. A higher number of epochs would result in a much longer training time but a much more accurate model. A lower number of epochs would result in a smaller training time but a less accurate model.

## **Model Validation**

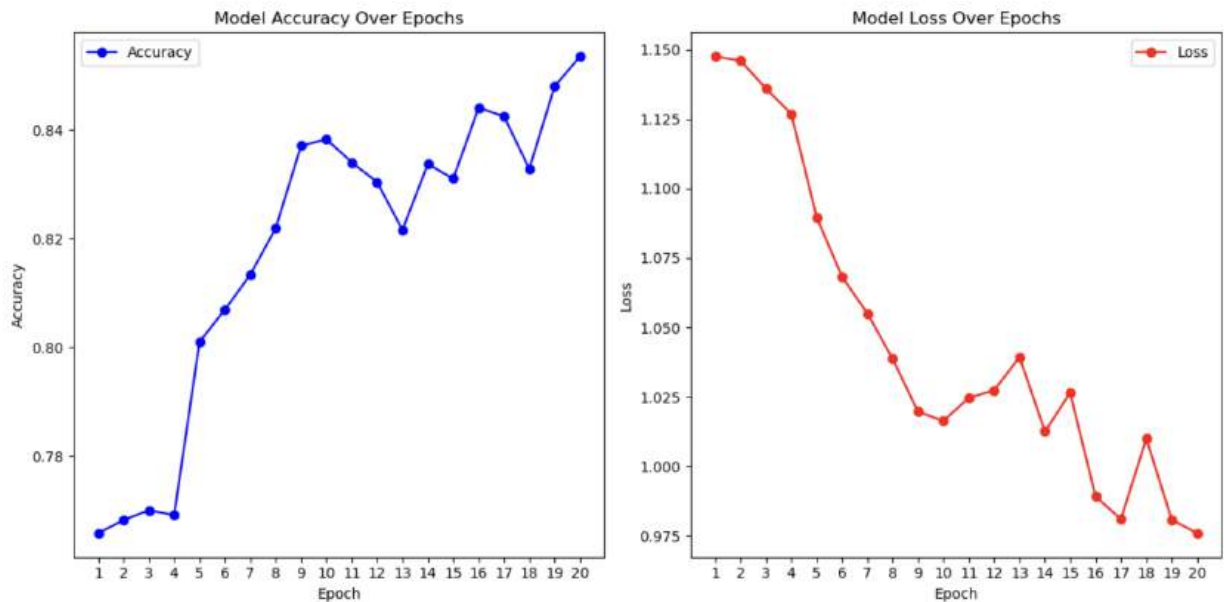
First, the model was run with a batch size of 4 and increasing learning rates while keeping the number of epochs constant. After the ideal learning rate was determined, that rate was kept constant while the batch size was multiplied by four in each trial. Using this process, the most effective learning rate and batch size were determined for the base model. Since the other EfficientNet models were iterations of this base model, the same optimal parameters were applied when conducting trials to determine the most effective model. Then, once the best model was found, the epoch number was drastically increased until the accuracy and loss displayed significantly diminished returns. This allowed for the best possible accuracy and the lowest possible loss from the optimal model.

To determine the most accurate model with the most optimal parameters, a base model is necessary. In this case, the efficientnet\_b0 model was used as the base model to determine the ideal parameter values.

The model's performance was judged by determining the percentage of images identified correctly for each class and then averaging those values together. Another way that the models' effectiveness was judged was through log loss.

To evaluate the accuracy and loss from the model a separate dataset that the model wasn't trained with is run through the model. Since there was only one training set, a dataset consisting of 3297 images was made from the training set known as a validation set. The model was not trained on this validation set and was only exposed to it when evaluating the final percentages, accuracy, loss, and log loss.

## Results



**Fig. 1.** These graphs represent how the accuracy and loss of the final model changed as the number of epochs increased. We can see that the accuracy is increasing and the loss is decreasing as the model trains.

The best-performing model was the `efficientnet_b1` model which beat out the `efficientnet_b0`, `efficientnetv2-b0`, and `efficientnetv2-b1` models. The winning parameters were a learning rate of 0.005 and a batch size of 16. This means that the lowest learning rate and a moderate batch size led to the most accurate results. The models were all compared with 10 epochs, but once it was determined that `efficientnet_b1` was the best model, it was run with 20 epochs to get the final percentage of 86.27%. It accepts up to images of 240 pixels by 240 pixels. It can be seen from the graphs that the accuracy and loss will still change a substantial amount past epoch 20, suggesting that the model can be even more accurate at identifying the classes.

To train the model with 10 epochs it took 3-4 hours. Training the model with 20 epochs took upwards of 8 hours. This was expected and is still a reasonable amount of time to train a

model, as it can be done overnight.

Model	Learning Rate	Batch Size	Antelope	Duiker	Bird	Blank	Civet	Genet	Hog	Leopard	Monkey	Prosim	Rodent	Average	Epochs
efficientnet_b0	0.005	4	41.99%	79.88%	50.79%	90.22%	81.52%	94.76%	66.05%	75.06%	72.53%	10			
efficientnet_b0	0.05	4	2.23%	14.50%	35.73%	75.15%	35.87%	68.56%	74.74%	27.90%	41.84%	10			
efficientnet_b0	0.01	4	43.81%	70.71%	54.83%	87.17%	89.13%	86.03%	62.11%	74.32%	71.01%	10			
efficientnet_b0	0.005	16	65.11%	79.59%	60.45%	94.30%	87.50%	90.17%	72.05%	75.56%	78.09%	10			
efficientnet_b0	0.005	64	44.22%	63.61%	45.17%	88.59%	76.63%	83.19%	61.73%	68.17%	10				
efficientnetv2-b0	0.005	16	72.41%	82.85%	48.09%	88.80%	90.76%	90.83%	67.91%	78.27%	77.49%	10			
efficientnetv2-b0	0.005	16	76.47%	71.89%	42.25%	86.56%	86.41%	84.50%	74.53%	67.90%	73.81%	10			
efficientnet_b1	0.005	16	68.97%	85.80%	49.66%	91.65%	88.59%	87.55%	74.53%	85.19%	78.99%	10			
efficientnetv2-b1	0.005	16	78.50%	78.70%	53.93%	93.28%	90.76%	94.32%	61.70%	74.07%	78.16%	10			
efficientnet_b1	0.005	16	72.41%	92.31%	66.07%	94.70%	94.57%	97.38%	85.30%	87.41%	86.27%	20			

**Table 2.** This table shows the experimentation process and the order in which the models and hyperparameters were tested. It also demonstrates how the final hyperparameters and model were decided on.

The model performed the best in the Leopard class with a percentage of 97.38%. This is most likely because of its easily identifiable pattern of spots that can be seen on any part of the Leopard. This allows the model to easily identify the Leopard, regardless of how much of the Leopard is shown in the image and where in the image it is shown.

The model struggled with the blank images the most. This could be because even though the image doesn't contain an animal in it, there are still various geographic features and foliage that are in the image which the model could be overwhelmed by and assume that it is an animal.

These results indicate that a lower learning rate is optimal for the most accurate identification. Although it takes longer to train the model, it overall leads to better results. These results also indicate that a batch size that balances both time to train and accuracy is the most effective.

## Conclusions

We trained a series of EfficientNet models on a camera trap dataset to identify eight different species, evaluating how different hyperparameter values for learning rate and batch size influence performance. Our findings indicate that a lower learning rate with a moderate batch size is the most effective for accurate identification. The model identified leopards, hogs, genet civets, and birds well but struggled with the blank class.

We successfully trained the final model in under eight hours on a single laptop. This resulted in a model that can accurately identify over 90% of images in four species classes. This efficiency is impressive and suggests that with longer training, the performance could improve even further (Fig 2). This emphasizes that computer vision is a powerful tool for conservationists to analyze massive datasets quickly and accurately. Our work demonstrates how tuning basic hyperparameters can lead to dramatic improvements in accuracy (Table 2). With the optimal values for learning rate and batch size, we significantly enhanced the model's performance.

## Limitations and Future Work

Although the model performed well, it struggled with certain classes. The class it had the most difficulty with was the blank class. Future work could focus on improving classification

accuracy for these challenging categories by incorporating more advanced data preprocessing techniques or exploring different model architectures.

Our results in Table 2 suggest that other models, besides our final one, have potential and may be worth experimenting with further. Additionally, exploring other hyperparameters such as dropout rates and weight decay could further enhance model performance.

One major limitation of the model is its potential variability in performance across different regions or camera sites. Although it works well for camera traps in the Tai National Park, its effectiveness in other environments is yet to be tested. Future studies could analyze the model's adaptability when it comes to different locations or environmental conditions.

In summary, our study demonstrates the vast potential that computer vision has in the wildlife conservation field. Its ability to process and classify large camera trap datasets efficiently and accurately makes it a significant asset to conservation work. By continuously refining hyperparameters and exploring new models, we can further enhance the accuracy and applicability of these tools in conservation efforts.

### **Acknowledgments**

We thank Google for providing notebooks for training the model and augmenting the data available here:

[https://colab.research.google.com/github/tensorflow/docs/blob/master/site/en/hub/tutorials/tf2\\_image\\_retraining.ipynb#scrollTo=FlsEcKVeuCnf](https://colab.research.google.com/github/tensorflow/docs/blob/master/site/en/hub/tutorials/tf2_image_retraining.ipynb#scrollTo=FlsEcKVeuCnf)

### Works Cited

- DrivenData. (n.d.). Conser-vision practice area: Image classification.  
<https://www.drivendata.org/competitions/87/competition-image-classification-wildlife-conservation/page/409/>
- Norouzzadeh, M. S., Morris, D., Beery, S., Joshi, N., Jovic, N., & Clune, J. (2020). A deep active learning system for species identification and counting in camera trap images. *Methods in Ecology and Evolution*, 12(1), 150–161. <https://doi.org/10.1111/2041-210x.13504>
- Willi, M., Pitman, R. T., Cardoso, A. W., Locke, C., Swanson, A., Boyer, A., Veldthuis, M., & Fortson, L. (2018). Identifying animal species in camera trap images using Deep Learning and Citizen Science. *Methods in Ecology and Evolution*, 10(1), 80–91. <https://doi.org/10.1111/2041-210x.13099>

## **How Was Lili Boulanger's Music a Sign of the Rise of Early Feminism in France During the Early 20th Century? By Jenny Zhu**

As a musician deeply committed to highlighting underrepresented voices, I have always been motivated to discover works by lesser-known women composers. Coming across Lili Boulanger, I was impressed when listening to her works, captivated by the complexity in structure that in my opinion, deserves much greater recognition. My own experience of performing her Nocturne for Violin and Piano allowed me to feel firsthand the depth and sensitivity she infused into her music. This performance of Boulanger's work has driven me to explore her unique position in music history—specifically, how her success as a composer was a sign of the rise of early feminist movements in France during the early 20th century.

### **Abstract**

Through the lens of Lili Boulanger's groundbreaking music, this paper explores how cultural values subtly influence accessibility and gender norms in the arts. Boulanger's achievements, set against the backdrop of early 20th-century France, highlight the intersection of music, gender, and societal expectations. Boulanger became the first woman to win the Prix de Rome in 1913 through her innovative compositional style with deep inspiration from Romantic music heritage while contributing to the broader French feminist movement. By embracing an inquisitive approach, including an analysis of her Nocturne for Violin and Piano (1911), I seek to highlight the importance of Boulanger's contributions not just to music, but to the feminist movement as a whole. This research celebrates her legacy and calls for greater recognition of women's contributions in historically male-dominated fields, underscoring the importance of accessibility and equity in cultural spaces.

### **Boulanger Sisters' Background**

Born on August 21, 1893 in Paris, Lili Boulanger (1893-1918) grew up in a musically cultured and well-connected family, with their father, Ernest Boulanger (1815-1900), being a successful opera composer, choral director, and singing teacher, won the Prix de Rome for his cantata "Achille" in 1835 (Find a Grave n/d, n/p). Ernest Boulanger had developed extensive musician connections within the musical sphere, like with Fauré, Massnet, and Saint Saens (Find a Grave n/d, n/p). Nadia Boulanger (1887-1979), Lili's older sister, intensely studied music since age five, and became a well-known composer, conductor, and music educator, studying at the Paris Conservatory and winning the second prize at the Prix de Rome in 1908 for her cantata "Le Sirène" (Potter, 1999, p.536). When she was only three, Lili was diagnosed with "intestinal tuberculosis," or Crohn's disease, and her health concerns led to a point where she was unable to consistently study music until she was 16 (Anderson, 2018, n/p). Nadia always took care of Lili and won as many prizes as possible to support her mother and sister after her father retired (Nones 2023, n/p). Even after their father's death in 1900, the Boulanger sisters continued to thrive in their culturally rich environment, continuing musical family tradition with their

mother's efforts to foster relationships between created artists (Williamson 2001, p. 4). After reaching the final round of the Prix de Rome in 1909 with her cantata *Roussalka*, Nadia began her teaching career that would last almost until her death in 1979, with Lili being her first student (Potter, 1999, p.537). Throughout her career, Nadia mentored many of the most renowned conductors and composers in the 20th century at the Paris Conservatory, such as Leonard Bernstein and Aaron Copland, as well as being the first woman to conduct major orchestras like the New York Philharmonic and Boston Symphony (La Phil, n/d.). Although Lili's life was cut short due to medical concerns, she became a gifted composer and multi-instrumentalist up until her death in 1918.

Nadia Boulanger is often regarded as one of the most influential women in classical music history, although she likely would not have wanted her achievements to be defined by her gender (Pentreath, 2020, n/p). Her success was not limited by being a woman, as she once said, "I've been a woman for a little over 50 years and have gotten over my initial astonishment. As for conducting an orchestra, that's a job where I don't think sex plays much part" (Pentreath, 2020, n/p). In addition to her remarkable career as a teacher, she was a trailblazer, becoming the first woman to conduct many major orchestras in the United States and Europe, including the BBC Symphony, Boston Symphony, Hallé Orchestra, and New York Philharmonic (Burton-Hill, 2017, n/p). Boulanger was also a pivotal mentor to Igor Stravinsky and a passionate advocate for his music, promoting it even when it faced skepticism (Burton-Hill, 2017, n/p). She was instrumental in presenting many significant premieres.

Despite her vast impact across multiple musical genres—classical, jazz, tango, funk, and hip-hop—Boulanger remains relatively unknown outside classical music circles (Burton-Hill, 2017, n/p). It is unlikely that a man with a similar influence on 20th-century music would be so overlooked. Boulanger was known for her strong, uncompromising, and dynamic personality. She was charismatic, loyal, and passionate, yet also complex, as she often grappled with feelings of jealousy toward her younger sister Lili (Burton-Hill, 2017, n/p). Both sisters began studying composition at the Paris Conservatoire at the age of 10, but Nadia did not gain the same recognition as a composer (Burton-Hill, 2017, n/p). After Lili's death, Nadia dedicated much of her career to preserving and promoting Lili's music, ensuring her works continued to be performed alongside more established compositions (Burton-Hill, 2017, n/p).

Living in both a privileged and musically gifted environment, Nadia describes her the ambiance in the Boulanger home and acknowledges her sister's musical talent in a conversation with Bruno Mosnsaingeon, French filmmaker, writer, and violinist:

At home everyone made music; music was the starting point and the center of our life. My younger sister Lili played music; we were extraordinarily close and attached but quite independent. She roamed about, musically; playing the piano a little, the violin a little, the organ a little, composing. She already had ideas, some of which were developed in a very moving way. I believe that her whole talent was rooted in her first knowledge of grief. When our father died, she was six years old. And at six she understood what death



was; that it is the grief of surviving someone you love. She never forgot that up to her own death; she never forgot any detail of our father's life. . . . She was so gifted that, when still a baby, at two and a half, she used to sing all the time. Moreover, much later, Fauré quite often used to come to accompany her gladly because she could sight read a melody of which she might not have been expected to understand anything but of which she seemed to understand everything (Monsaingeon, 1985, 21-22 in Williamson, 2001, 3).

Recognizing Lili's talents in composition, Nadia reconsiders her own path. Although Nadia had initially sought a prominent position as a composer, her sister's undeniable talent led her to become Lili's primary tutor, giving up her own compositional aspirations to support Lili's burgeoning career (Burton-Hill 2022, n/p). As a result, Nadia shifted her focus to conducting and performance, thus avoiding direct comparisons to her sister's more lauded compositional skills (Burton-Hill, 2017, n/p). Lili's victory in the Prix de Rome not only highlighted her own extraordinary capabilities but also underscored the shifting gender dynamics within the French musical establishment, challenging traditional gender roles and marking an early triumph for feminist ideals in the arts.

Generally, though, both Nadia's and Lili's compositional styles are influenced by and have similar literary tastes as Debussy (Potter, 1999, p. 537). For example, in Nadia's setting of Grandmougin's poem *Les sirenes*, she absorbs his musician style through use of pedal points and modally tinged melodic lines (Potter, 1999, p. 538). However, the work remains unpublished, which implies that she had lost interest and self-confidence in the project that took too long to complete. Similarly, Lili's setting, composed six years later, is evidently influenced by Debussy and Nadia, and can be considered more satisfying than Nadia's because of a surer sense of form (Potter, 1999, p. 538). During Lili's life, Debussy dominated the musical scene in Paris. Although aspects of Debussy can be recognized in Lili's work, a large number of her works shows her fascination with pedal points and ostinati, which was largely un-Debussyan (Palmer, 1968, p. 227). Additionally, with Lili's composition style including techniques of other great composers at that time, including Fauré and Wagner, coupled with her own creativity, Lili created her own unique musical language (Williamson, 2001, n/p).

### **The First Woman to Win the Prix De Rome**

Established in 1663, the Prix de Rome was a prestigious scholarship offered by the French Académie des Beaux-Arts, extending to musical composition in 1803 (Anderson, 2018, n/p). Despite the French Revolution's promise of "universal rights of men," female artists faced significant barriers in competing for this award even though there were some in the Académie during the Ancien Regime, reflecting broader societal limitations on women's participation in public life (Anderson, 2018, n/p). The concept of Republican Motherhood, reinforced a gendered division of spheres, allowing women participation only in "private" musical settings like salons while excluding them from "public" domains such as the Académie, symphony halls, and opera

houses (Anderson, 2018, n/p). In 1903, when Minister Chaumié opened the Prix de Rome to women, resistance from Académie members was shown through the blocking of Juliette Toutain's participation as the first female artist (Anderson, 2018, n/p). Unfortunately, Toutain's social class, scandal, and engagement, such as the rule that contestants must be unmarried, led her to an unsuccessful attempt in the competition (Anderson, 2018, n/p).



Fig. 1: Contestants of the 1904 Prix de Rome.

After their father's death in 1900, Nadia Boulanger dedicated herself to competing in the Prix de Rome, entering the prestigious competition in 1906 and 1907 but fell short in both attempts (Anderson, 2018, n/p). In the 1908 competition, Nadia submitted an unconventional piece scored for string quartet, using an instrumental head motif by Saint-Saëns instead of the expected choral fugue for four voices (Anderson, 2018, n/p). Camille Saint-Saëns even accused her in a letter of trying to "create a sensation" rather than adhering to "naturalness and simplicity," and Nadia had little supporters after flaunting the rules (Leonard, 2007, p.89). Although seen as unorthodox by the competition's traditional expectations, the judges recognized Nadia's high skill and advanced her to the second round (Anderson, 2018, n/p). However, in the final round, despite her work being the most competent, Nadia only received the Deuxième Second Grand Prix (Anderson, 2018, n/p). Nadia's persistence and innovative spirit, although not always rewarded by the competition, marked her as an assertive composer in the male-dominated musician world at the time.

In contrast, Lili Boulanger became the first woman to win the Prix de Rome with her cantata "Faust et Hélène" in 1913, with Nadia's efforts paving the way for her sister's success. The piece, drawing on a specific moment in Goethe's retelling of Dr. Faust, when he asks to see Helen of Troy, captured the judges' attention and marked a significant milestone for women in

music (Wagstaff 2020, n/p). Given the conservative nature of the judges, it is improbable they would award the first prize to a woman perceived as challenging authority. Lili's greater compositional knowledge, calm demeanor, modesty, and softer personality—contrasting with her sister Nadia's often irascible nature—likely increased her chances of success (Leonard, 2007, p.89). Nonetheless, the participation of women in the Prix de Rome during this period reflects their growing presence in the musical world. As the first woman to receive this prestigious award, Lili Boulanger quickly gained international recognition, with newspapers describing her as "an international celebrity almost overnight" (Williamson, 2001, p.5). By July 7, 1913, accounts of her triumph had reached the United States; shortly thereafter, Lili signed a publishing agreement with the Italian publishing house Ricordi (Williamson, 2001, p.5). As part of her award, Boulanger was granted residency at the Villa de Medici in Rome, but she was obliged to return to Paris due to the outbreak of World War I (Williamson, 2001, p.5). Nevertheless, Lili's victory in the Prix de Rome broke the glass ceiling for women composers (Wagstaff 2020, n/p).

### **Shifting French Atmosphere**

On August 3rd, 1914, France entered into the First World War. Already suffering demographically, due to the low birth rate at the end of the 18th century, coupled with a rural, stagnant, and aging population, France lost around 1.7 deaths from the war, nearly 4.35% of their population (Chemins de mémoire n/d, n/p). Specifically in the military, over 4.2 million soldiers were wounded, over 500,000 soldiers were imprisoned, and approximately 1.4 million soldiers were killed (Beaupré, 2014, n/p) .

During that year, France's role in the war began with a number of disastrous battles, with August and September being the deadliest months (Beaupré, 2014, n/p). For instance, after the Battle of the Marne, only after six weeks of combat, the French army had already suffered around 100,000 losses (Beaupré, 2014, n/p). Over this six week period, the average number of French losses rose from 900 soldiers per day to approximately 2,400 soldiers per day (Beaupré, 2014, n/p). For soldiers, the war led to hardships in their daily life and during combat. Typically, everyday life for French soldiers included long walks and tiresome labor, usually in harsh outdoor conditions, and also periods of inactivity, discouragement, and boredom in diseased filled trenches (Beaupré, 2014, n/p). In fact, since the war significantly changed the physical appearances of men, French soldiers were nicknamed "poilu," meaning "hairy," with their bodies symbolizing manliness (Beaupré, 2014, n/p).

For civilians during the first weeks of the war, German atrocities led to over 900 deaths in the French people and dozens of destroyed villages, especially in the Meurthe-et-Moselle department. German atrocities led to over 900 deaths in the French people and dozens of destroyed villages, especially in the Meurthe-et-Moselle department. As a result, the population increased the flow of refugees, out of terror, and also established a unified sense of French support for its soldiers. From 1914 until 1918, "German barbarism" and references to the atrocities of 1914 were overall the most popular topics published in propaganda (Beaupré, 2014, n/p). Additionally, the massive mobilization for the war effort - 1.7 million men - separated

French families, with many families unfamiliar with the lack of manpower in their households, and the poorer living conditions due to material shortages (Beaupré, 2014, n/p). The winter of 1916-1917, for example, led to a shortage in heating supplies, food, and an increase in prices, ration cards, and tickets on the home front; prices, after remaining relatively stable during 1916, increased by 25% in the first quarter of 1917 (Beaupré, 2014, n/p). At home, civilians learned about the sufferings on the front lines through publications of war literature and narratives, which became popular genres (Beaupré, 2014, n/p). The increase in wounded men coming home increased the need for hospitals, and as a result, primary and secondary schools, barracks, and even homes turned into hospitals (Beaupré, 2014, n/p).

During World War I, French women were not allowed to enlist in the military as soldiers, but nevertheless, over 200,000 French women worked in the French Army, often working the same jobs as soldiers (ORR, 2017, pp. 42-75). Because of France's shortage of soldiers, heavy losses of 1.4 combat deaths, and weak demographics, women were vital to the Allied victory of the war. In fact, As soon as war broke out in 1914, France's prime minister Rene Viviani called on women to "replace in the workplace those who are on the battlefield" (France 24, 2018, n/p). However, many French male officers showed hostility and mistrust to women's service as civilian employees during and shortly after the war (ORR, 2017, pp. 42-75). Over the 1920s, though, many military leaders had begun to accept the capabilities of women to take on these roles and considered them as part of the military community (ORR, 2017, pp. 76-110). Moreover, conflicts between military generals and civilian leaders over army organization, laws regarding the length of conscript service, and the extension of certain political rights to soldiers expedited the emergence of women's military identity (ORR, 2017, pp. 76-110).

According to historian Richard J. Evans, during the late 1860s to late 1880s, anticlericalism was a driving force behind the emergence of bourgeois feminism in France (Evans, 1982, p. 947). However, during this era, the feminist movement was both small and poorly supported, but after this era, the feminist movement in France included Protestant women leaders (Evans, 1982, p. 948). This progression shows that the anticlerical aspect of French feminism was only a phase ended by the 1890s, before becoming more of a "moralistic orientation" (Evans, 1982, p. 948). A major shortcoming of the anticlerical phase of French feminism was the lack of support for these feminists, even amongst those who supported republican and anticlerical ideals (Evans, 1982, p. 948). Furthermore, early feminists in France remained a minority within the larger social and political movement they belonged to.

Compared to other areas in France, especially some northern cities, Paris was largely undamaged from World War I; the city did not suffer from German occupation either (Parisian Fields 2015, n/p). Instead, Parisians mainly felt the effect of the war through shortages and privations, mourning over sons' and husbands' deaths from trench warfare, and an ambiance of uncertainty heightened each subsequent year (Parisian Fields, 2015, n/p). Although Paris held lots of wounded soldiers and refugees, life continued to go on in cafés, shops, and luxuries offered in the black market for those who could afford them; most people, however, experienced a dismal period of waiting with uncertainty, since news was scarce (Parisian Fields 2015, n/p).

Nevertheless, the wartime “Comité-Franco-Américain,” a society of young French musicians that included sisters Nadia and Lili Boulanger, had sought to provide comforts in the form of morale-boosting events and newsletters for troops drawn from the cultural life of the city and country (Leonard, 2007, p. 19).

### **Nadia Boulanger’s Efforts**

Nadia Boulanger continued teaching at the Paris Conservatory during the war, providing stability and continuity in education for her students. Specifically, Boulanger mentored many young women, offering guidance, encouragement, and professional connections. She was known to have a teaching style that combined honest feedback and a dedication to exploring the curiosity and voice of each pupil (Salazar 2017, n/p). Additionally, Nadia exhibited an impressive work ethic; Nadia says: “Anyone who acts without paying attention to what he is doing is wasting his life. I’d go so far as to say that life is denied by lack of attention, whether it be to cleaning windows or trying to write a masterpiece” (Salazar 2017, n/p). According to Quincy Jones, Nadia Boulanger had a singular way of encouraging and eliciting each student’s own voice – even if they were not yet aware of what that voice might be (Salazar 2017, n/p).

Nadia Boulanger organized a series of concerts in Paris to support the war effort, reflecting her commitment to both her country and the arts during a period of intense hardships (Brooks, 2013, pp. 19-40). These concerts not only provided cultural sustenance but also served as acts of patriotism, helping to maintain morale in war-torn France, and the impact of these events extended beyond the immediate context, as they forged important artistic collaborations, including her relationship with Igor Stravinsky (Francis, 2010, p. 2). Stravinsky’s involvement, particularly through the performance of his works, showed the role of music in sustaining cultural identity and resilience during the war (Burton-Hill, 2017, n/p). Boulanger’s efforts to bring his compositions to the public during this time significantly contributed to the preservation and promotion of modernist music, further expanding Stravinsky’s influence in the European and American musical landscapes, even long after the war ended (Francis, 2010, p.3).



Fig. 2: Nadia Boulanger and Igor Stravinsky, 1937.

Despite Lili's fragile health and social pressure during that time, she managed to compose over 50 works during her short career; however, her compositions are often overshadowed by Nadia's because of her death at such a young age (Wagstaff 2020, n/p). Nevertheless, in the years after Lili's death in 1918, her sister Nadia tirelessly promoted her music as a performer and conductor, organizing several memorial concerts to honor her sister's work (Burton-Hill 2022, n/p). For instance, shortly after Lili's death, Nadia arranged a significant memorial concert at the Salle Gaveau in Paris in March 1919, featuring many of Lili's compositions (Brooks, 2013, n/p). Additionally, Nadia played a crucial role as an artistic advisor in the first-ever recording of Lili's most significant works for chorus and orchestra, helping her compositions reach a broader audience (Nones 2023, n/p). This recording, featuring the Chorale Elisabeth Brasseur and the Lamoureux Orchestra under Igor Markevitch's direction, was released on the Everest label in 1960 (Nones 2023, n/p). Since its release, the recording has continuously stayed in the catalog in the 60 years since (Nones 2023, n/p). In a concert at the Théâtre du Châtelet in 1921, Nadia conducted a concert that included Lili's "Pie Jesu," "D'un soir triste," and "D'un matin de printemps," showcasing the depth and variety of Lili's work (Rogers, 2021, p.230). This is merely one of the over 70 performances of Lili's work that Nadia was involved in between 1919 and 1973, either as a performer, conductor, or organizer/overseer of the concert (Rogers, 2021, pp. 229-236). Moreover, Nadia included Lili's music in her programs when she conducted orchestras and music festivals around the world, like in Belgium and the United States, fostering international exposure.



Fig. 3: First-recording of Lili Boulanger's Significant Chorus and Orchestra works, 1960.

As a prominent teacher at the American Conservatory in Fontainebleau beginning in 1921, Nadia included Lili's music in the conservatory's annual concerts, which not only kept Lili's music alive but also exposed it to international students and musicians. In the Conservatory's inaugural year, 25 concerts were given by the vocal ensemble students of the Conservatoire Américain, either as singers or piano accompanists; performed works included Lili's compositions (Leonard 2007, 39). Boulanger regularly introduced her students to fresh music pieces and arranged for them to meet significant composers. Some of her American students, like Louise Talma, showed a clear influence from French music overall, particularly from Lili Boulanger and Fauré (Leonard 2007, 63). In 1939, Nadia helped establish the Lili Boulanger Memorial Fund at the American Conservatory in Fontainebleau including the Lili Boulanger Memorial Prize to support the performance and publication of Lili's works; winners of the prize were presented at concerts at the American embassy and at Radio-Diffusion Française (Leonard 2007, 110). Lastly, Nadia often performed Lili's compositions herself, both as a conductor and pianist.

Nadia Boulanger played a crucial role in promoting her sister Lili's work both in France and internationally, yet Lili's compositions are still relatively underappreciated compared to her sister, and contemporaries like Fauré and Debussy. Despite drawing inspiration from these leading composers, Lili's innovative approach within traditional frameworks showcases her exceptional talent and highlights the broader shifts in gender roles during that period, with women extending beyond domestic spheres to the work force and even into the arts, exercising greater creative freedom.

Lili Boulanger's victory in the Prix de Rome not only marked a personal achievement, but also a shift in female participation and recognition in French performing arts. Her inability to marry due to poor health motivated Lili to study composition seriously in order to win the Prix de Rome with "Faust et Hélène," the piece showing the lasciviousness of women despite Lili's lack of romantic life (Goss 2013, n/p). This victory catalyzed an ongoing period of creativity, where Lili continuously overcame her physical conditions to compose innovative works of beauty and mystery (Goss 2013, n/p). In the following section, I will analyze Lili Boulanger's Nocturne for Violin and Piano, illustrating her sophisticated compositional technique that reinforces her high level of expertise.

### **Lili Boulanger's Nocturne for Violin and Piano (1911)**

First published in 1911, Lili Boulanger's "Nocturne" for Violin and Piano in F Major showcases delicate yet unresolved melodies intertwined with whole-tone harmonies, representative of her typical style. This piece draws inspiration from the concept of a nighttime scene. The piano accompaniment starts softly as the violin introduces the melody; as the violin becomes more agitated and intense, the piano grows louder, creating a dynamic interaction. While it is typical for a nocturne to evoke a nighttime scene or memory, Lili's unique harmonic language stands out, especially in her use of unresolved harmonies. Throughout the piece, the

piano follows the violin's movements closely, mirroring its delicacy and intensity and thereby shaping the piece's dramatic structure and voice.

The piece is structured in an ABA' format. In Section A, the piano establishes a repeated note C as an ostinato, creating a jumping rhythm with a wide range that encircles a constantly evolving melody. This approach brings the piano to an equal footing with the violin, enveloping the latter in a mysterious, somewhat unclear atmosphere, like darkness. In contrast, the B section allows the violin to break free from this circular motif; this section is marked by a more chromatic and minor tonal color, while the transition back to the A' section reintroduces a major chord, with repeated C octaves in the piano. Even though the structure follows a mostly traditional form and nocturne style, Lili introduces innovative elements within this framework.

**Deux Morceaux pour Violon et Piano**  
 1. Nocturne  
 (1911)  
*à ma chère Marie Danielle Parenteau*  
 Lili Boulanger  
 1893 - 1918

The image shows the first system of the musical score for Section A of Nocturne. It consists of three staves: a single treble clef staff for the violin and a grand staff (treble and bass clefs) for the piano. The tempo is marked 'Assez lent' with a metronome marking of a quarter note equal to 63. The piano part features a prominent ostinato of repeated octaves of the note C, which is highlighted by a red box labeled 'A'. The violin part has a melodic line with some notes circled in red. Performance markings include 'pp très doux et soutenu', 'p doux', 'à l'aise', 'mf', and 'expressif'. A '2 Ped.' instruction is present at the beginning.

Fig. 4: mm. 1-8, Section A of Nocturne.

The piece begins with a sparse accompaniment, opening with repeated octaves in the piano that rise and fall throughout the section. As illustrated in Figure 2 (mm. 1-13), Section A



features a pedal point on the note C, which is repeated across different octaves and with the same pattern. The violin, although following a repetitive structure, presents a melody that climbs and falls with direction. For instance, in mm. 3-4, the violin reaches B-flat, and in mm. 7-8, it ascends to D. This creates a balanced relationship between the violin and piano, both parts in unity and of equal importance. However, the continuous repetition of C raises the question of why Lili selected the note C to repeat, if the key of the piece is F major, making C the dominant?

Although Lili employs a traditional harmonic scheme similar to that of many late 19th-century German composers, she incorporates more impressionistic colors through frequent use of minor seventh chords. Unlike Debussy, who often puts less importance on traditional harmonic structures, incorporates more mm7/minor colors and less tonic and dominant compared to Lili. Lili's compositions retain some conventional elements, while still exploring innovative harmonic techniques. Although her style reflects some influence from Debussy, she does not go further beyond his innovation and does not completely abandon the German musical tradition.

The image shows a musical score for Section B of Nocturne, measures 21-27. The score is in F major and 3/4 time. It features a piano accompaniment with a constant C pedal point in the bass and a violin melody. The violin melody consists of eighth-note runs that ascend and then descend. Performance markings include 'augmentez', 'et animez', 'rit.', 'sans lenteur', 'ff', 'très vibrant', and 'revenez au mouve-'. Red and purple circles highlight specific notes in the violin part, and red brackets group the eighth-note runs.

Fig. 5: mm. 21-27, Section B of Nocturne.

Measure 14 marks the transition to the B section, where the repetition of the note C is temporarily abandoned, introducing a more chromatic and dissonant texture; this section stands in contrast to the brighter, more open color of the A section. In measure 30, the pedal point on C returns, marking the beginning of the A' section. However, the violin introduces the note C four measures earlier, in measure 26, building virtuosic intensity as the texture thickens, the violin beginning to climb. Starting in measure 21, the violin embarks on a sequence of seven tuplets with each reaching in an ascending series of half notes—G, A, B-flat, B natural, and finally C—gradually building tension before resolving to the top, note C.

The image displays a musical score for Section A' of Nocturne, spanning measures 31 to 42. The score is presented in four systems, each containing a vocal line and a piano accompaniment. The key signature is one flat (F major), and the time signature is 3/4. The piano part features a prominent pedal point on the note C in the bass register. The vocal line includes lyrics such as "en dehors, doucement chanté", "cédez", "un peu en dehors", and "en s'éloignant". The score is divided into four systems, with red boxes highlighting specific passages in measures 31-33, 34-36, 37-39, and 40-42. The first system (measures 31-33) shows the vocal line with lyrics "en dehors, doucement chanté" and the piano part with a dynamic marking of *pp*. The second system (measures 34-36) includes the instruction "mettez la sourdine" and the vocal line with lyrics "cédez" and "un peu en dehors", with a dynamic marking of *pp* and the instruction "très doux". The third system (measures 37-39) features the vocal line with lyrics "en s'éloignant" and "cédez", and the piano part with a dynamic marking of *pp*. The fourth system (measures 40-42) shows the vocal line with a dynamic marking of *ppp* and the piano part with a dynamic marking of *ppp*.

Fig. 6: mm. 31-42, Section A' of Nocturne.

Interestingly, there is no F chord in the A section, as seen in Fig. 4; instead, the emphasis is on the dominant note C, yet not the C chord. The first F chord appears only towards the end of the piece, in measure 36, with the piece ultimately resolving to F in the final two bars. This F chord foreshadowed in measure 32 creates a sense of tension and surprise, similar to the harmonic structures found in works by composers like Schumann's *Dichterliebe* and Brahms' late piano works. The choice of a C pedal point can be seen as analogous to a narrative journey—starting from a distant place, venturing further away, and finally finding the way back home. Even though Lili's *Nocturne* is in F major, the tonic F does not appear until the end, contributing to an overall feeling of anxiety and unresolvedness. This might suggest Lili's intention to evoke the unresolved, anxious emotions associated with nighttime. Despite this, the piece is meticulously organized; by setting a key of F major without immediately employing a tonic function, Lili demonstrates her advanced compositional technique.

## **Conclusion**

Lili Boulanger's legacy today is one of profound, though often understated, significance in the classical music world. Despite her brief life, Lili's compositions are celebrated for their emotional depth, innovative harmonies, and unique voice, distinguishing her as a remarkable talent of her time. Her music, characterized by its emotional and technical sophistication, serves as a testament to the evolving role of women in the arts and public life. Furthermore, sister Nadia Boulanger's dedication to promoting Lili's music played a crucial role in preserving her legacy, introducing Lili's compositions to wider audiences internationally. Today, Lili Boulanger is recognized as a pioneering female composer, and her contributions continue to be studied and admired in both academic and performance circles. Her legacy is particularly significant in discussions about women in classical music, serving as an inspiration for female composers and musicians striving for greater recognition. By breaking the gender barriers of her time and winning the prestigious *Prix de Rome*, Lili Boulanger challenged the male-dominated world of classical composition and proved that women could achieve creative and intellectual heights previously reserved for men. While some framed her as a tragic figure because of her medical struggles or solely as a feminist icon, Boulanger broke the glass ceiling through her endless dedication to composition, reflecting the broader societal shifts that were reshaping women's roles in early 20th-century France.

## Works Cited

- Anderson, Will. *Lili Boulanger and the Prix de Rome*. Untune the Sky, 27 Jan. 2018, [untune-the-sky.com/2018/01/26/lili-boulanger-prix-de-rome/](http://untune-the-sky.com/2018/01/26/lili-boulanger-prix-de-rome/).
- Beaupré, Nicolas. "France." International Encyclopedia of the First World War, October 8, 2014. <https://encyclopedia.1914-1918-online.net/article/france/>.
- Brooks, Jeanice. "Nadia Boulanger between the Wars." Chapter. In *The Musical Work of Nadia Boulanger: Performing Past and Future between the Wars*, 19–40. Musical Performance and Reception. Cambridge: Cambridge University Press, 2013. <https://www.cambridge.org/core/books/abs/musical-work-of-nadia-boulanger/nadia-boulanger-between-the-wars/A64A53EC1F0D3A6E35078BEB42089A0A>.
- Burton-Hill, Clemency. "The Greatest Music Teacher Who Ever Lived." BBC News, April 19, 2017. <https://www.bbc.com/culture/article/20170308-the-greatest-music-teacher-who-ever-lived>.
- Evans, Richard J. "Feminism and Anticlericalism in France, 1870-1922." *The Historical Journal* 25, no. 4 (1982): 947–49. <http://www.jstor.org/stable/2638639>.
- "Ernest Henri Alexandre Boulanger (1815-1900)." Find a Grave. Accessed September 13, 2024. [https://www.findagrave.com/memorial/126797944/ernest-henri\\_alexandre-boulanger](https://www.findagrave.com/memorial/126797944/ernest-henri_alexandre-boulanger).
- France 24. "World War I, a Revolution for Women." France 24, October 30, 2018. <https://www.france24.com/en/20181030-world-war-i-revolution-women>.
- "Get to Know: The Boulangers." LA Phil. Accessed 10 Sept. 2024. [www.laphil.com/about/watch-and-listen/get-to-know-the-boulangers#:~:text=Displaying%20influences%20of%20Faur%C3%A9%20and,instrumentation%2C%20and%20skillful%20text%20setting](http://www.laphil.com/about/watch-and-listen/get-to-know-the-boulangers#:~:text=Displaying%20influences%20of%20Faur%C3%A9%20and,instrumentation%2C%20and%20skillful%20text%20setting).
- Goss, Thomas. "Lili Boulanger in Her Own Right." *Orchestration Online*, 21 Sept. 2013, [orchestrationonline.com/lili-boulanger-in-her-own-right/](http://orchestrationonline.com/lili-boulanger-in-her-own-right/). Accessed October 1, 2024.
- Leonard, Kendra Preston. *The Conservatoire Américain - A History*. Lanham, Maryland: The Scarecrow Press, Inc., 2007. <https://www.census.gov/history/pdf/reperes112018.pdf>.
- Luthardt, Rachel. "Influential Women Composers: The Legacy of Antonia Brico and Nadia Boulanger." FUSION. Accessed September 13, 2024. <https://www.fusionmagazine.org/influential-women-composers-the-legacy-of-antonia-brico-and-nadia-boulanger/>.
- Nones, Philip. "Famed Musician and Educator Nadia Boulanger's Perceptive Description of French Composer Florent Schmitt's Artistry (1925). | Florent Schmitt." Florent Schmitt, December 1, 2023. <https://florentschmitt.com/2023/12/01/famed-musician-and-educator-nadia-boulangers-perceptive-description-of-french-composer-florent-schmitts-artistry-1925/>.
- ORR, ANDREW. *Women and the French Army during the World Wars, 1914–1940*. Indiana University Press, 2017. <https://doi.org/10.2307/j.ctt1zxx0r0>.
- Palmer, C. (1968). Lili Boulanger 1893-1918. *The Musical Times*, 109(1501), 227–228. <https://doi.org/10.2307/953482>.
- Parisian Fields. "Paris in the First World War." Parisian Fields, November 8, 2015. <https://parisianfields.com/2013/10/27/paris-in-the-first-world-war/>.

- Pentreath, Rosie. "Meet Nadia Boulanger, The Inspiring Woman Behind the 20th Century's Greatest Composers." *Classic FM*, June 23, 2020.  
<https://www.classicfm.com/discover-music/women-in-music/nadia-boulanger-composer-conductor-music-teacher/>.
- Potter, Caroline. "Nadia and Lili Boulanger: Sister Composers." *The Musical Quarterly* 83, no. 4 (1999): 536–56. <http://www.jstor.org/stable/742616>.  
[https://digital.library.unt.edu/ark:/67531/metadc3022/m2/1/high\\_res\\_d/dissertation.pdf](https://digital.library.unt.edu/ark:/67531/metadc3022/m2/1/high_res_d/dissertation.pdf),
- Rogers, Jillian C. "In Search of a Consolatory Past: Grief and Embodied Musical Memory," in *Resonant Recoveries: French Music and Trauma Between the World Wars*. New York, Oxford, 2021; online edn, Oxford Academic, 18 Feb. 2021.
- Rosenstiel, Léonie. *The Life and Works of Lili Boulanger*. Farleigh Dickinson University Press, 1978.
- Salazar, Eric. "Nadia Boulanger: Mentor of Modern Composition." *Classical Music Indy*, December 28, 2017. <https://classicalmusicindy.org/nadia-boulanger-mentor-of-modern-composition/>.
- Wagstaff, Malinda. "Musical Women of the Suffrage Movement." *WXXI Classical*, February 28, 2020.  
<https://www.wxxiclassical.org/blog/2020-02-28/musical-women-of-the-suffrage-movement>.
- "1914: A Demographically Weakened France." *Chemins de mémoire*. Accessed September 13, 2024.  
<https://www.cheminsdememoire.gouv.fr/en/1914-demographically-weakened-france>.
- Williamson, Deborah. "A Performer's Analysis of Lili Boulanger *Clairières dans le ciel*: Song Cycle for High Voice and Piano; a Lecture Recital Together with the Role of Blanche in *Dialogues of the Carmelites* by F. Poulenc and Two Recitals of Selected Works by H. Purcell, F. Schubert, S. Prokofieff, E. Chausson, W. A. Mozart, R. Schumann and G. Fauré." Dissertation, University of North Texas, 2001.

## **The Credit Scoring System’s Relationship with Financial Inclusion in Developing Countries By Khanh Nam Nguyen (student), Alan Zhang (mentor)**

### **Abstract**

Many developing countries, such as Vietnam, aim to reach their financial inclusion goals to help poor people access formal and credible sources of credit. However, is financial inclusion achievable without tampering with the sustainability of creditors’ portfolios or is it inevitable between the “affordability” and “sustainability” of credit? This research discovers the importance of both the “sustainable” and “affordable” aspects of credit. It offers insight into the case study of Vietnam, Kenya, and India to discover this inverse relationship that countries must face when seeking financial inclusion.

**Keywords** Economics; Credit Ranking; Developing Countries; Finance; Financial Inclusion

### **Introduction**

Financial inclusion has been established as one of many developing economies’ primary goals. Vietnam, for instance, approved a national strategy towards financial inclusion by 2030 by targeting non-cash transactions to increase by 20-25% annually.<sup>1</sup> This requires credit transactions to become more widespread and access to loans to become easily accessible for rural and poor people; however, the current credit scoring system employs rather strict selective methods to ensure the safety of lenders’ portfolios that can easily single out many potential borrowers. This strict selective method ensures the “sustainability” of credit transactions.

The credit scoring system is inextricably linked to the achievement of financial inclusion, but whether it is a supportive or negative relationship is yet to be determined. An operational and efficient system allows people straightforward access to lines of credit and affordable financial services without putting in major collaterals. However, the current system can be detrimental to achieving financial inclusion, especially in developing economies, where access to and knowledge of financial services still needs to be improved. Credit scoring as it stands, according to the multidimensional definitions of financial inclusion that will be listed below, effectively needs more features that dissuade financial inclusion goals in many developing economies. Under most definitions, financial inclusion is the access to different financial services of different demographic populations. This definition has recently been broadened to refer to financial services’ access, use, cost, and quality of financial services. The “access” dimension refers to the proximity and opportunity of these financial services, while the “use” dimension typically means the actual uses of the services.<sup>1</sup> Access to credit scoring systems requires proper paperwork and documentation from borrowers, which can be problematic in developing economies. Without proper access to formal lines of credit, informal credit lines are often utilized, which leads to grave uncertainty on the part of the borrower and the lender. The effect of informal credit plays a crucial role in deciding the importance of the developing credit scoring system, which is discovered later in this article.

Other multidimensional definitions of financial inclusion include “cost” of financial services, such as the monetary and nonmonetary fee of accessing these services, and “quality,” which refers to whether the financial services meet the consumers' needs and how well-informed a customer is to the service.<sup>1</sup> In terms of a credit scoring system, the “cost” of the service typically refers to the fee of accessing credit lines out of proximity to a bank or a bank’s branch. For rural citizens, the fee required to complete a trip into urban areas to access a bank to open a credit account might dissuade most citizens from opening bank accounts.<sup>2</sup> This challenge is further accentuated by the lack of trust in banking services for many uninformed citizens who decline to employ a service they need help understanding. This leads to increased gold and cash transactions in rural areas- tangible objects that could be easily understood.<sup>2</sup> The impact of the transparency of the credit scoring system is highlighted later in this article.

Finally, the “quality” of the credit services refers to the extent to which a customer’s need is met and how informed they are about the service.<sup>3</sup> This definition of financial inclusion directly contrasts with the status quo of implementing the credit scoring system. It is majorly due to the credit scoring system’s fallible nature, which will be justified through a study of the system’s history, that has prevented many borrowers from accessing lines of credit and, in the same process, hamper the lenders’ ability to manage risks in credit transactions. The flaws of the current credit scoring system are reasons that affect the “affordability” of credit lines, and we will understand why these flaws create a major issue in the relationship between “affordable” and “sustainable” credit.

All of these aspects discussed above are present in the assessment of the impact of credit scoring system on financial inclusion; however, the issue does not cease to exist simply with regards to definitions but broadens out to a greater dilemma that many financial services face: whether these services are operated at a sustainable rate for the lender, or whether these services should be provided at an “affordable” rate for the borrower, “affordable” in terms of ability to access without many obstacles.<sup>3</sup>

Despite the pressing need of many developing economies to achieve financial inclusion, one must remember the benefit of the current credit scoring system and its role in keeping loaners’ assets safe and protected. This applies to banks whose primary source of risks is credit, making the credit scoring system a powerful tool in managing assets and granting banks the ability to make rational decisions. With the implementation of a credit scoring system, loaners can avoid uncertainty in their transactions which negatively impacts their asset management, thus leading to loaners having to rely on their highly fallible decision-making ability.<sup>4</sup>

The lack of a proper credit scoring system is a challenge; for instance, in the USA, a credit scoring system allows credit card companies to make valid estimates of someone’s credit behavior, deciding whether they should grant them credit cards. On the other hand, Russia faced considerable challenges due to the absence of this system.<sup>5</sup> The availability of this system in financial sectors of developing economies certainly improves the strength of lenders’ portfolios through selective methods that often ensure calculable risks.

The current credit scoring system employs a selective process to ensure the safety of lenders' portfolios. Different methods of credit scoring systems result in either low or high selective power for the lender. However, either way, lenders can always seek to optimize the profitability of their portfolio. This might benefit the lenders, but judging from a macroeconomic view, this selective process often singles out many people who need proper credit history documentation or records. This once again raises the question of whether developing economies should opt for the "sustainability" of credit services or "affordability" for citizens of their countries.

The article discovers the benefit of the credit scoring system to loaner's portfolio sustainability through the study of the credit scoring system's development history to understand why "sustainability" is such an important aspect when considering financial inclusion. The article also elaborates on the backs of the current credit scoring system to understand why the "sustainability" of the current credit scoring system is negatively influencing the "affordability" of credits. Is there a way that both "sustainable" and "affordable" credit can exist to facilitate financial inclusion? To further discover this inverse relationship that the current credit system creates, the article employs an empirical analysis of the development of the credit scoring system, using the case study of Vietnam's credit scoring system to assess how the strict bureaucratic credit scoring can hinder poor people's ability to access credit. The article also considers Kenya's farmers' preference for formal credit and why the credit scoring system is necessary to facilitate this preference. The article also gives an insight into solutions offered by Indian economists in an attempt to enable credit "affordability" for citizens of India to assess whether both "sustainability" and "affordability" can be achieved to ensure an efficient financial inclusion achievement for developing countries. Finally, the potential trade-off that is inevitable when having both "sustainable" and "affordable" credit is highlighted through an analysis of proposed solutions by Indian economists.

## **Discussion**

### ***The benefit of the current credit scoring system on loaner's portfolios sustainability***

As the introduction mentions, "sustainability" is an important aspect of any financial service and applies to the credit scoring system. The current credit system serves loans particularly well in making rational decisions, thus making their credit transactions more sustainable and safer. The credit scoring system's calculable risk ensures that a loaner's portfolio can be safeguarded against uncertainty.<sup>6</sup> The section below will analyze the sustainable benefit that the credit scoring system provides for loaners through its development history.

The credit scoring system has been established as a tool amongst loaners to monitor and potentially convert uncertainty to risk in market economies. Under complex conditions that require outstanding cognitive demands, a loaner can comprehend the certainty of their action (about their monitoring of money or utility) and make rational financial decisions; however, human cognition is not always trusted to engage in complicated assessments of their actions



which are often wreathed with uncertainty.<sup>6</sup> Under conditions of uncertainty, information on the consequences of one's actions is highly restricted, which leaves loaners in desperate search of a method to monitor the probability of outcomes of their actions.<sup>6</sup> In a financial situation, having limited information often results in poor decision-making and can cost loaners financial loss. Loaners, therefore, highly value rational decision-making to protect their assets. This demands that loaners devise a tool to assess the risk of their actions on their behalf and which course of action leads to the best outcome ("best" is defined in terms of money or utility). Without the invention of such tools, businesses conducted through credit transactions or loans given out would be difficult to assess and sustain.

Loaners turn to third parties that can provide information on potential borrowers to resolve any uncertainties and convert them into assessable risks. The credit scoring system was created to replace any possible loan uncertainties with manageable risks. Without this creation, loaners would rely heavily on their common sense and become overly confident in their ability to make the right choice without being aware of underlying uncertainty that could leave loaners blindsided. When "there is no valid basis of any kind for classifying instances" actors must resort to estimates, creating conditions of uncertainty as mentioned above.<sup>7</sup> The most important feature of the credit scoring system is that it creates a homogenous ranking basis for loaners to compare their potential borrowers.<sup>8</sup>

The need for a credit scoring system was exemplified after the financial crisis of 1837 when credit issues became a pressing matter in the business world.<sup>8</sup> Firms began gathering customers' information to comprise a guidebook to help loaners determine the risks of giving credit to specific clients. This primitive system, although flawed and appeared unsystematic to today's researchers, still managed to "absorb uncertainty," making the information appear unambiguous to loaners.<sup>9</sup> A homogenous ranking system was created to rank customers based on reports of their credit histories, with credit ratings having ordinal categories that, as mentioned above, strengthen the ability to reduce certainty to risk.<sup>10</sup> Uncertainty management was mitigated, and loaners' portfolios received much better safeguards than before, allowing the credit scoring system to evolve into what it is today. Loaners could now rank their potential clients based on a similar foundation, making the system appear unbiased and strictly selective in favor of loaners.<sup>10</sup>

Later in the 1960s, attempting to eliminate social ties involvement in the credit lending process, Fair, Isaac, and Company devised formal credit scoring techniques that strictly took advantage of statistics to determine a consumer's credibility.<sup>11</sup> Banks were able to process much larger applications for credit cards than before without the need for loan officers, making the credit lending process more sustainable.<sup>12</sup> This extrication of the credit scoring system from the realm of social ties was also exemplified in the switch from trust (retailers' open-book accounts and community-based credit reporting) to more rational calculation (nationwide credit bureaus).<sup>12</sup> Passage of the Equal Credit Opportunity Act Amendments endorsed credit scoring systems when instructions on using these systems were explicitly included in Regulation B, which implements the Act.<sup>13</sup> In the amendments hearings, creditors argued that implementing

credit scoring systems would free credit decisions from credit evaluators' arbitrary and capricious behavior.<sup>14</sup> Furthermore, the costs to sustain the credit scoring system were significantly reduced because less skilled personnel were required, making the credit scoring system even more sustainable.<sup>14</sup> Without the heavy involvement of social ties and human error, the credit scoring system became much more trustworthy, allowing the credit card market to become sustainable.

Moving away from social ties in credit decision-making also resulted in larger changes for the credit scoring system. In the 1990s, due to low federal borrowing rates (allowing banks to borrow cheaply) and the deregulation of caps on interest rates and fees (enabling banks to raise prices on consumer loans), creditors began targeting poorer and riskier consumers who were previously denied credits.<sup>15</sup> This change requires creditors to alter their credit evaluation methods by directing greater focus to behavior in borrowing history, habits of repayment, and consumption patterns rather than characteristics (marital status, age, occupation), which created more accessible credit lines and more sustainable credit for consumers.<sup>15</sup> This change in evaluation method also created an opportunity for the switch from "control by screening" to "control by risk," as Martha Poon described.<sup>16</sup> "Control by screening" previously involved only two categories, which are creditworthy or not creditworthy' while "control by risk" allowed for a wide range of calculable risks for creditors so that different rates could be charged according to the risks calculated. By the 2000s, defaults were being predicted not only on credit history and income but also on the trivial spending habits of consumers to determine whether credit could be granted.<sup>16</sup> This change in the credit scoring system was significant for aiding in managing loaners' assets, as the ultimate goal of the latter is to convert uncertainty into calculable risks to protect and safeguard their assets.

### ***Credit scoring issues affecting the "availability" and "affordability" of credit***

Having discussed the sustainability of the credit scoring system, the question this paper seeks to address is how the current credit scoring model negatively impacts potential consumers in developing economies. Credit scoring is a prerequisite for financial inclusion.<sup>17</sup> As discussed above, the credit scoring system's nature is highly selective in providing creditors with safeguards against incalculability. However, the selective nature of the credit scoring system also means that it often singles out many people in developing economies.

The affordability of credit depends on the availability of characteristics on which the credit scoring system is based. These characteristics are usually credit history, updated IDs, and household certificates. Doncha Marron mentions that the "bureaucratic administration of limited, categorical, quantified data" performed by current credit scoring systems added "new depth and specificity of data".<sup>18</sup> This "new depth and specificity of data" means that creditors could assess loaners' ability to pay back more accurately, disregarding the previous superficial process of assessment that only reports on a borrower's behaviors with raw unprocessed data. This alteration benefits creditors because they can make rational credit decisions with much calculable risks rather than incalculability caused by human error; however, this bureaucratic nature

hampers many citizens from developing economies from accessing credit because of their lack of paperwork and credit reports. This issue is a major issue in the case of Vietnam, where a recent shift from nine-digit IDs to twelve-digit IDs led to confusion amongst many citizens.<sup>19</sup> Additionally, a senior officer from CIC mentioned that “borrowers may try to conceal their information by using different papers to try and apply for different loans”.<sup>19</sup>

Besides being bureaucratic, the credit scoring system is also affected by bias in the scoring process. The credit scoring system typically samples a historic applicant pool to decide whether to grant credits.<sup>20</sup> However, because many applicants were denied credit, a system based on a population of accepted applicants creates significant bias when there is a corresponding population of denied applicants.<sup>20</sup> Bias estimates have been obtained after assessing these models; furthermore, it is impossible to estimate in which direction the bias lies. Developers of this system have tried to rectify this model by separating the sample of the denied into “good” and “bad” about the actual goods and bads.<sup>20</sup> The actual and denied “goods” are grouped, and the actual and denied “bads” are grouped to create an augmented model that generates less bias.<sup>20</sup> Although creditors have attempted to rectify this problem, it is important to address this issue as it is a major component that flaws many credit scoring systems since biased estimates are still obtained with these alternative models.

Another major issue many credit scoring models face is sample size. Credit scoring systems usually use insufficiently large samples to achieve reliability in assigning point values.<sup>20</sup> For example, for the occupation characteristic of a credit scoring system employed by a major oil company, the occupations of farm foreman and laborers, enlisted personnel, clergymen, entertainers, farmers and ranchers, and government and public officials received few points. However, the sample sizes on which the point scores are based were three, twenty-three, four, four, three, and three.<sup>20</sup> Similarly, when zip code is used as a characteristic in assessing credit scores, using 3,000 or fewer subjects results in the point scores for many zip codes based on very few data points, resulting in a disproportionate assessment.<sup>20</sup> A more diverse population of applicants could somewhat rectify this problem should the credit scoring system be more accessible. The presence of many potential borrowers might not only rectify this flaw in many credit scoring models but also facilitate financial inclusion.

### ***The “affordability” of credit’s impacts on financial inclusion***

As mentioned above, the greatest dilemma that lenders and creditors have to consider is whether financial inclusion’s “affordability” will negatively impact the “sustainability” that credit scoring systems have created for loaners. Financial inclusion is the “delivery of banking services at an affordable cost to the vast sections of disadvantaged and low-income groups.”<sup>21</sup> The financially excluded will be defined as households who are denied credit despite their demands. Financial inclusion plays an important role in developing economies because their large farmer populations rely on a constant stream of capital. Therefore, financial inclusion focuses on bringing these disadvantaged groups to opportunities for formal and sustainable credit from commercial banks, cooperative banks, RRBs, NABARD SHG-linkage and other self-help

groups, and credible microfinance institutions.<sup>21</sup> Access to formal credit lines will help increase the productivity and sustainability of farmers and other vulnerable groups.

The utilization of credit scoring systems, however, usually results in highly selective methods, as mentioned above, that might limit the use of credits amongst citizens of developing countries, especially farmers. Even though farmers prefer loans from banks due to larger amounts in loan, longer grace periods, and longer repayment periods, these loans are hard to obtain; therefore, farmers usually resort to other credit options that might be riskier than formal credit.<sup>22</sup> These informal credits take advantage of pre-existing social relations, mainly based on trust, and in many cases, the interest rates are colossally higher than formal banking rates.<sup>23</sup> Michel Letart defined this informal sector of the economy as “non-declared activities, performed with little capital and lots of unskilled labor, in a very small scale, with no respect for regulation”.<sup>24</sup> We will consider the case study of Kenya and Vietnam to assess the widespread use and risk of these informal credit lines.

Garment producers in Kenya possess a “pecking order” when considering credit sources, which prioritize bank loans, MFIs (micro-finance institutions), and finally, more informal sources such as ROSCAs (rotating savings and credit associations) or families and friends. ROSCAs were highlighted as the most important source of credit for farmers as they yield the highest average loan size (44,960 in Kenyan shillings).<sup>25</sup> This demonstrates the major impact that informal credit lines have on various borrowers. Informal credit lines are rife in Kenya and a major issue in Vietnam. Vietnam’s formal credit sector has been expanding significantly with the advent of many banks that aim to provide loans to farmers, such as Agribank or the Bank for Social Policies (VBSP). However, the problem of too much bureaucracy arises when giving out credit as many banks require collateral (red book, land title) and administrative documents (such as permanent residence certificates and poor household cards), which many borrowers do not possess.<sup>26</sup> This has prompted many people who need proper documentation in Vietnam to resort to informal credit. However, the underlying uncertainty of these loans is dangerous because a loan of a few hundred thousand dong (VND) can reach a few million in weeks or months, driving borrowers into a dangerous spiral and ultimately resulting in over-indebtedness. These absurd interest rates are illegal according to Article 163 of Vietnam’s criminal code: “Those who provide loans at an interest rate ten or more times higher than the maximum interest rate prescribed by law, which is exploitative, shall be subject to a fine of one to ten times the interest amount or non-custodial reform to up to one year.”<sup>26</sup> The uncertain nature of these informal credit lines creates unsustainable loans, which entirely contradict formal creditors’ aims to achieve sustainable credits. Informal credit certainly increases borrowers’ opportunity to access credit, but its unsustainable nature will be detrimental to borrowers and ultimately make achieving financial inclusion much more difficult. As a result, it should be acknowledged that there are different paths to financial inclusion, such as informal credits; however, formal credit lines are much more sustainable in creating conditions for financial inclusion.

### ***The trade-off between “sustainability” and “affordability” when targeting financial inclusion***

As mentioned above, formal credit access is a more sustainable approach for both lenders and borrowers because it employs a highly selective scoring process to ensure the efficiency of credit loans. Nevertheless, because of such a selection process, many disadvantaged consumers convert to informal credit lines, which are much riskier and more uncertain. Questions remain about whether traditional credit scoring methods should be altered to create a more favorable condition for the disadvantaged to access formal credit. A full-scale financial inclusion strategy would mean more people would receive access to credit, which would call for a less selective credit scoring method and riskier approaches from creditors. This may negatively affect the efficiency of many creditors when giving out credit. This section explores whether the “affordability” of credit will negatively impact the “sustainability” of loaners' portfolios. From a policy perspective, the question is to bank the “unbanked but bankable” (by reducing market imperfections) or bank the “unbanked even if unbankable” (by reducing market competition and handing out credit through less selective methods).<sup>27</sup>

“A Hundred Small Steps,” a program released under the chairmanship of Raghuram G Rajan (RRC) criticized banks in India for intentionally leaving out the poor and said that mandatory rural banking branches do not effectively service the poor.<sup>28</sup> The RRC pointed out that banks turn towards the bankable in the priority sector and tend to migrate towards them more than trying to meet the needs of the “unbanked but unbankable.”<sup>29</sup> However, it is important to note that the main reason banks do this is due to the risk that many potential consumers pose to banks' overall portfolio safety since banks' primary goals are to ensure an efficient system that can be sustained. Moreover, the interest rate ceiling reduces bankers' desires to service the excluded. The credit scoring system role once again comes into place to serve the interest of bankers as it requires many of these potential credit consumers to possess the paperwork to properly assess their risks, thus resulting in high negligence of poor consumers. Suppose banks were to disregard this process and bank the “unbankable” to create opportunities for access to credits; in that case, they might suffer financial uncertainty that might negatively impact their portfolio.

The definition of financial inclusion is once again re-emphasized to contradict ideas that support financial inclusion through mass credit distribution: “Individuals and firms are not denied access to basic financial services based on motivations other than efficiency criteria.”<sup>30</sup> In this case, rather than choosing to bank the “unbankable,” bankers and creditors prioritize using a credit scoring system to secure their investing portfolio. RRC also mentioned that the problem with credit access to poor people is also limited due to hidden charges from creditors or banks to poor people to cover the difference in the market rate.<sup>31</sup> This approach, however, signifies a disregard for traditional credit scoring methods which implies credit availability for the poor but is restricted by traditional credit scoring methods. This shows that the approach to financial inclusion that would facilitate creditors' portfolios would be to bank the “unbanked but bankable” rather than bank the “unbanked but unbankable” through improvements in traditional credit scoring methods.

The RRC has also proposed the formation of Small Finance Banks(SFBs) with strong linkage to mainstream banks which are local to understand the needs of local businesses.<sup>31</sup> Their report mentions that SFBs have to be small and that the center of decision-making must be close to the loan officer to facilitate the approval of requests without much documentation, delay, and loss of information.<sup>31</sup> Not only will the formation of such banks undoubtedly provide more credit availability for poor and disadvantaged people, but these will also eschew the current credit scoring system and opt for a more human-run system that considers uncertainty caused by human influence. As mentioned above, the primary aim of creditors is to safeguard and sustain their portfolio while also trying to make the process of giving out credit sustainable by eliminating social ties and human error.

With this proposed solution of an SFB, poor people will certainly benefit from the abundance of credit, though creditors will also suffer greater risks in their portfolios. For example, the case of The Bank for Social Policies (VSBP) in Vietnam was called “Bank for the Poor” because it aimed to lend credit to the poorest in Vietnam, who are typically denied credit.<sup>32</sup> Contrary to RRC’s assumption that these SFBs’ low-cost structure will allow loans to be profitable, the VSBP suffered unsatisfactory financial results. Its low productivity obliged the Vietnamese government to subsidize it, and international donors progressively withdrew their financial support despite the bank reaching millions of people.<sup>32</sup> The efficiency of banking the “unbankable” is poor in this situation and emphasizes the need for an orthodox credit scoring system to manage loans.

The approach of banking the “unbanked but bankable” will certainly not be as practical as banking the “unbanked but unbankable” from the perspective of poor and disadvantaged borrowers. However, from the perspective of creditors, it will ensure the safety of their portfolio and allow sustainability. Credit scoring can be greatly improved by increasing the number of potential borrowers. As mentioned above, when discussing the credit scoring system’s problems, the lack of data and bias impacts a complete assessment of clients' repayment ability when making credit scores. By sampling a historical pool of applicants, credit scores are determined. However, samples of previous customers who were granted credit will create an unbalanced comparison because many customers who were granted credit are rich and financially stable. The influx of new potential customers from less advantaged backgrounds will allow samples to be more diverse, thus improving the current credit scoring system. As mentioned, credit scoring systems also face considerable trouble from insufficiently large sample sizes. These sample sizes will contain an unbalanced number of occupations and demographic data, thus creating difficulties in the credit scoring system. The introduction of new customers from a variety of occupations and demographic backgrounds will certainly improve the efficiency of the credit scoring system.

## **Conclusion**

Financial inclusion is a major goal many developing economies aim to achieve because formal credit lines are safer, more credible, and offer more exciting rates than informal credit

lines for many disadvantaged borrowers. However, financial inclusion can create issues regarding existing credit tools, especially the credit scoring system that provides creditors safeguards to their portfolios. In the article, the emphasis on “sustainability” was constantly stressed to highlight how important it was for creditors and borrowers to sustain a credit relationship that would benefit both sides. However, the credit scoring system’s bureaucratic nature is a major issue that can hamper financial inclusion due to its selective nature to ensure creditors’ portfolios’ safety. The case study of Vietnam certainly showed how complicated the credit scoring process can be with the requirements of multiple legal documents. The current credit scoring system’s flaws also contribute to preventing financial inclusion. Surprisingly, these problems stemmed from the lack of data that could be provided if more poor citizens could access credit. The improvement of the credit scoring system could greatly facilitate the achievement of financial inclusion. However, financial inclusion, in this case, will not be as widespread. At the same time, disregarding the credit scoring system can create a much more effective financial inclusion process (in the sense that it provides more easily accessible credit). Still, it will undoubtedly render the relationship between creditors and borrowers unsustainable.

### **Acknowledgments**

I would like to thank my mentor Alan Zhang for having guided me through this journey and my paper would not have been completed without his dedication.

## Works Cited

- Nguyen, T. PM approves Vietnam national strategy for financial inclusion until 2025. VietNamNet News. <https://vietnamnet.vn/en/pm-approves-vietnam-national-strategy-for-financial-inclusion-until-2025-611649.html> (accessed 2023-11-27).
- Lainez, N.; Thi, B.; Doai, T.; Khanh, T.; Linh, L.; Phuong, T.; Pannier, E. Financial Inclusion and Consumer Finance in Vietnam: Challenges and Approaches to Credit Scoring; 2021. [https://www.iseas.edu.sg/wp-content/uploads/2021/08/ISEAS\\_Perspective\\_2021\\_112.pdf](https://www.iseas.edu.sg/wp-content/uploads/2021/08/ISEAS_Perspective_2021_112.pdf)
- Pesqué-Cela, V.; Tian, L.; Luo, D.; Tobin, D.; Kling, G. Defining and Measuring Financial Inclusion: A Systematic Review and Confirmatory Factor Analysis. *Journal of International Development* **2021**, 33 (2), 316–341. <https://doi.org/10.1002/jid.3524>.
- Carruthers, B. G. From Uncertainty toward Risk: The Case of Credit Ratings. *Socio-Economic Review* **2013**, 11 (3), 525–551. <https://doi.org/10.1093/ser/mws027>.
- Guseva, A. Into the Red; Stanford University Press, 2008.
- Carruthers, B. G. From Uncertainty toward Risk: The Case of Credit Ratings. *Socio-Economic Review* **2013**, 11 (3), 525–551. <https://doi.org/10.1093/ser/mws027>.
- Knight, F. Risk, Uncertainty and Profit.; Vernon Press: Wilmington, 1921.
- Carruthers, B. G. From Uncertainty toward Risk: The Case of Credit Ratings. *Socio-Economic Review* **2013**, 11 (3), 525–551. <https://doi.org/10.1093/ser/mws027>.
- March, J.G. and Simon, H.A. (1958) Organizations. Wiley, New York. - References - Scientific Research Publishing. Scirp.org. [https://www.scirp.org/\(S\(i43dyn45teexjx455qlt3d2q\)\)/reference/ReferencesPapers.aspx?ReferenceID=1099306](https://www.scirp.org/(S(i43dyn45teexjx455qlt3d2q))/reference/ReferencesPapers.aspx?ReferenceID=1099306).
- Carruthers, B. G. From Uncertainty toward Risk: The Case of Credit Ratings. *Socio-Economic Review* **2013**, 11 (3), 525–551. <https://doi.org/10.1093/ser/mws027>.
- Capon, N. Credit Scoring Systems: A Critical Analysis. *Journal of Marketing* **1982**, 46 (2), 82. <https://doi.org/10.2307/3203343>.
- Guseva, A.; Rona-Tas, A. Uncertainty, Risk, and Trust: Russian and American Credit Card Markets Compared. *American Sociological Review* **2001**, 66 (5), 623. <https://doi.org/10.2307/3088951>.
- United States. Congress. Equal Credit Opportunity Act Amendments of 1976. Stlouisfed.org. <https://fraser.stlouisfed.org/title/equal-credit-opportunity-act-amendments-1976-1039> (accessed 2023-11-18).
- Capon, N. Credit Scoring Systems: A Critical Analysis. *Journal of Marketing* **1982**, 46 (2), 82. <https://doi.org/10.2307/3203343>.
- McClanahan, A. Bad Credit. *Representations* **2014**, 126 (1), 31–57. <https://doi.org/10.1525/rep.2014.126.1.31>.
- Poon, M. From New Deal Institutions to Capital Markets: Commercial Consumer Risk Scores and the Making of Subprime Mortgage Finance. *Accounting, Organizations and Society* **2009**, 34 (5), 654–674. <https://doi.org/10.1016/j.aos.2009.02.003>.



- Lainez, N. Informal Credit in Vietnam: A Necessity rather than an Evil. *Journal of Southeast Asian Economies* **2014**, 31 (1), 147–154.
- Marron, D. “Lending by Numbers”: Credit Scoring and the Constitution of Risk within American Consumer Credit. *Economy and Society* **2007**, 36 (1), 103–133. <https://doi.org/10.1080/03085140601089846>.
- Lainez, N.; Thi, B.; Doai, T.; Khanh, T.; Linh, L.; Phuong, T.; Pannier, E. *Financial Inclusion and Consumer Finance in Vietnam: Challenges and Approaches to Credit Scoring*; 2021. [https://www.iseas.edu.sg/wp-content/uploads/2021/08/ISEAS\\_Perspective\\_2021\\_112.pdf](https://www.iseas.edu.sg/wp-content/uploads/2021/08/ISEAS_Perspective_2021_112.pdf)
- Capon, N. Credit Scoring Systems: A Critical Analysis. *Journal of Marketing* **1982**, 46 (2), 82. <https://doi.org/10.2307/3203343>.
- Dev, S. M. Financial Inclusion: Issues and Challenges. *Economic and Political Weekly* **2006**, 41 (41), 4310–4313.
- Akoten, J.; Sawada, Y.; Otsuka, K. The Determinants of Credit Access and Its Impacts on Micro and Small Enterprises: The Case of Garment Producers in Kenya. *Economic Development and Cultural Change* **2006**, 54 (4), 927–944. <https://doi.org/10.1086/503585>.
- Lainez, N. Informal Credit in Vietnam: A Necessity rather than an Evil. *Journal of Southeast Asian Economies* **2014**, 31 (1), 147–154.
- Philippe, H. Michel Lelart, de La Finance Informelle à La Microfinance. *Revue Tiers Monde* **2006**, 47 (187), 673–673.
- Akoten, J.; Sawada, Y.; Otsuka, K. The Determinants of Credit Access and Its Impacts on Micro and Small Enterprises: The Case of Garment Producers in Kenya. *Economic Development and Cultural Change* **2006**, 54 (4), 927–944. <https://doi.org/10.1086/503585>.
- Lainez, N. Informal Credit in Vietnam: A Necessity rather than an Evil. *Journal of Southeast Asian Economies* **2014**, 31 (1), 147–154.
- Pesqué-Cela, V.; Tian, L.; Luo, D.; Tobin, D.; Kling, G. Defining and Measuring Financial Inclusion: A Systematic Review and Confirmatory Factor Analysis. *Journal of International Development* **2021**, 33 (2), 316–341. <https://doi.org/10.1002/jid.3524>.
- Jain, A. A Hundred Small Steps Ii a HUNDRED SMALL STEPS a Hundred Small Steps Report of the Committee on Financial Sector Reforms Government of India Planning Commission. [www.academia.edu](http://www.academia.edu).
- Dasgupta, R. Two Approaches to Financial Inclusion. *Economic and Political Weekly* **2009**, 44 (26), 41–44.
- Amidžić, G.; Massara, A.; Mialou, A. Assessing Countries’ Financial Inclusion Standing: A New Composite Index. *SSRN Electronic Journal* **2014**. <https://doi.org/10.2139/ssrn.2407529>.
- Dasgupta, R. Two Approaches to Financial Inclusion. *Economic and Political Weekly* **2009**, 44 (26), 41–44.
- Lainez, N. Informal Credit in Vietnam: A Necessity rather than an Evil. *Journal of Southeast Asian Economies* **2014**, 31 (1), 147–154.

## Breast Cancer Survey By Arnav Rajadhyax

Breast cancer is the second most common type of cancer worldwide and the most prevalent cancer among women. Breast cancer has a particularly severe impact on developing countries like India. In 2022, India has the third highest rate of breast cancer amongst women and the highest breast cancer mortality rate globally (Breast cancer statistics: World cancer research fund international 2024).

A notable concern is the high incidence of breast cancer in young Indian women, which deviates from the typical trend seen globally where about 80% of breast cancer cases occur in women aged 45 or older (Risk factors for breast cancer). A recent study by Apollo Hospital, Mumbai found that 25% of breast cancer cases in India were in women aged 40 or younger (Breast cancer strikes younger Indian women at an alarming rate: Study 2023). Several factors contribute to this trend. Lifestyle changes, particularly the rise in obesity, are commonly cited by doctors as contributing to the increase in breast cancer among younger women. Additionally, a significant factor is the lack of awareness among uneducated women. Stigmas associated with mammograms deter many women from undergoing regular screenings, resulting in late-stage diagnoses and poorer outcomes.

The idea for this survey came to me upon reading another study which talked about the lack of awareness about breast cancer in rural areas and the need for educational resources. To emphasise the importance of education in recognising the dangers of breast cancer and the necessity for regular mammograms, guided by Dr Mona Mehta, I designed a survey aimed at women living in urban areas. I consulted with Dr Mehta on appropriate questions to ask and which questions are too personal. My objective was to find a link between education, access to resources, and awareness of breast cancer.

### Survey

The survey was conducted at Lilavati Hospital in Mumbai (from 1st to 8th of June, 2024). The Lilavati Hospital is a private hospital and a research centre.

35 patients who came for a regular mammogram check in were requested to take the survey and their responses are below. The results provide valuable insights into the educational and socio-economic background of the respondents and their knowledge about breast cancer.

### Key Survey Findings

1. **Age Distribution:** A majority of the respondents (68.58%) were below 50 years of age. This indicates a significant number of younger women are seeking mammograms.

AGE	>50	<50
	31.4%	68.6%

2. **Marital Status:** An overwhelming 94.28% of respondents were married, which might reflect societal norms influencing health-seeking behaviour.

MARITAL STATUS	MARRIED	SINGLE	DIVORCED
	94.3%	2.9%	2.9%

3. **Education Level:** Most respondents were well-educated, with 48.57% being college graduates and 28.57% holding postgraduate degrees. Only 2.86% had not graduated from school.

LEVEL OF EDUCATION	Not graduated from School	High School	College Graduate	Post Graduate
	2.9%	20%	48.6%	28.6%

4. **Occupation:** The majority (60%) were housewives, with 37.14% being professionals. >91% of respondents had at least one child.

OCCUPATION	Housewife	Professional	Business	Other
	60%	37.1%	0%	2.9%

5. **Pregnancy:** 8.6% of respondents did not have Children. Other parameters related to pregnancy and child rearing are given below.

Number of Children	0	1	2	3	4
	8.6%	17.1%	60%	11.4%	2.9%

Age at 1 <sup>st</sup> Pregnancy	N/A	<30 yrs	>30 yrs
	8.6%	77.1%	14.3%

HISTORY of BREAST feeding	YES	NO
	85.7%	14.3%

6. **Other key parameters:**

Onset of Menstruation	Before 12	After 12
	22.9%	77.1%

Menopausal	N/A	45-50 yrs	50-55 yrs	50-55 yrs	>55yrs
	31.4%	40%	28.6%	28.6%	0%

7. **Family History Breast Cancer:** 85.7% did not have a family history of breast cancer, which is significant given the high awareness and self-initiative in getting mammograms.

<b>FAMILY HISTORY – BREAST CANCER</b>	<b>YES</b>	<b>NO</b>
	<b>14.3%</b>	<b>85.7%</b>

### 8. Mammogram related details:

First Mammogram: Only 22.9% were undergoing their first mammogram, indicating that most women were returning for follow-ups, which suggests good awareness and adherence to screening recommendations. Most of the respondents were self referred – This also indicates a level of self awareness, response to public health preventive campaigns.

<b>1<sup>st</sup> Mammogram</b>	<b>YES</b>	<b>NO</b>
	<b>22.9%</b>	<b>77.1%</b>
<b>Reasons for coming to Mammogram</b>	<b>Recommended by Doctor</b>	<b>Came on own</b>
	<b>14.3%</b>	<b>85.7%</b>

Understanding the symptoms: There was a high level of understanding of symptoms of breast cancer as evident from the table below:

<b>Lump in Breast</b>	<b>Lump in the Armpit</b>	<b>Nipple Discharge</b>	<b>Pain in Breast</b>	<b>None</b>
<b>91.4%</b>	<b>74.3%</b>	<b>82.9%</b>	<b>65.7%</b>	<b>2.9%</b>

### 9. Response of patients related to usefulness of Mammograms, symptoms:

<b>Should every woman above a certain age undergo Mammogram</b>	<b>Yes</b>	<b>No</b>
	<b>100%</b>	<b>0%</b>

<b>Does Mammogram help to pick up early cancer</b>	<b>Yes</b>	<b>No</b>
	<b>94.3%</b>	<b>5.7%</b>

<b>Will you recommend other women to undergo Mammograms?</b>	<b>Yes</b>	<b>No</b>	<b>Maybe</b>
	<b>85.7%</b>	<b>5.7%</b>	<b>8.6%</b>

### Learnings from the Survey

Breast Cancer Awareness in Urban Areas: The responses revealed a general trend of education and breast cancer awareness among the women who participated. Comparing this with studies on rural populations reveals a stark contrast. In rural areas, awareness and knowledge about breast cancer are significantly lower. For instance, the study that inspired my survey revealed that in rural Central India a third of the respondents had not heard about breast cancer, and more than 90% were unaware of breast self-examination (Gangane 3). The study also highlighted significant delays in seeking medical consultation due to lack of awareness and

socio-economic barriers. The survey results highlight the importance of education in raising awareness about breast cancer as well as the several key differences between urban and rural populace as regards Breast Cancer.

1. **Education and Awareness:** Urban, educated women show a better understanding of breast cancer symptoms and the importance of early detection. In contrast, rural women, often with lower levels of education, have limited knowledge about breast cancer, which leads to delays in diagnosis and treatment.
2. **Health-Seeking Behaviour:** Urban women are more likely to seek medical advice at the first sign of symptoms, whereas rural women often delay due to a lack of awareness and stigmas.
3. **Screening Practices:** Regular mammograms are more common among urban women due to better access to healthcare facilities and greater awareness. Rural areas often lack organised screening programs as well.

## **Conclusion**

The survey underscores the critical role of education in enhancing breast cancer awareness and the necessity of targeted awareness campaigns in rural areas. Efforts should focus on:

1. **Educational Programs:** Implementing breast cancer awareness programs tailored to rural populations can significantly improve knowledge and early detection practices.
2. **Improving Access:** Enhancing healthcare infrastructure in rural areas to provide easier access to screening and diagnostic facilities.
3. **Community Engagement:** Utilising community health workers to disseminate information and encourage regular screenings can bridge the gap between urban and rural healthcare practices.

By addressing these disparities, we can improve early detection rates and reduce the mortality associated with breast cancer in rural India.

### Works Cited

- Breast cancer statistics: World cancer research fund international (2024) WCRF International. Available at: <https://www.wcrf.org/cancer-trends/breast-cancer-statistics/> (Accessed: 28 July 2024).
- Gangane, N. (no date) Breast cancer in rural India: Knowledge, attitudes, practices. Available at: <https://www.diva-portal.org/smash/get/diva2:1187627/FULLTEXT01.pdf> (Accessed: 28 July 2024).
- Risk factors for breast cancer (no date) WebMD. Available at: <https://www.webmd.com/breast-cancer/overview-risks-breast-cancer> (Accessed: 28 July 2024).
- Standard, B. (2023) Breast cancer strikes younger Indian women at an alarming rate: Study, Business Standard. Available at: [https://www.business-standard.com/india-news/breast-cancer-strikes-younger-indian-women-at-a-alarming-rate-study-123102300753\\_1.html](https://www.business-standard.com/india-news/breast-cancer-strikes-younger-indian-women-at-a-alarming-rate-study-123102300753_1.html) (Accessed: 28 July 2024).

## **A Cultural Awakening for South Korean Guide Dogs By Catrina Cho**

### **Abstract**

Although man's domestication of dogs began about 150,000 years ago, the earliest account of the unique relationship between a blind person and a dog was discovered in a 1st-century mural in the Roman Herculaneum ruins. More illustrations of dogs leading the blind have been found, sprinkled throughout history, including on a 13th-century Chinese scroll and in other instances in Europe. With the increase of blinded soldiers returning from WWI, German Doctor Gerhard Stalling was inspired to systematically train dogs to help the blinded, establishing the world's first guide dog school for the blind in 1916. Although guide dog schools have since been established and are growing in number around the world, restoring agency and dignity to blind individuals, East Asia, however, is lagging behind in its implementation. This paper will discuss the cultural and social stigmas against guide dogs in South Korea, and discuss possible solutions to increase public acceptance and support for guide dogs.

### **A Snapshot of South Korea's Current Situation**

According to South Korea's Samsung Fire & Marine Insurance Guide Dog School (SGDS), guide dogs are "service dogs specially trained to help visually impaired people move around obstacles." In other words, they are overall "responsible for the safety of people with visual impairment." However, according to a news article that interviewed Park Tae-jin, the current general director of Samsung Guide Dog, guide dogs are not only present for practical support for their handlers - individuals who receive the service of guide dogs. They bring emotional stability and comfort to their handlers, as the handlers gain "confidence by empowering them to be active caregivers." Regardless of this loving companionship, guide dogs and handlers are not always and fully welcomed in every nation's society as much as they should be. Unfortunately, this problem is prevalent in some Asian countries. Diving deeper into South Korea, cases where guide dogs were not treated with hospitality can be easily found. For instance, South Korea's first visually impaired lawmaker Jeong Hwa-won had to enter the plenary chamber without his guide dog during his term from 2004 to 2008. His guide dog was banned, "based on a clause in the law of the National Assembly which states an assembly person should not bring objects or food that could cause a disturbance into the plenary chamber and committee meeting rooms." Years later, as Kim Ye-ji, a visually impaired politician, joined the Assembly, the National Assembly Secretariat has been working on improving the guide dog entry regulations. With support from multiple politicians from different parties, the rule regarding the entry of guide dogs to the National Assembly was modified in 2020, now allowing guide dogs to enter the chamber without any hindrance.

However, South Korea's cultural perception of guide dogs has not yet been entirely changed. In 2020, there was a publicized case in which a trainee guide dog and its trainer were refused entry to Lotte Mart. It was revealed that a Lotte Mart employee yelled at the dog and its trainer to leave and blamed the trainer for bringing a guide dog as she was not blind. This issue

went viral as witnesses uploaded photos of the trainee guide dog sitting with a frustrated face. With the prevalent use of social media, this recent case has significantly increased South Korea's awareness of the presence and role of guide dogs. In order to address this ongoing problem, the first step is to raise the public's awareness of guide dogs - especially through active interactions with guide dogs or trainee guide dogs in public.

For a comparative analysis into the lack of public support for guide dogs in Korea, we can look to the West for marked differences in cultural norms. Unlike other Western countries, South Koreans have a clear preference for dog sizes. In regard to South Koreans' dog preferences, the 2024 top-ranked dog breeds based on popularity revealed all small-sized breeds, such as Pomeranians, Bichon Frises, or Maltipoos. All popular breeds happen to be small-sized in nature. In South Korea, large dogs are not really welcomed in general, in comparison to small dogs. Some people get overly surprised and agitated when they bump into big dogs in the apartment elevator or even on the street. Moreover, there are even South Korea's public spaces that only permit entry to dogs based on having a smaller size. Usually, dogs that weigh around 10-15 kg would be the biggest size that are allowed to enter those public spaces, making it difficult for owners of larger dogs to spend time with their dogs outside of their houses. The discrimination against larger dogs, which happen to be the breeds for most guide dogs, constitutes part of the social stigma that guide dogs face in Korea. As a volunteer temporarily raising a trainee guide dog, I myself have experienced these social discrepancies daily. My dog trainee, a friendly labrador retriever with a yellow vest that indicates that she is a trainee guide dog, would render some people uneasy in public spaces. Not expecting to encounter this larger dog in public, people who were not conscious about their surroundings would be quite surprised and offended when my dog carefully passed by. These were all observations made in South Korean urban settings.

### **Global Context Given South Korea's Short History and Recent Developments**

To consider South Korea's position in regards to global history and culture, we take a step back to the very first systematic attempt in training dogs to assist visually impaired people, at the 'Les Quinze-Vingts' hospital for the blind in Paris of 1780. As mentioned before, Dr. Gerhard Stalling trained dogs to aid blinded WWI soldiers. One day, when the doctor returned from leaving the dog and a blind patient alone due to an urgent issue, he witnessed the dog looking after the patient. Dr Stalling then studied and developed different methods, creating guides to train dogs. In 1916, he opened the very first guide dog school for the visually impaired. His foundation of the guide dog school influenced and led countless places to create such schools, in Bonn, Breslau, Dresden, and Essen. Progress has continued until the present, with guide dog schools created all over the world. To put into context, currently, South Korea has only one guide dog school, which was established in 1993 by Samsung. By 2017, the Samsung Guide Dog School had graduated approximately 200 guide dogs. Despite the relatively late start of the guide dog system, SGDS has been making progressive and noticeable growth in its short history, which significantly boosted the welfare of visually impaired people of South Korea. However, certain



aspects are hindering the further growth of South Korea's guide dog system or SGDS, such as animal protection rights.

One of the cultural stigmas that foreigners may have about South Korea is the consumption of dog meat. In the past, some Koreans consumed dog meat when food and resources were scarce due to famine and war. Currently, dog meat production and consumption are no longer prevalent in South Korea. According to Sangkyung Lee, a campaign manager at Humane Society International, "dog meat is only eaten by a very small proportion of the population." Meanwhile, "millions of dogs had to endure unimaginable pain and suffering" despite the small proportion of consumption. In 2024, a new law officially banned breeding, butchering, distributing, or selling dogs. This law will take effect after a three year grace period, which ends in 2027. Sangkyung Lee stated, "This ban marks a significant turning point in South Korea's attitude to animal protection." As noted, although this law is quite recent, it reveals how animal protection rights have overall been improving, by gaining public and political support in South Korea. This change in the public's opinion towards dogs overall will aid in the improvement of raising awareness for guide dogs in Korea.

### **Raising Awareness to Combat Stigma**

To get an expert's opinion, I interviewed one of the SGDS trainers, Nayoung Mok. When asked what she believed to be the biggest concern regarding the guide dog system in Korea, she stated that it was the guide dog's public entrance. She revealed that after the guide dog public entrance law was created within the disabled person welfare law, this problem has been improved noticeably. However, she argues that a considerable number of South Koreans are still unaware of the role that guide dogs have and unaware that they are allowed in all public spaces. She informs that many public spaces are refusing guide dogs (particularly the dogs in training) from entering, by claiming that the puppy walkers - the volunteers who raise and train the trainee guide dogs for about a year - are not disabled. Additionally, there are many misconceptions among South Koreans. Many people believe that guide dogs are miserable, and that they are exploited to help disabled individuals. However, trainer Mok refutes this claim. She states that guide dogs are essentially companion dogs. The guide dogs only work as guide dogs with their harnesses on and do duties as guide dogs for about two hours a day. The remainder of their days are very similar with other pets. The two to three hours a day spent as a guide dog are not any different from non-guide dogs going on a walk with their owners. In fact, many behaviors and skills that guide dogs learn are what normal pets should also have in a community living with humans.

While writing this research paper, the biggest question that developed was on the relationship and impact between guide dogs and the visually impaired. Trainer Mok stated that guide dogs for the visually impaired are not to be viewed as merely a form of GPS or road navigation. Instead, guide dogs are perceived as life companions for the visually impaired, providing the opportunity for them to better and easily socialize with other people in society. Overall, guide dogs are a bridge that connects society with those who often face difficulty being

well connected with the community. Having learned firsthand that South Korea has a lack of awareness towards guide dogs, in comparison with countries such as the United States and Europe, there is a call to action to change the negative stigma towards guide dogs and the system itself.

In regard to solutions or even small efforts that could change the current situation, Mok stated that although legal actions could be significant, she personally believes that the most important step is increasing exposure so the South Korean society has a better image about guide dogs. One way to achieve this was for guide dogs in training and puppy walkers to be especially cautious of not creating disturbance to others while in public. Mok also added that in the past, South Korea had a negative perception towards dogs, making the situation challenging. However, the awareness of guide dogs has improved significantly throughout South Korea's short guide dog history of 30 years. Mok hopes that based on the efforts of current staff, trainers, and puppy walkers, South Korea's social and cultural perception of guide dogs would become more positive in the near future.

## **Conclusion**

As stated above, South Korea has made and is making great strides towards its treatment of guide dogs. However, I believe that raising people's awareness of guide dogs and their entire system would be the most efficient and effective method to combat social stigma. The public's indifference seems to be the biggest factor that delays South Korea's guide dog system. For example, the most common way would be through frequent exposure, in that the "puppy walkers" continue going out in public. Not only does going to public spaces frequently with a guide dog in training help the puppies socialize, it also has a tremendous effect on bystanders who encounter these puppies. Currently in South Korea, there are about 80 guide dogs in training with their puppy walkers. As 80 is a small number, it is quite difficult for the average bystander to bump into these puppies. However, efforts of puppy walkers traveling to different public spaces would allow more people to interact with the dogs. As most Koreans have never met guide dogs or guide dogs in training, many do become genuinely interested when encountering these puppies with the special vests on. People who see the dogs' orange or yellow vest become intrigued and will oftentimes ask the puppy walkers about the vest's significance or search online for more information. These interactions do generate more interest from the public, ultimately leading to increased awareness and acceptance. Initially, this simple theory did seem implausible. However, as I have had my guide dog in training for about a year and went out as often as possible, I discovered that my theory rang true. Many bus drivers on buses we took now know what my dog does. Increasingly, I've seen awareness grow through interactions with the grocery store cashiers, the elderly neighbors and young children in our building, the security officers, and even cafe staff. In just one year of introducing my dog around the community, my neighborhood has increased awareness and hospitality towards guide dogs. Although not instantaneous, this paper proposes that through steady commitment, puppy walkers' efforts in socializing the guide

dogs in training in public will create an open environment in which South Korea can soon encounter both an improved guide dog system and improved welfare for the visually impaired.

## Works Cited

- “Guide Dogs in Korea: Public Perception Improves but Challenges Remain.” Arirang News, YouTube, 9 Jan. 2017, [www.youtube.com/watch?v=4s8LJSGRV\\_o](https://www.youtube.com/watch?v=4s8LJSGRV_o).
- “History of Guide Dogs.” International Guide Dog Federation, 24 Sept. 2021, [www.igdf.org.uk/guide-dogs/history-of-guide-dogs/](http://www.igdf.org.uk/guide-dogs/history-of-guide-dogs/).
- Jang, Lina, and Korea Bizwire. “Blind Lawmaker Calls for More Attention to Guide Dogs.” The Korea Bizwire, 29 Apr. 2022, [koreabizwire.com/blind-lawmaker-calls-for-more-attention-to-guide-dogs/217465](http://koreabizwire.com/blind-lawmaker-calls-for-more-attention-to-guide-dogs/217465).
- Kim, Rahn. “Guide Dog to Be Allowed into National Assembly.” Koreatimes, 20 Apr. 2020, [www.koreatimes.co.kr/www/nation/2024/05/113\\_288128.html](http://www.koreatimes.co.kr/www/nation/2024/05/113_288128.html).
- Koh, Ewe. “South Korea Finally Passes Historic Ban on Dog Meat Industry.” Time, Time, 9 Jan. 2024, [time.com/6553409/south-korea-passes-dog-meat-trade-ban/](https://time.com/6553409/south-korea-passes-dog-meat-trade-ban/).
- Lee, Ji-Young, and Sarah Chea. “Lotte Mart Slammed for Illegally Kicking out Trainee Guide Dog.” Korea JoongAng Daily, 30 Nov. 2020, [koreajoongangdaily.joins.com/2020/11/30/business/industry/Lotte-Mart-guide-dog-puppy-walking/20201130190900576.html](http://koreajoongangdaily.joins.com/2020/11/30/business/industry/Lotte-Mart-guide-dog-puppy-walking/20201130190900576.html).
- Nayoung Mok (SGDS Trainer), personal communication, July 16, 2024.
- No, Kyung-min. “[Eye] ‘Dogs Cherish Time with Human Companions, Visually Impaired or Not.’” The Korea Herald, The Korea Herald, 20 Apr. 2024, [www.koreaherald.com/view.php?ud=20240418050794](http://www.koreaherald.com/view.php?ud=20240418050794).
- Rashid, Raphael, and Justin McCurry. “How South Korea Fell in Love with Dogs and Banned Their Slaughter for Meat.” The Guardian, Guardian News and Media, 19 Jan. 2024, [www.theguardian.com/world/2024/jan/19/south-korea-dog-meat-ban](https://www.theguardian.com/world/2024/jan/19/south-korea-dog-meat-ban).
- Song, Seung-hyun. “Behind the Making and Marketing of ‘trendy’ Dog Breeds.” The Korea Herald, The Korea Herald, 15 May 2024, [www.koreaherald.com/view.php?ud=20240515050067#:~:text=As%20of%202024%2C%20the%20most,coton%20de%20Tulears%2C%20and%20Maltipoos.&text=The%20popularity%20of%20certain%20dog,often%20triggered%20by%20media%20exposure](http://www.koreaherald.com/view.php?ud=20240515050067#:~:text=As%20of%202024%2C%20the%20most,coton%20de%20Tulears%2C%20and%20Maltipoos.&text=The%20popularity%20of%20certain%20dog,often%20triggered%20by%20media%20exposure).
- “What Is Guide Dog?” Samsung Fire&Marine Insurance Guide Dog School, [www.guidedog.co.kr/eng/social/guideDogSchool.do](http://www.guidedog.co.kr/eng/social/guideDogSchool.do). Accessed 16 Sept. 2024.

# RNA Immunoprecipitation for RNA Fragments Containing the m7G Cap

## By Sophie Kwon and Serene Kwon

### Abstract

RNA transcripts produced by eukaryotic RNA polymerase II have the characteristic m7G nucleotide at the 5' end. Specialized methods are devised to identify RNA sequences following the m7G cap, which are used for the determination of transcription start sites by RNA polymerase II. Currently, the best method to identify unknown transcription start sites is streptavidin-coated magnetic beads to select a biotinylated m7G derivative by a series of chemical reactions. In this study, we used antibodies instead of streptavidin to select RNA fragments containing the m7G cap to streamline the procedure by removing the reactions to biotinylate the m7G cap. PCR amplification reactions demonstrated that RNA immunoprecipitation performed well with commercially available antibodies against the m7G nucleotide.

### Introduction

Three distinct RNA polymerases are engaged in the specific transcription of different classes of RNA in the nuclei of eukaryotic cells [1]. RNA polymerase I is specialized in the transcription of the large precursor of the most abundant ribosomal RNAs. Polymerase III makes precursors of various short RNAs including tRNAs and 5S rRNA. RNAs transcribed by these RNA polymerases compose most of the RNAs in quantity. In contrast, RNA polymerase II produces a wide variety of long RNA precursors that are further processed into small nuclear RNAs, small nucleolar RNAs, microRNAs, long noncoding RNAs, and mRNAs encoding proteins. The RNAs transcribed by each RNA polymerase species have characteristics. The long ribosomal RNA precursor is the only RNA made by RNA polymerase I. The precursor RNAs transcribed by polymerase III are relatively short and maintain the triphosphate at the 5' end. The triphosphate group is usually not present in mature RNAs by nucleolytic cleavage. The phosphate at the 5' end is monomethylated at the 5' end of U6 RNA and 7SK RNA. In contrast, 7-methyl guanosine (m7G) is added to the first nucleotide of RNA polymerase II transcripts during the early stage of transcription (Fig 1) [2].

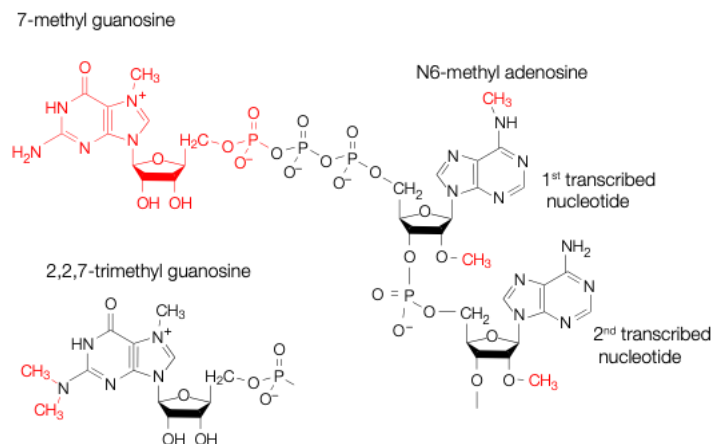


Fig 1: the structures of the m7G cap and the trimethylguanosine cap. Structures denoted the red color represents chemical structures added after transcription initiation.

Initially, all RNAs made by polymerase II have a 5' triphosphate, to which m7G is added by the 5'-5' linkage (m7G cap). The modified RNA structure is further modified by adding methyl groups to the first transcribed and the second transcribed nucleotides.

While the 5' ends of precursor RNAs transcribed by RNA polymerases I and III are not present in mature RNAs, the m7G cap is maintained in mRNAs and long noncoding RNAs. Therefore, the m7G cap is used to determine the transcription start sites of RNAs transcribed by RNA polymerase II. Precise determination of the transcription start site provides insights into the mechanisms and regulation of transcription initiation by RNA polymerases. Current methods to determine transcription start sites generally utilize one of the following approaches [3]. First, RNAs without the m7G cap are removed by using enzymes to degrade unprotected RNAs at the 5' end and the 5' end of the remaining RNAs are determined. Notably, the reaction to remove the unprotected 5' end should be complete to avoid false signals. Alternatively, RNAs containing the m7G cap are selected from total RNAs. To select RNAs containing the m7G cap by using streptavidin-coated beads, the m7G structure is biotinylated by tedious chemical reactions. A systematic comparison of these methods showed that the selection of biotinylated m7G was the best method.

Arguably, the simplest method to select the m7G cap is RNA immunoprecipitation using a specific antibody for m7G that was first reported in 1977 [4]. However, transcription start sites have not been determined by using a specific antibody for m7G. In this study, we investigated whether RNA fragments containing m7G are enriched by antibodies for m7G. The immunoprecipitated RNA fragments were converted into complementary DNAs (cDNAs) using reverse transcriptase. The DNAs were circularized and amplified by polymerase chain reaction. The amplified DNAs were quantified and visualized by gel electrophoresis. Based on the experimental data, we reasoned that RNA fragments containing m7G were efficiently enriched and RNA immunoprecipitation could be applied for transcription site determination.

## Material and Methods

Total RNA from HeLa cells was a gift from Dayoung Yoon. Dynabead mRNA purification kit (Invitrogen, #61006) was used according to the manufacturer's protocol. For RNA fragmentation, 2 mg of total RNA was incubated in a 20 ml fragmentation buffer (10 mM Tris-HCl pH 7.4, 10 mM ZnCl<sub>2</sub>) at 94°C for 5 min. The fragmented RNA was then purified using the Monarch RNA cleanup kit (50mg) (NEB) and 0.7 mg of RNA was incubated with 3 mg antibodies against m7G (MBL, RN016M, RN017M) or non-specific IgG bound to protein A/G Dynabeads 4°C for 2 hours. Beads were washed three times with a wash buffer (10 mM Tris-HCl, 150 mM NaCl, and 1% Empigen BB). RNA on magnetic beads was treated with 1 unit FastAP (Thermo Fisher Scientific) at 37°C for 20 min. The RNA fragments bound to antibodies were incubated with 20 pmole preactivated 3' adapter (5' AppNNNNNCGTCGTGTAGGGAAAGAGTGTAGATC biotin 3') overnight at 16°C in 15 ml RNA ligase buffer containing 50 mM Tris-HCl pH7.5, 10 mM MgCl<sub>2</sub>, 1 mM DTT, 15% PEG8000, 0.5 ml RNasin (Promega), 1.5 unit T4 RNA ligase I (NEB) and 300 unit truncated T4 RNA ligase II (NEB). RNA bound to beads was recovered by phenol/chloroform extraction and ethanol precipitation. The RNA-adapter hybrid was converted into complementary DNAs (cDNAs) by using a 2 pmole primer (5' phosphate CCTGTGGTTCGTAGCATCAGCTAGATCTACTCTTTC CCTACACGA 3') in 20 ml buffer containing 10 mM Tris-HCl pH8.3, 50 mM KCl, 5 mM MgCl<sub>2</sub>, 0.1 M DTT, 0.5 mM dNTPs, 1 ml RNasin and 10 unit SuperScript III reverse transcriptase (Invitrogen) according to manufacturer's recommendation. The cDNA was size-fractionated by electrophoresis on a 12% polyacrylamide gel containing 8 M urea and circularized by 100 unit circLigase ssDNA ligase (Biosearch technologies) at 60°C for 60 min. Circularized cDNAs were PCR-amplified using 1 unit AmpliTaq Gold DNA polymerases (Applied Biosystems) with 10 pmole PCR primers (CAAGCAGAAGACGGCATAACGAGATAAAGCTAGCTGATGCTACGACCACAGG, AATGATACGGCGACCACCGAGATCTACTCTTTCCTACACGACG). The PCR conditions were as follows: The reaction volume was 20 ml containing 1x PCR buffer provided by the manufacturer, 1.5 mM MgCl<sub>2</sub>, 0.4 mM dNTPs, and 1 ml circularized DNA. After activation of DNA polymerase by heating at 95°C for 5 min, DNA samples were amplified by 23 cycles of denaturation at 95°C for 15 seconds, annealing at 95°C for 15 seconds, and extension at 72°C for 30 seconds. PCR products were quantified using a Qubit fluorometer (Invitrogen) and separated by electrophoresis on a 2% agarose gel.

## Results

rRNAs and tRNAs, which comprise most of the RNAs in the cell, are highly modified. Although 5' ends of rRNAs and tRNAs are not modified, m7G modifications are present at internal positions. Since the presence of the m7G cap on mRNAs is useful for the determination of transcription start sites, mRNAs were enriched by the selection of RNAs containing consecutive adenylate nucleotides at the 3' end. Total RNA extracted from HeLa cells was hybridized with the oligo (dT)<sub>25</sub> residues covalently coupled to the surface of magnetic beads

(Invitrogen, Dynabeads mRNA purification kit) at room temperature for 5 min. The magnetic beads were washed twice and mRNA bound to beads was eluted by heating the sample at 65°C for 2 min. Total RNA, unbound RNA, and purified RNA were size-fractionated by electrophoresis on a 1% agarose gel (Fig 2). While most of the visible RNAs were rRNAs and tRNAs in total RNA and unbound RNA, mRNAs were enriched in RNA eluted from magnetic beads.

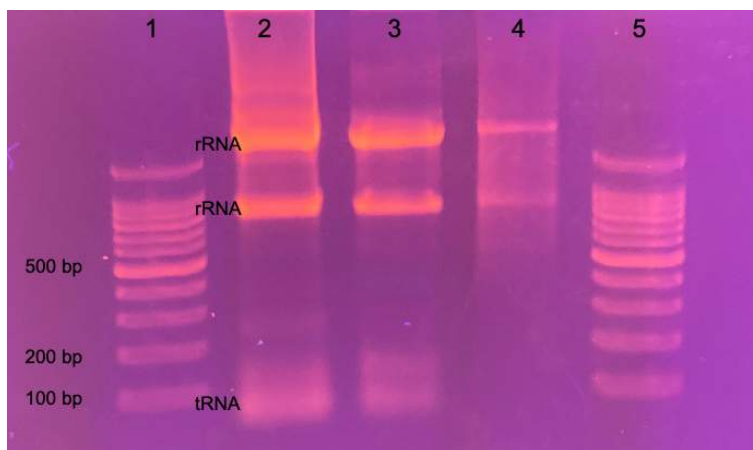


Fig 2: rRNAs and tRNAs are abundant in total RNA (lane 2) and unbound RNA (lane 3). Lane 4 shows that mRNAs that are heterogeneous in size are enriched after mRNA purification. Lanes 1 and 5 show DNA size markers. The total amounts of RNAs are 120 mg (total RNA), 72 mg (unbound RNA), and 2.2 mg (purified RNA).

For simultaneous PCR amplification of a variety of RNA fragments, the RNA fragments or the cDNAs should be flanked by sequences complementary to PCR primers. Since the m7G cap blocks the 5' end of the mRNA, determination of the 5' end requires special methods. If an adapter is linked to the 5' end of capped RNA, the m7G cap should be removed. Alternatively, an adapter is linked to the 3' end of cDNA produced by reverse transcription RNA. In this study, we used a method delineated in Fig 3. Instead of attaching an added adapter to the 3' end of cDNA, the 5' end of the primer used for reverse transcription was linked to the 3' end of cDNA by using a thermostable RNA ligase I having good single-strand DNA ligase activity [5]. Because the reaction rate of intramolecular ligation is independent of cDNA and linker concentrations, the reaction is efficient even with very low cDNA concentration.



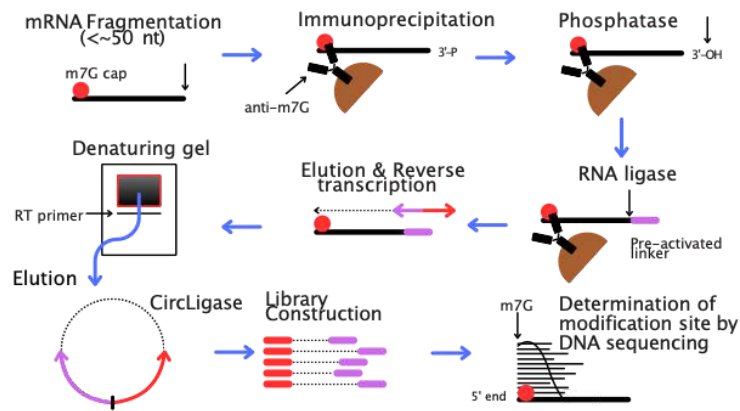


Fig 3: the strategy to make cDNA

There are two different mouse monoclonal antibodies available from MBL. An antibody (RN016M, clone 150-15) has a strong preference for the m7G cap. In contrast, the other antibody (RN017M, clone 4141-13) reacts with both 5'-terminal and internal m7G modifications in RNA. As a negative control, a parallel reaction with a non-specific IgG antibody was conducted. Ten percent of circularized cDNA was used for each PCR reaction with 23 cycles. The PCR-amplified samples were visualized by electrophoresis on a 2% agarose gel.

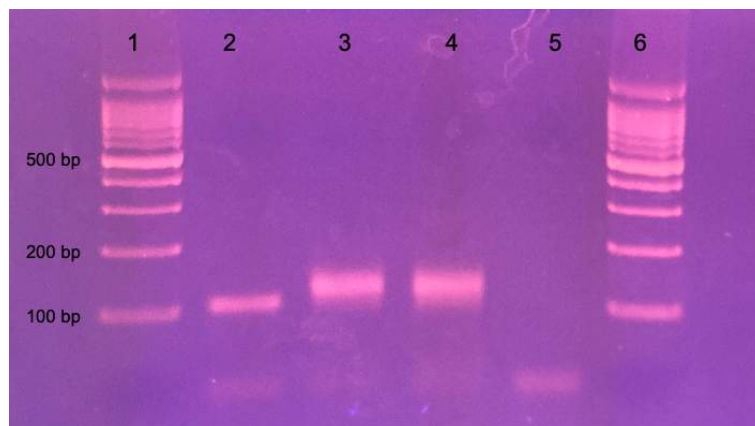


Fig 4: the distributions of size-fractionated PCR products

In Fig 4, The size of the band in the negative control reaction (lane 2) was approximately 100 base pairs, which suggests that the band was primer dimers devoid of cDNA. The PCR reaction without circularized DNA produced no primer dimers. In contrast, PCR products of reactions conducted with circularized cDNA derived from RNA selected by m7G-specific antibodies (lane 3 RN016M, lane 4 RN017M) were distributed in upper regions, corresponding to 120 to 150 base pairs, suggesting that cDNAs ranging from 20 to 50 nucleotides were amplified. Although the antibody (RN016M) has a preference for the m5G cap, the size distributions of PCR products were similar.

## **Discussion**

Information on transcription start sites is very important for understanding the mechanisms and regulation of transcription initiation by RNA polymerases, especially polymerase II. Although the amount of transcripts produced by RNA polymerase II is small, their diversity is incomparable to that of RNA produced by other RNA polymerases. These conditions pose difficulties in experiments for precise determination of the transcription start site. Several methods have been devised to determine start sites, but these methods have drawbacks. The methods are complicated and time-consuming. Determination of internal locations of 6-methyladenylate (m6A) on mRNA was achieved using an antibody that captures RNA fragments containing the modified nucleotide [6]. In this study, the procedure designed to determine the locations of m6A was also applied to m7G because antibodies against the m7G nucleotide have been available for a long time.

The experiment conducted in this study can be completed in 4-5 days from mRNA purification to PCR amplifications. Just as m6A locations on the human mRNA were determined by high-throughput DNA sequencing, positions of the m7G cap could be determined on a genome-scale by applying a high-throughput sequencing method. It would be interesting to compare the performance of the method used in this study with that of the best method that selects biotinylated m7G by using streptavidin-coated magnetic beads [3].

## **Acknowledgments**

We thank our mentor, Dayoung Yoon at the Department of Biomedical Sciences, College of Medicine, Korea University, for her guidance and providing of resources.

### Works Cited

- [1] R. G. Roeder and W. J. Rutter, "Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms", *Nature*, 1969, Vol. 224, Issue 5216, Pages 234-7.
- [2] Y. Furuichi, M. Morgan, A. J. Shatkin, W. Jelinek, M. Salditt-Georgieff and J. E. Darnell, "Methylated, blocked 5 termini in HeLa cell mRNA", *Proc. Natl. Acad. Sci. U. S. A.*, 1975, Vol. 72, Issue 5, Pages 1904-8.
- [3] X. Adiconis, A. L. Haber, S. K. Simmons, A. Levy Moonshine, Z. Ji, M. A. Busby, et al., "Comprehensive comparative analysis of 5'-end RNA-sequencing methods", *Nat. Methods*, 2018, Vol. 15, Issue 7, Pages 505-511.
- [4] T. W. Munns, M. K. Liszewski and H. F. Sims, "Characterization of antibodies specific for N6-methyladenosine and for 7-methylguanosine", *Biochemistry*, 1977, Vol. 16, Issue 10, Pages 2163-8.
- [5] T. Blondal, A. Thorisdottir, U. Unnsteinsdottir, S. Hjorleifsdottir, A. Aevarsson, S. Ernstsson, et al., "Isolation and characterization of a thermostable RNA ligase 1 from a *Thermus scotoductus* bacteriophage TS2126 with good single-stranded DNA ligation properties", *Nucleic Acids. Res.* 2005, Vol. 33, Issue 1, Pages 135-42.
- [6] D. Dominissini, S. Moshitch-Moshkovitz, S. Schwartz, M. Salmon-Divon, L. Ungar, S. Osenberg, et al., "Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq", *Nature*, 2012, Vol. 485, Issue 7397, Pages 201-6.

## Household Readiness for Induction Cooktop Adoption in India By Evan Obbard

### Abstract

This research project aims to discover whether households in India have sufficient electrical capacity to switch from gas cooking to induction cooking and, if so, to what extent. It focuses on the technical restrictions of households in peri-urban areas in Maharashtra, India. This paper was conducted due to the environmental and human health-related benefits of an induction stove compared to other forms of cooking such as wood, gas, or electric, and the potential advantages of incorporating induction stovetops into households worldwide. Data on respondent demographics, household composition, electrical system details, appliances, and safety conditions was collected from 123 households picked at random in multiple districts in Maharashtra. The main findings indicate that there is ample electrical capacity in the Main Circuit Breaker to carry an additional induction stovetop of varying wattages in almost all households tested in this study.

### Introduction

Wood-based cooking is still widely used in many parts of the world, particularly in rural and peri-urban areas of developing nations. It is currently the most widely used cooking fuel in India<sup>1</sup>. However, its environmental and health impacts are severe. The burning of wood releases large amounts of particulate matter (PM<sub>2.5</sub>), carbon monoxide (CO), and black carbon, all of which contribute to both air pollution and climate change<sup>2,3</sup>. Approximately 2.3 billion people rely on wood and other biomass for cooking, leading to the release of an estimated one gigaton of CO<sub>2</sub> per year<sup>4</sup>. Black carbon, in particular, is a potent short-lived climate pollutant that absorbs sunlight and contributes significantly to global warming, ranking second only to CO<sub>2</sub> in its contribution to climate change<sup>5</sup>.

The health implications are also severe. Exposure to indoor air pollution from wood stoves is linked to respiratory diseases, including chronic obstructive pulmonary disease (COPD), lung cancer, and childhood pneumonia<sup>4</sup>. The WHO estimates that indoor air pollution from biomass fuel use causes approximately 3.2 million premature deaths annually, with women and children being disproportionately affected<sup>4</sup>. Additionally, the collection of firewood leads to deforestation, reducing biodiversity and further exacerbating climate change<sup>30</sup>. Deforestation also contributes to soil erosion and decreases water quality, compounding environmental degradation<sup>6</sup>.

Gas-based cooking, while often considered cleaner than wood, still has significant environmental and health drawbacks. It is currently the second most widely used cooking fuel in India<sup>1</sup>. Natural gas stoves release methane<sup>7</sup>, a potent greenhouse gas with a global warming potential over 80 times greater than CO<sub>2</sub> during its first 20 years in the atmosphere<sup>8</sup>. Even when turned off, gas stoves have been found to leak methane into the atmosphere<sup>9</sup>, contributing to climate change. A 2022 study found that U.S. gas stoves alone emit 2.6 million tons of methane annually, equivalent to the emissions from approximately 500,000 gasoline-powered cars<sup>10</sup>. In

addition to methane, gas stoves emit nitrogen dioxide (NO<sub>2</sub>)<sup>11</sup>, which can reach levels indoors that exceed outdoor air quality standards<sup>12</sup>. Prolonged exposure to NO<sub>2</sub> is associated with respiratory issues, including asthma, particularly in children<sup>13</sup>.

The extraction and transportation of natural gas also contribute to its environmental impact. Hydraulic fracturing (fracking), often used to extract natural gas, releases methane during the process and can lead to groundwater contamination<sup>14</sup>. Furthermore, gas infrastructure tends to disproportionately impact marginalized communities, contributing to environmental justice concerns<sup>15, 16</sup>.

Transitioning to cleaner cooking technologies, such as induction cooktops, is essential for addressing both the public health and environmental challenges posed by traditional cooking methods. Induction cooking, which uses electromagnetic fields to heat cookware, eliminates the harmful emissions associated with wood and gas combustion. Because induction cooktops transfer energy directly to the cookware, they are far more efficient, using about 90% of the energy produced compared to just 40-60% for gas stoves<sup>17</sup>. This increased efficiency means less energy is wasted, reducing overall energy consumption and household carbon emissions. Induction cooktops also improve indoor air quality by eliminating combustion gases, a significant public health benefit<sup>17, 18</sup>.

India, with its rapidly growing population<sup>19</sup> and increasing energy demands<sup>20</sup>, is an extremely important place for the transition to cleaner technologies. This research focuses on peri-urban and urban areas of Maharashtra, where access to sustainable cooking methods could have profound impacts on both the environment and public health. In these regions, a shift to induction cooking is more plausible than in rural areas, due to greater access to technologies and electricity<sup>21</sup>.

However, one potential constraint to the incorporation of induction stovetops into households is the electrical capacity of the Main Circuit Breaker. The Main Circuit Breaker, or MCB, is a safety device that controls the flow of electricity in a household. It protects electrical circuits from overloads and short circuits by interrupting the flow of electricity when it exceeds a specified limit. Since an induction stovetop is powered by electricity, it contributes to the electrical load, thus making the amount of electricity it uses a possible constraint to its function in a household.

In this study, the primary objective is to evaluate the electrical capacity of households in peri-urban and urban regions of Maharashtra, India, specifically to determine their ability to support the integration of induction cooktops. The research focuses on identifying potential technical restrictions, particularly concerning the Main Circuit Breaker (MCB) capacity, which regulates the electrical load within households. By assessing this constraint, the study aims to provide insights into the feasibility of adopting induction cooking as a cleaner alternative to traditional gas and wood-based methods.

By showing that households in peri-urban and urban areas in Maharashtra have the electrical capacity to incorporate induction cooktops, both government and commercial efforts can be made to accelerate adoption. This also has the benefit of informing homeowners of the

benefits of induction stovetops, as well as the possible incorporation of these stovetops into their households.

## **Methods**

### **Research Design**

The research followed an observational design, aiming to collect data from households regarding their electrical infrastructure, appliance usage, and safety conditions. No interventions were made, and the data was collected at a single point in time, making it a cross-sectional study.

### **Population/Demographics**

The participants in this study were the heads of their households. Households surveyed were in Belatgavan, Chehadi, or Deshwandi, all of which are peri-urban districts outside Nashik, a city in Maharashtra. Maharashtra is the second most populous state in India<sup>29</sup>. These households were randomly selected inside these three districts, ensuring a representative sample of the local population.

### **Sample Selection**

The data was collected by Professor Narasimha Rao, the Associate Professor of Energy Systems at the Yale School of the Environment; however, the data has not been published. Rao states, "It is currently a work in progress." To collect all data, 123 households within Nashik and other towns close by were chosen at random. Prof. Narasimha et al. surveyed the head of the household on demographics, familial relations, and electrical system details. For this paper, data was collected on the electrical output of different appliances using a wattage meter. This tracks the wattage of each appliance. In addition, the MCB has a power rating, which was used to measure the MCB capacity. The non-coincident load, or the electrical load of all electrical appliances in each household, was found by adding the electrical load of all appliances.

### **Data Analysis**

Data was analyzed by adding up the electrical output of all the electrical appliances in a given household. The electrical output of all appliances in each household was tracked. The electrical capacity of the microwave, TV, refrigerator, washing machine, electric geyser, immersion rod, kettle, mixer, and rice cooker were tracked.

The electrical output was determined in watts. The wattage was found by considering the pressure, or force, at which the electrical current was flowing (Volts), as well as the rate at which it was flowing (Amps). All appliances were flowing with a force of 240 Volts, so the amps of each appliance were multiplied by 240 volts, resulting in the value in watts. These wattage values represent the total power flow rate through a given system, with the systems being each of the electrical appliances. The wattage values of all appliances in a household were added up to calculate the non-coincident load. This non-coincident load was compared to the MCB value for

each household in the data set obtained by Professor Rao. These values were compared with and without an additional induction stove, the value of which would be added on top of the pre-existing non-coincident load. 36 households were removed from the analysis due to data collection errors, such as a non-coincident load of 0 watts or no input MCB capacity value. This left 97 out of the 123 households to be analyzed.

### **Graph & Chart Analysis + Creation**

Javascript was used to analyze the electrical capacity of households' MCBs and compare it to the non-coincident load with the addition of various value induction stovetops. HTML was used to visually represent the charts and graphs, as it allows for a user interface. This analysis was implemented through the use of VisualStudioCode.

Additional induction stovetop watt values of 1300, 1800, 2100, and 2500 were chosen due to these being some of the most commonly used induction stove power values<sup>31, 32, 33</sup>. This provides for the most useful and accurate results.

#### **A. Graph for Non-Coincident Load Without Induction Cooker Added**

JavaScript was used to analyze the available electrical capacity of households' MCBs. The first step involved calculating the non-coincident load for each household without an induction cooker. This was accomplished by summing the wattages of all appliances that could potentially be used simultaneously, represented by the formula:

$$\text{Total Power} = P_{\text{fridge}} + P_{\text{TV}} + P_{\text{microwave}} + P_{\text{kettle}} + P_{\text{mixer}} + P_{\text{rice cooker}} + P_{\text{immersion rod}} + P_{\text{washer}} + P_{\text{geyser}}$$

where each P represents the power consumption of the respective appliance. After calculating the total power consumption, this value was divided by the MCB wattage capacity to determine the percentage of the load that the MCB could safely handle.

This analysis focused on a range from 0% to 100% of the non-coincident load in use because only up to 100% of appliances can be running. To quantify this, a boolean system was implemented where each household received a "true" designation if the calculated MCB percentage was under 100% and "false" if it was over. By aggregating these boolean values across all households, the percentage of households that could sustain their non-coincident loads safely was calculated.

The graph, created using JavaScript for dynamic data handling and displayed in HTML, shows the percentage of households on the y-axis against the respective non-coincident load percentages on the x-axis. This visualization provides clear insights into the electrical safety margins for households without the added load of an induction cooker, illustrating how many households remain within safe operating limits and emphasizing the importance of considering typical usage patterns in assessing electrical capacity.

## **B. Graph for Non-Coincident Load With Induction Cookers Added**

Building on the previous analysis, this process was repeated to evaluate how the addition of induction cookers—specifically at wattages of 1300W, 1800W, 2100W, and 2500W—affected household electrical capacity. Using the same JavaScript and HTML framework, the non-coincident load was recalculated by determining the percentage of the non-coincident load that each household could theoretically sustain, again ranging from 0% to 100%.

The rationale for this range remained consistent: households rarely operate at maximum capacity, and this analysis sought to capture a realistic picture of household electrical safety. The total power calculation was modified as follows:

$$\text{Total Power} = (\text{Non-Coincident Load} \times \text{Percentage}) + P_{\text{induction}}$$

$P_{\text{induction}}$  is the fixed wattage of the induction cooker, operating independently of the household's non-coincident load percentage. This distinction was essential because the induction cooker would always be running at its full wattage.

After calculating the total power consumption for each scenario, this value was compared against the MCB capacity using the same boolean evaluation as before. The “true” values were aggregated across households, and data was plotted using HTML for the graphical representation, with the y-axis representing the percentage of households and the x-axis showing the percentage of non-coincident load.

### **Chart Creation**

In addition to the overall graph, a detailed chart was created that specifically visualized the relationship between MCB capacity and non-coincident load with and without an induction cooker. This chart employed a similar methodology to the graph but focused on grouping households based on their load percentages into discrete buckets for clarity. These charts are useful for understanding how many households can sustain their non-coincident load.

## **C. Chart With No Induction Cooker**

Each bucket represented a range of percentages (e.g., 0-5%, 5-10%, etc.), allowing for an easier interpretation of how many households fell within each range. The total % of the MCB power that the non-coincident load used determined which bucket the household falls into. For example, if the entire non-coincident load used 52% of the MCB capacity, it would fall into the 50-55% bucket. The greater the number of households in a given bucket, the taller the column became to visualize the number of households that fell into each percentage category. The axes for this chart displayed the percentage ranges on the x-axis and the number of households on the y-axis. This design aimed to facilitate comparisons across different load scenarios while providing a visual summary of household safety profiles.

## **D. Charts for Induction Cookers Added (1300W, 1800W, 2100W, 2500W)**

The individual charts correspond to each induction stovetop wattage. The methodology



mirrored that of the chart for 0W, but it focused on the impacts of each specific induction cooker value. For each wattage scenario, the percentage of the MCB that was being used when the entire non-coincident load was running was calculated, and given that percentage, the household would fall into a certain bucket. If the household falls into a bucket over 100%, this means that the MCB is not able to sustain the entire coincident load. Visually, these buckets are red to convey how the household is not able to sustain that induction stovetop in addition to their entire non-coincident load.

The axes of these charts were structured similarly to the previous ones: the x-axis represented the percentage of non-coincident load, while the y-axis indicated the number of households capable of sustaining that load with the induction cooker added.

### Table Creation

Tables were created for both the charts and the graph. The tables give a numerical breakdown of the charts and graphs, allowing for a more direct correlation between the data analyzed and their implications.

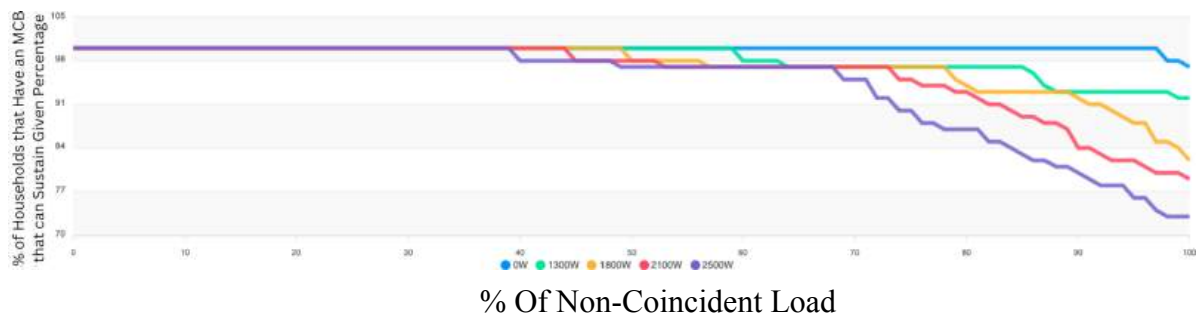
### Tables for Figure 1

Figure 1 contains five different graphs, all overlapping. Due to this, five different tables were created, one for each of the induction stovetop values.

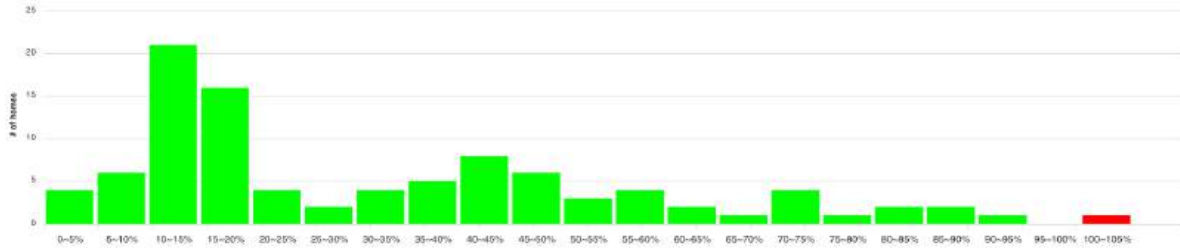
### Table for Charts

A table was created for all charts. This table allows for an easier comparison between each chart and a visualization of the decrease in homes that are able to sustain their entire non-coincident load while using an increasingly higher wattage induction stove.

## RESULTS

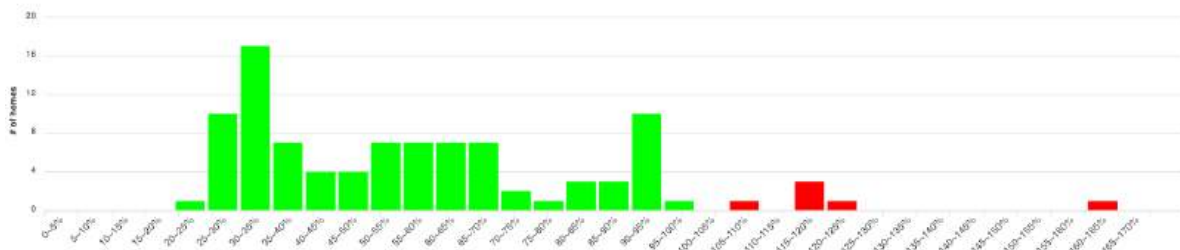


**Figure 1. Percentage of Households Whose MCB can Sustain Percentage of Non-Coincident Load - Various Induction Stovetops Running**



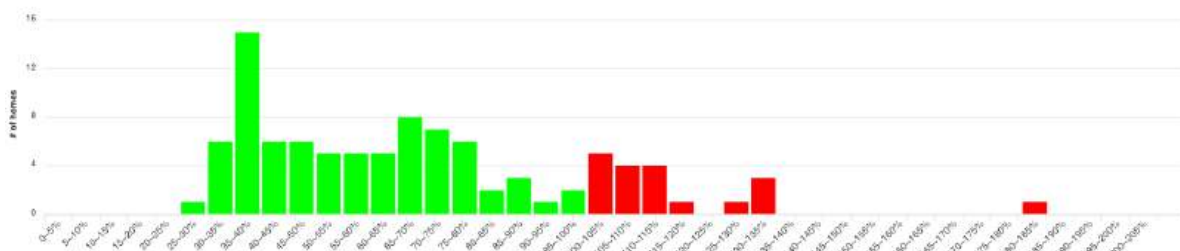
**Percentage of MCB Capacity Utilized by Full Non-Coincident Load**

**Figure 2. # Of Homes that Use Certain Percentages of their MCB While Running Non-Coincident Load- No Additional Induction Stove**



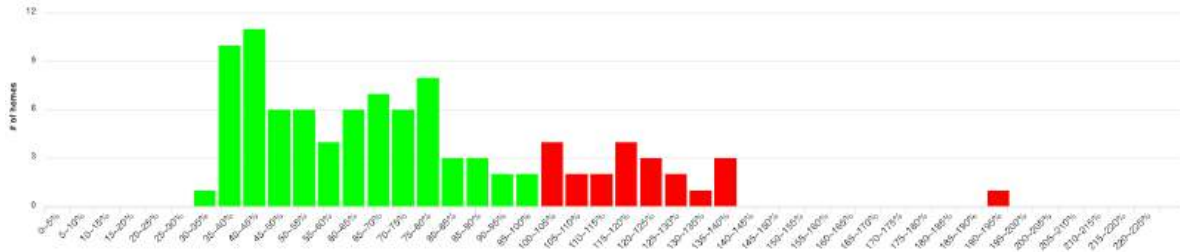
**Percentage of MCB Capacity Utilized by Full Non-Coincident Load**

**Figure 3. # Of Homes that Use Certain Percentages of their MCB While Running Non-Coincident Load- Additional 1300 Watt Induction Stove**



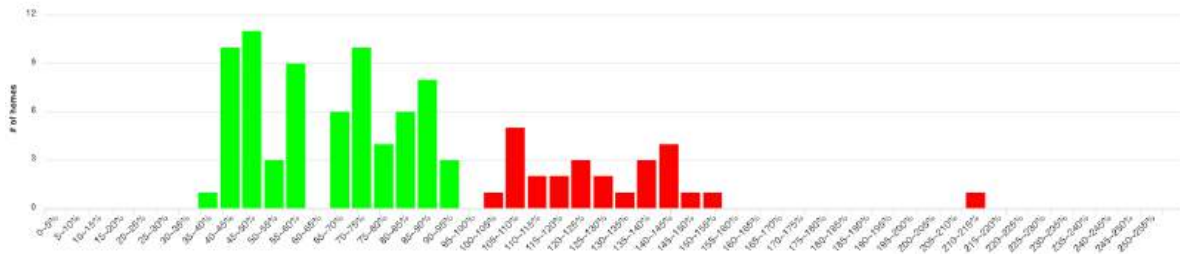
**Percentage of MCB Capacity Utilized by Full Non-Coincident Load**

**Figure 4. # Of Homes that Use Certain Percentages of their MCB While Running Non-Coincident Load- Additional 1800 Watt Induction Stove**



Percentage of MCB Capacity Utilized by Full Non-Coincident Load

Figure 5. # Of Homes that Use Certain Percentages of their MCB While Running Non-Coincident Load- Additional 2100 Watt Induction Stove



Percentage of MCB Capacity Utilized by Full Non-Coincident Load

Figure 6. # Of Homes that Use Certain Percentages of their MCB While Running Non-Coincident Load- Additional 2500 Watt Induction Stove

Table 1: Percentage of Households Whose MCB can Sustain Percentage of Non-Coincident Load - No Induction Stove

Percentage of Non-Coincident Load:	Percentage of Households that can Sustain:
0-97	100
98-99	98
100	97

Table 2: Percentage of Households Whose MCB can Sustain Percentage of Non-Coincident Load - 1300 Watt induction Stove

Percentage of Non-Coincident Load (1300 W):	Percentage of Households that can Sustain:
0-59	100
60-63	98
64-85	97
86	96
87	94
88-98	93
99-100	92

**Table 3: Percentage of Households Whose MCB can Sustain Percentage of Non-Coincident Load - 1800 Watt Induction Stove**

Percentage of Non-Coincident Load (1800W):	Percentage of Households that can Sustain:
0-49	100
50-56	98
57-78	97
79	95
80	94
81-89	93
90-92	91
93	90
94	89
95-96	88
97-98	85
99	84
100	82

**Table 4: Percentage of Households Whose MCB can Sustain Percentage of Non-Coincident Load - 2100 Watt Induction Stove**

Percentage of Non-Coincident Load (2100 W):	Percentage of Households that can Sustain:
0-44	100
45-52	98
53-73	97
74-75	95
76-78	94
79-80	93
81	92
82-83	91
84	90
85-86	89
87-88	88
89	87
90-91	84
92	83
93-95	82
96	81
97-99	80
100	79

**Table 5: Percentage of Households Whose MCB can Sustain Percentage of Non-Coincident Load - 2500 Watt Induction Stove**

Percentage of Non-Coincident Load (2500 W):	Percentage of Households that can Sustain:
0-39	100
40-48	98
49-68	97
69-71	95
72-73	92
74-75	90
76-77	88
78-81	87
82-83	85
84	84
85	83
86-87	82
88-89	81
90	80
91	79
92-94	78
95-96	76
97	74
98-100	73

**Table 6: Percentage of Households Able to Sustain Non-Coincident Load - Varying Induction Stovetops Running**

Added Induction Stovetop Power (Watts)	Percentage of Houses Able to Sustain Entire Non-Coincident Load:
0	98.9% (96/97)
1300	93.8% (91/97)
1800	80.4% (78/97)
2100	77.3% (75/97)
2500	73.2% (71/97)

**Results**

This study finds that the majority of households are able to sustain their entire non-coincident load, while simultaneously running a high-power induction stove (Table 6). All households can sustain 50% of their coincident load without an induction stove or with a 1300W induction stove (Figure 6, Table 1, Table 2). With an 1800W or 2100W cooktop, 98% of households can sustain 50% of their load (Figure 6, Table 3, Table 4), while 97% can do so with a 2500W cooktop (Figure 6, Table 5). Because the majority of households in this study have enough excess electrical capacity for an induction stove, many homes have the ability to switch from gas- or solid-fuel cooking to induction. This information is useful because it lets governments know that they can subsidize and commercialize induction stovetops since they now know houses can sustain them. This is also useful because it lets companies know that they can create and sell them, it lets stores know they can stock them, and it lets customers know that they can buy them, given that the reason they were holding back is that they didn't know if their houses could.

Increased data on peak loads throughout the day in households could strengthen this study. Since there is no data on the maximum percentage of the MCB that is used at a certain point in the day, the results finding that over 95% of households can sustain 50% of their non-coincident load while running an 2500 watt or less induction cooktop (Figure 6, Table 1, Table 2, Table 3, Table 4, Table 5) does not allow for complete certainty that this is a realistic percentage of a households' MCB in use. It seems reasonable to assume that it is extremely unlikely that households are using more than 50% of their MCB capacity due to the number of appliances that would be required. However, without additional studies on this, it cannot be determined for certain that this is realistic.

## **Discussion**

This study strengthens the evidence base showing the readiness of urban and peri-urban Indian households to adopt induction cooking. This paper examined one of the many technical barriers to adoption. To fully understand the technical barriers to the adoption of these more environmentally friendly and human-healthy stoves, additional studies could look at other technical constraints, such as the sanctioned load, theoretical plug points for an induction cooktop, wire capacity, and so on. These barriers can be analyzed with this same data set, as information on these factors was recorded. However, for a broader analysis, it would be beneficial to get data for a larger set of households.

This study also only delves into the technical side of adoption. In order to further acknowledge the implications of induction, looking at the political and social constraints to adding these cooktops is necessary. This could include research on economic factors surrounding the incorporation of induction stoves. One could delve into the additional funds required for households to purchase induction stoves and research if households have excess funds for these stoves.

## **Political Constraints**

One of the most critical political constraints for incorporating induction stoves is energy policy and grid capacity. India's electricity grid, particularly in rural and peri-urban areas, faces challenges related to reliability, capacity, and distribution<sup>22, 23, 24</sup>. As the demand for electricity increases with the adoption of electric cooking technologies like induction stoves, it is necessary that the energy grid can handle the additional load. Governments need to invest in upgrading grid infrastructure, including expanding renewable energy sources, to ensure that the grid remains stable and can provide consistent power to homes. Without these upgrades, rolling blackouts and power shortages could make the widespread adoption of induction stoves unfeasible. Additionally, energy policies that support renewable energy could reduce the environmental impact of increased electricity consumption, which aligns with broader goals for decarbonization and energy efficiency.

Another significant political constraint is government subsidies and tariffs. The affordability of induction stoves largely depends on the cost of electricity, the availability of subsidies, and the import tariffs on induction cookers. Political will is needed to provide financial incentives or subsidies to low-income households, encouraging them to switch from traditional cooking methods (such as wood or gas) to electric cooking. Furthermore, reducing import duties on induction stoves could make them more affordable for the general population. Without political action to provide these incentives, the initial cost of adopting induction stoves could prevent many households from making the switch, stalling progress in modernizing cooking infrastructure.



## **Social Constraints**

One of the most prominent social barriers is cultural resistance to change in cooking practices<sup>25</sup>. Many households in India, particularly in rural areas, do not utilize electricity when cooking<sup>26,27</sup>. These methods are deeply embedded in daily life, and the transition to induction stoves may not be readily accepted without education or demonstration of benefits. For example, open-flame cooking is often associated with better flavor in some dishes<sup>28</sup>, and shifting to induction stoves might be perceived as altering important cultural traditions. For the adoption of induction stoves to be successful, community education and outreach campaigns are essential to demonstrate the benefits of this technology, such as faster cooking times, safety, and reduced indoor air pollution.

Awareness and education are also critical social constraints. Many households may not be aware of how induction stoves work or their long-term economic and health benefits. Some families may resist using induction stoves due to a lack of understanding of their functionality or because they fear unfamiliar technology. Educational initiatives, possibly government-led or supported by NGOs, could help bridge this gap by informing families about the advantages of induction cooking, such as its energy efficiency and ease of use. Workshops or subsidies that make induction stoves accessible, paired with demonstrations, could be useful in shifting social attitudes toward electric cooking.

Additionally, this analysis took place with only 97 households in various towns in Maharashtra, India, the majority of which were in Nashik, India. Due to this, additional studies on household electrical capacity could be taken to confirm this analysis and expand it to other regions.

## **Conclusion**

In conclusion, as households increasingly seek sustainable living solutions, the integration of induction cooking technology is a promising path forward. This research not only highlights the feasibility of these transitions but also highlights the prevalence of climate change in this day and age. As we move forward, climate change becomes a bigger and bigger threat to humanity. By switching to more renewable technology, such as induction stoves, carbon emissions can be mitigated and progress towards a greener future can be made.

### Work Cited

1. "Percentage Share of Households by Type of Cooking Fuel - India Total 2011." ResearchGate, [https://www.researchgate.net/figure/Percentage-Share-of-Households-by-Type-of-Cooking-Fuel-India-Total2011-Source\\_fig1\\_299454400](https://www.researchgate.net/figure/Percentage-Share-of-Households-by-Type-of-Cooking-Fuel-India-Total2011-Source_fig1_299454400). Accessed 29 Sept. 2024.
2. "Black Carbon." Climate & Clean Air Coalition, <https://www.ccacoalition.org/short-lived-climate-pollutants/black-carbon>. Accessed 29 Sept. 2024.
3. "Smoke from Residential Wood Burning." Environmental Protection Agency, <https://www.epa.gov/indoor-air-quality-iaq/smoke-residential-wood-burning>. Accessed 29 Sept. 2024.
4. "Household Air Pollution and Health." World Health Organization, <https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health>. Accessed 29 Sept. 2024.
5. "EEA Black Carbon Report Highlights Health and Climate Change Impacts." International Institute for Sustainable Development, 10 Dec. 2013, <https://sdg.iisd.org/news/eea-black-carbon-report-highlights-health-and-climate-change-impacts/>. Accessed 29 Sept. 2024.
6. "Deforestation: Facts, Causes & Effects." National Geographic, <https://www.nationalgeographic.com/environment/article/deforestation>. Accessed 29 Sept. 2024.
7. Stanford Woods Institute for the Environment. "Stanford Scientists Find Climate and Health Impacts of Natural Gas Stoves Are Greater Than Previously Thought." Stanford University, 27 Jan. 2022, <https://woods.stanford.edu/news-events/news/stanford-scientists-find-climate-and-health-impacts-natural-gas-stoves-are-greater>. Accessed 29 Sept. 2024.
8. "Methane: A Crucial Opportunity in the Climate Fight." Environmental Defense Fund, <https://www.edf.org/climate/methane-crucial-opportunity-climate-fight>. Accessed 29 Sept. 2024.
9. Borgeson, Merrian. "Gas Stoves Emit Pollution Even When Not in Use." Natural Resources Defense Council, 27 Jan. 2022, <https://www.nrdc.org/bio/merrian-borgeson/gas-stoves-emit-pollution-even-when-not-use>. Accessed 29 Sept. 2024.
10. "Gas Stoves Found to Emit Methane Equivalent to 500,000 Cars." World Economic Forum, 4 Feb. 2022, <https://www.weforum.org/agenda/2022/02/gas-stoves-climate-change-study-methane/>. Accessed 29 Sept. 2024.
11. "People with Gas and Propane Stoves Breathe More Unhealthy Nitrogen Dioxide." Stanford News, 10 May 2024, <https://news.stanford.edu/stories/2024/05/people-with-gas-and-propane-stoves-breathe-more-unhealthy-nitrogen-dioxide/>. Accessed 29 Sept. 2024.

12. "Nitrogen Dioxide's Impact on Indoor Air Quality." Environmental Protection Agency, <https://www.epa.gov/indoor-air-quality-iaq/nitrogen-dioxides-impact-indoor-air-quality>. Accessed 29 Sept. 2024.
13. "The Health Impact of Nitrogen Dioxide." National Center for Biotechnology Information, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7112409/>. Accessed 29 Sept. 2024.
14. "Environmental Impacts of Fracking." Greenpeace USA, <https://www.greenpeace.org/usa/fighting-climate-chaos/issues/fracking/environmental-impacts-water/>. Accessed 29 Sept. 2024.
15. "Fossil Fuel Racism." Greenpeace USA, <https://www.greenpeace.org/usa/reports/fossil-fuel-racism/>. Accessed 29 Sept. 2024.
16. "Methane Gas Leaks Present Environmental Justice Concerns." Environmental Defense Fund, 11 May 2022, <https://blogs.edf.org/energyexchange/2022/05/11/methane-gas-leaks-present-environmental-justice-concerns/>. Accessed 29 Sept. 2024.
17. "Making the Switch to Induction Stoves or Cooktops." U.S. Department of Energy, 11 May 2023, <https://www.energy.gov/articles/making-switch-induction-stoves-or-cooktops>. Accessed 29 Sept. 2024.
18. Coren, Michael J., et al. "Are Gas Stoves Dangerous? We Tested Against Induction Cooktops to Find Out." *The Washington Post*, 16 Jan. 2024, <https://www.washingtonpost.com/climate-environment/2024/01/16/gas-stoves-induction-cooktop-health/>. Accessed 29 Sept. 2024.
19. "India Population 2024 (Live)." *World Population Review*, <https://worldpopulationreview.com/countries/india>. Accessed 29 Sept. 2024.
20. "India to Post an 8% Rise in Electricity Consumption in 2024: IEA." *Business Today*, 19 July 2024, <https://www.businesstoday.in/industry/energy/story/india-to-post-an-8-rise-in-electricity-consumption-in-2024-ia-437788-2024-07-19>. Accessed 29 Sept. 2024.
21. "Impacts of Energy Access: Differences between Rural and Urban Energy Access." *Energypedia*, [https://energypedia.info/index.php?title=Impacts\\_of\\_energy\\_access:\\_Differences\\_between\\_Rural\\_and\\_Urban\\_Energy\\_Access](https://energypedia.info/index.php?title=Impacts_of_energy_access:_Differences_between_Rural_and_Urban_Energy_Access). Accessed 29 Sept. 2024.
22. "Ensuring Grid Reliability in India." RMI, <https://rmi.org/ensuring-grid-reliability-in-india/>. Accessed 29 Sept. 2024.
23. Raizada, Swasti, et al. "Mapping India's Energy Policy 2023." *International Institute for Sustainable Development*, <https://www.iisd.org/story/mapping-india-energy-policy-2023/>. Accessed 29 Sept. 2024.
24. Jain, M., Rao, A.B., and Patwardhan, A. "Energy Efficiency Policies in India: Implications for Climate Change Mitigation." In *Climate Change Signals and Response*, edited by Venkataraman, C., Mishra, T., Ghosh, S., and Karmakar, S. Springer, Singapore, 2019, [https://doi.org/10.1007/978-981-13-0280-0\\_18](https://doi.org/10.1007/978-981-13-0280-0_18). Accessed 29 Sept. 2024.

25. "The Clean Cooking Revolution in India: Solutions to Overcome Persistent Barriers." EKI Energy Services Limited,  
<https://enkingint.org/the-clean-cooking-revolution-in-india-solutions-to-overcome-persistent-barriers/>. Accessed 29 Sept. 2024.
26. Dhamija, Dhruvika. "How Many Indian Households Use Clean Fuel for Cooking?" CEDA, 10 May 2023,  
<https://ceda.ashoka.edu.in/how-many-indian-households-use-clean-fuel-for-cooking/#:~:text=In%20rural%20areas%2C%20almost%20half,their%20primary%20fuel%20for%20cooking>. Accessed 29 Sept. 2024.
27. Agrawal, Shalu, et al. "Are Indian Homes Ready for Electric Cooking?" Council on Energy, Environment and Water, 15 Oct. 2021,  
<https://www.ceew.in/publications/are-indian-homes-ready-for-electric-cooking-transition>. Accessed 29 Sept. 2024.
28. Nolin, Kristine. "The Smoky Science Behind What Makes Food Grilled Over an Open Flame Taste So Good." PBS, 2 July 2023,  
<https://www.pbs.org/newshour/science/the-smoky-science-behind-what-makes-food-grilled-over-an-open-flame-taste-so-good>
29. "Maharashtra Population 2024." World Population Review,  
<https://worldpopulationreview.com/cities/india/maharashtra>. Accessed 29 Sept. 2024.
30. Shin, Yunne-Jai, et al. "Actions to Halt Biodiversity Loss Generally Benefit the Climate." *Global Change Biology*, vol. 28, no. 9, 2022, pp. 2846-2874,  
<https://doi.org/10.1111/gcb.16109>. Accessed 29 Sept. 2024.
31. Agarwal, Dibesh, and Consumer Advise Team. "Best Wattage for Induction Cooktop: 1200 vs 1800 vs 2000 Watts." Consumer Advise, 14 Feb. 2024,  
[www.consumeradvise.in/best-wattage-for-induction-cooktop/](http://www.consumeradvise.in/best-wattage-for-induction-cooktop/).
32. S, Nimesh. "Induction Stove Wattage: Best Watts for Indian Cooking." Lazy Dad Reviews, 26 Oct. 2022, [www.lazydadreviews.org/induction-cooktop-wattage/](http://www.lazydadreviews.org/induction-cooktop-wattage/).
33. Pothan, Parvathy. "6 Best Induction Cooktops in India." Everything Better, 8 Sept. 2024,  
[everythingbetter.in/best-induction-cooktop-stove-in-india/](http://everythingbetter.in/best-induction-cooktop-stove-in-india/).

## **Evaluating Plant and Bird Species Diversity and Abundance and Trash Pollution at Three Nature Reserves in Singapore By Yiqing Gong**

### **Abstract**

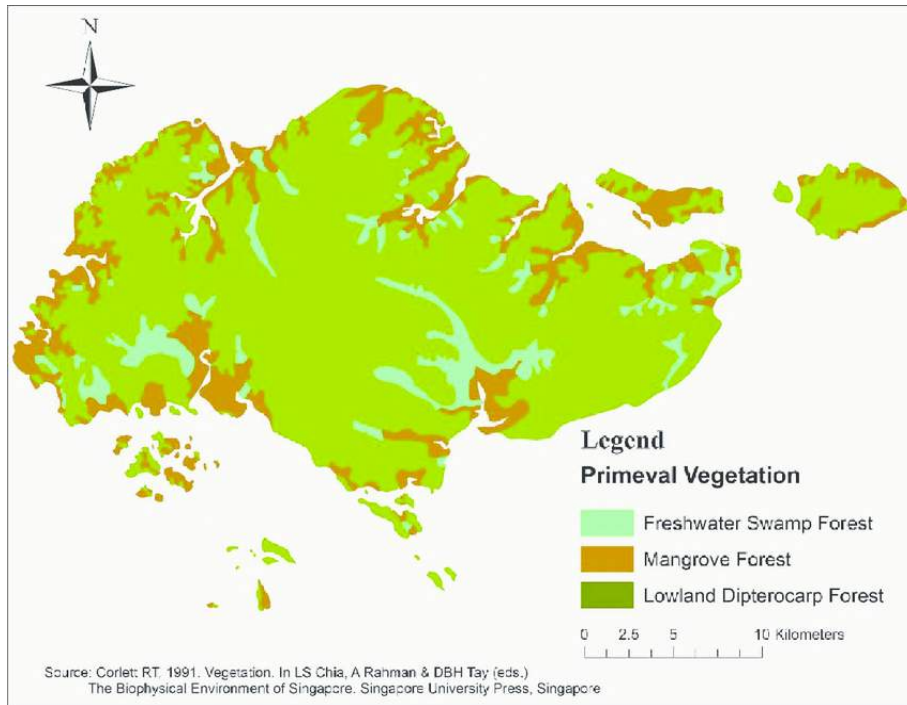
Singapore contains over 40,000 non-microbial organisms, many of which are nationally and/or regionally threatened. Significant ecosystems in Singapore include primary/old secondary forest, coastal hill secondary forest, and mangroves, each representing a unique portion of Singapore's native tropical ecology. This led to Labrador Nature Reserve, Bukit Timah Nature Reserve, and Sungei Buloh Wetland Reserve being chosen as field sites for understanding bird and plant diversity and abundance, and trash pollution. Bird and plant species were recorded using a transect and/or trail observation method at least thrice per field site. Data shows species diversity, species abundance, and locations of wildlife and trash pollution, allowing for analysis and comparisons between the ecosystems, as well as relationships between the data sets. This survey found a total of 31 bird species and 20 plant species across the three ecosystems and that primary/old secondary forest was overall more species diverse than coastal hill secondary forest over the period of surveying. This included seven notable migratory bird species present at the mangrove ecosystem, as well as defining organisms of the ecosystems and globally threatened species. Trash pollution was also found to be distributed more heavily in areas with heavy tidal movement at the mangrove ecosystem, however there was no relationship observed between trash pollution and number of birds. A clear positive relationship was found between plant and bird species diversity, however the three ecosystems differ greatly with regards to the relationship between plant and bird abundance. Overall, this study seeks to answer a number of ecological questions regarding biodiversity, ecosystem change through time, and species interactions, as well as collect data regarding the influence of human activity on Singapore's native ecology.

**Keywords.** Singapore, biodiversity, primary/old secondary forest, coastal hill secondary forest, mangrove, trash pollution

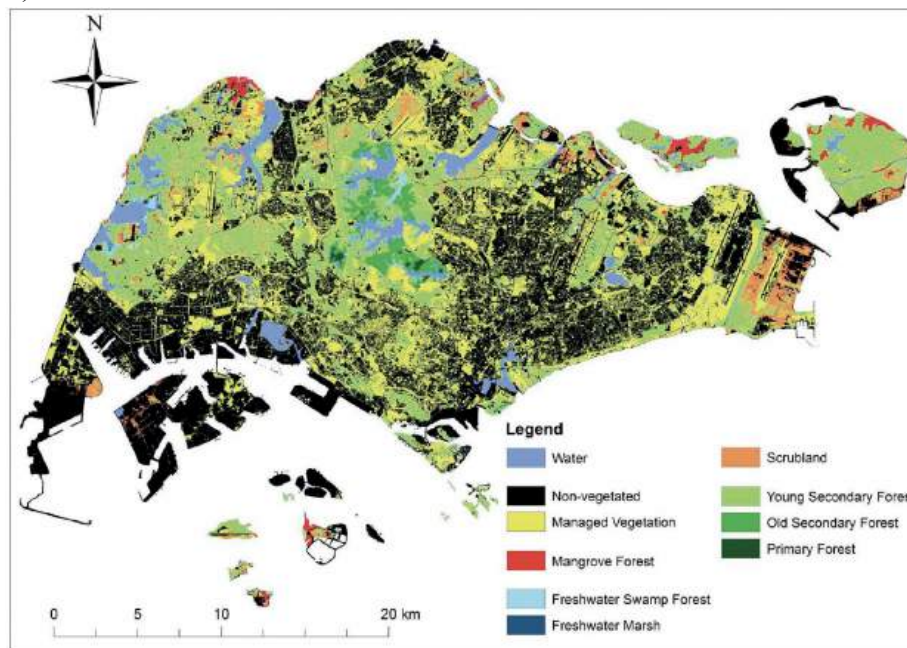
### **Introduction**

The city-state of Singapore is situated in South East Asia below the Malaysian Peninsula and near the equator. Prior to large-scale human interference, Singapore's coasts were covered with mangrove forest (13% of total forest cover), and its inland portion was dominated by lowland dipterocarp forest and some freshwater swamp forest (Corlett 411) (figure 1). As large-scale clearing of forests and land reclamation have transformed the biogeography of Singapore, only a few fragments of old-growth forest remain (figure 2). This has unfortunately restricted species (particularly endemic species highly specialized to a certain ecosystem) to comparatively small areas of habitat (Noreen and Webb 1). Today Singapore contains four government-designated Nature Reserves and 20 Nature Areas that represent many of these ecosystems, which include primary and secondary forest, mangroves, freshwater swamp forest,

and coral reefs (National Parks Board Singapore 12). They therefore play an instrumental part in protecting and ensuring diverse ecosystems and places of study.



(Yee et al. 206)



**Fig. 2: Current vegetation of Singapore**

(Yee et al. 209)

Singapore’s ecosystem diversity created high species diversity, with 3729 vascular plant

species and 404 bird species as of 2018 (National Parks Board Singapore 12), and the International Union for Conservation of Nature (IUCN) classifies some of these species between Near Threatened and Critically Endangered (IUCN). This diversity of species, both endemic and introduced, makes Singapore an ecologically significant and interesting area to research.

There are also many concerns with regards to Singapore's ecology, one of which is the threat of introduced and invasive species. Existing studies note that invasive species biologically invade ecosystems largely by outcompeting native species due to possessing different adaptations and competing with native species for resources such as nesting space, ultimately jeopardizing native biodiversity (Peh 1086). Specifically, Lim et al. found that the *Corvus splendens* and *Acridotheres javanicus* populations have grown substantially since their introduction, becoming widely observed across Singapore (Lim et al. 692). Tan and Tan highlight that extended periods of human disturbance, which is occurring in most areas in Singapore, encourages the proliferation of introduced species, suggesting that these species are a significant threat towards biodiversity of endemic species (Tan and Tan 60). This emphasizes the value of nature reserves as remnants of larger ecosystems. Despite being quite fragmented and covering small areas, they are the few remaining areas in Singapore that are minimally disturbed, therefore resisting invasive species and providing native species with habitat. Furthermore, most nature reserves in Singapore are quite fragmented and cover small areas.

Another concern is the impact of trash pollution. One study has found microplastics distributed in all waters surrounding Singapore, including those made from polypropylene, polyethylene, thermoplastic polyester, and foam (Curren and Leong 4). Some impacts of trash pollution on ecosystems include animals consuming and becoming entangled in trash pollution which decreases movement, feeding, and reproduction, threatening their populations. Trash pollution has also harmed plants through decreasing their growth, diversity, and health (Gondal et al. 6902-6903). These effects are particularly pronounced in aquatic ecosystems, which makes the nature reserves in Singapore particularly valuable to study as they all lie within close proximity of water, and the lack of existing research specific to Singapore in particular prompts the study of how trash pollution impacts local ecology and discussion for why keeping trash pollution under control is important for the health of aquatic ecosystems.

The main motivation of this study was to analyze plant and bird species diversity across different ecosystems in Singapore, as well as to investigate the influence of trash pollution on these ecosystems. This study is driven by the broad questions:

How do the primary and old secondary, secondary, and mangrove ecosystems present at Bukit Timah, Labrador, and Sungei Buloh Wetland Reserves in Singapore differ in terms of plant and bird species and trash pollution? Are there any relationships?

In order to address these driving questions, I studied three specific nature reserves in Singapore (figures 3-5, table 1), which play an instrumental part in protecting and ensuring diverse ecosystems and places of study.

Labrador Nature Reserve (field site A with 2 survey locations) is situated at the south end of Singapore and consists of secondary coastal hill forest (Low and Lim 1). These ecosystems tend to be characterized by vegetation growing near the sea, and organisms are tolerant to salt and poor soil (Ng and Siew). This field site is known to contain: plants *Tristaniopsis obovata* (critically endangered) (Ashton 274), *Syzygium grande*, *Ixonanthes reticulata*, *Rhodamnia cinerea*, *Terminalia catappa*, *Dipteris conjugata*; birds *Malacocincla abbotti*, and *Loriculus galgulus* (Yam et al. 265).

Bukit Timah Nature Reserve (field site B with 2 survey locations) is situated in the central region of Singapore and consists of primary and old secondary forest (both long-term undisturbed and recently planted trees) (Ho et al. 41). These ecosystems tend to be characterized by endemic dipterocarp trees and high species diversity (Goh). This field site is known to contain: plants *Shorea curtisii* (endangered), *Dipterocarpus sublamellatus* (endangered); birds *Pycnonotus zeylanicus* (critically endangered), *Cyornis brunneatus* (vulnerable), *Irena puella*, and *Psilopogon rafflesii* (near threatened) (Lim 204).

Sungei Buloh Wetland Reserve (field site C with 1 survey location) is situated in the North West of Singapore next to the Malaysian Peninsula and consists of mangrove trees surrounding wet mudflats (National Parks Board Singapore). These ecosystems tend to be characterized by a mixture of freshwater and saltwater. This field site is known to contain: plants *Bruguiera gymnorhiza*, *Heritiera littoralis* (endangered), *Ceriops zippeliana*, *Cassine viburnifolia* (endangered), *Sonneratia alba*, *Xylocarpus granatum*; birds *Tringa totanus*, *Tringa nebularia*, *Numenius phaeopus*, *Pluvialis fulva*, *Egretta garzetta*, and *Haliaeetus leucogaster* (Nai).

As primary/old secondary forest (Bukit Timah Nature Reserve/Field Site B) is older than secondary coastal hill forest (Labrador Nature Reserve/Field Site A), I hypothesized that I would find higher plant and bird species diversity in the former ecosystem. I hypothesized that I would find higher trash pollution abundance at Sungei Buloh Wetland Reserve/Field Site C, because it is the only field site directly next to a body of water. I also hypothesized that I would find a positive relationship between the number of plant and bird species as they may engage in beneficial relationships, as well as a negative relationship between trash pollution and the number of bird species as high density of trash pollution may deter birds from living in certain areas.

At Field Site A, I anticipated finding representative coastal hill forest plant species (*Tristaniopsis obovata*, *Syzygium grande*, *Ixonanthes reticulata*, *Rhodamnia cinerea*, *Terminalia catappa*, *Dipteris conjugata*) and more generalist bird species, as it is a secondary forest. At Field Site B, I expected to find primary/old secondary forest representative plant species (*Shorea curtisii*, *Dipterocarpus sublamellatus*) and some highly significant endemic bird species, as it more closely resembles Singapore's original rainforests. At Field Site C, I expected to find representative mangrove plant species (*Bruguiera gymnorhiza*, *Heritiera littoralis*, *Ceriops zippeliana*, *Cassine viburnifolia*, *Sonneratia alba*, *Xylocarpus granatum*) and migratory bird species that are not found in the other sites, as it represents a more specialized ecosystem.



## Materials and Methods

Labrador Nature Reserve was Field Site A and had two survey locations; Bukit Timah Nature Reserve was Field Site B and had two survey locations; Sungei Buloh Wetland Reserve was Field Site C and had one survey location. Each survey location was selected to be surveyed for bird and plant diversity and abundance of trash (table 1).



**Fig. 3:** Annotated map of Singapore from Google Earth, showing field sites

*Field Sites and Survey Locations*

**Table 1:** Field sites, survey locations within a field site, and descriptors.

<b>Field Site Code, Official Nature Reserve Name, Descriptors</b>	<b>Survey Location Code(s) and Descriptors</b>	
<p>A: Labrador Nature Reserve (1.2666° N, 103.8022 ° E)</p> <ul style="list-style-type: none"> <li>- Coastal hill forest</li> <li>- Young and old secondary forest</li> <li>- Southern Singapore</li> </ul>	A1	Along Nature Footpath, a worn stone footpath with extensive surrounding vegetation and further inside reserve
	A2	Along WWII Footpath, similar to A1 with but with a wide artificial clearing at 50m end
<p>B: Bukit Timah Nature Reserve (1.3486° N, 103.7767° E)</p> <ul style="list-style-type: none"> <li>- Lowland dipterocarp forest</li> <li>- Primary and old secondary forest</li> <li>- Central Singapore</li> </ul>	B1	Along Catchment Trail, sandy footpath near Macritchie Reservoir and close to Bukit Timah Expressway, tall surrounding vegetation
	B2	Along Cave Path, hilly and rocky footpath off the main path and with vegetation on cliffs
<p>C: Sungei Buloh Wetland Reserve (1.4466° N, 103.7234° E)</p> <ul style="list-style-type: none"> <li>- Wetlands, mangroves, mudflats</li> <li>- Northwestern Singapore</li> </ul>	C	Along Migratory Bird Trail, between the Main Hide and Main Bridge, between Sungei Buloh Besar and Buloh Tidal Ponds
	Only one survey location was used, as the Migratory Bird Trail is a circular footpath.	

Field Site A (Labrador Nature Reserve)



Survey Location A1



Survey Location A2

Field Site B (Bukit Timah Nature Reserve)



Survey location B1



Survey location B2

Field Site C (Sungei Buloh Wetland Reserve)



Survey location C

**Fig. 4:** Map of field sites and pictures from survey locations

**Survey Locations**

Field Site A (Labrador Nature Reserve)



(National Parks Board Singapore)

Field Site B (Bukit Timah Nature Reserve)



(National Parks Board Singapore)

Field Site C (Sungei Buloh Wetland Reserve)



(National Parks Board Singapore)

**Fig. 5:** Survey locations on official National Parks Board Singapore maps

### ***Sampling Methods***

At each survey location, a 40-60 meter transect was measured using a transect tape. Surveys were carried out from December 2023 to January 2024. The criteria for choosing survey locations was that they had to be along a pre-existing footpath to minimize disturbance to the environment by going off the path. The survey locations were selected by mapping out and numbering possible 40-60m transects (survey locations) throughout the field sites and choosing them using a random number generator.

**Visual plants survey (survey locations A1-2, B1-2, C):** a camera (0.5 lens on iPhone 13 Pro Max) was used to video record vegetation on either side of a transect while the researcher walked slowly along the transect, recording data for alternating 5m segments (e.g. 0-5m, 10-15m, etc.).

**Audio bird survey (survey locations A1-2, B1-2):** As birds are quite difficult to spot visually at these two sites, audio data was used. Along the same alternating 5m segments on the transects as stated earlier, birdsong was audio recorded (Voice Memo app on iPhone 13 Pro Max) as the researcher walked slowly along the transect.

**Visual bird survey (location C):** each bird that could be visually seen throughout the Migratory Bird Trail (along the outer edge of the mangrove and on the mudflats in the center) in two hours was videotaped and photographed (iPhone 13 Pro Max).

**Visual trash pollution survey (location C):** No trash pollution was spotted at survey locations A1-2 or B1-2 in any surveys, so this portion of the study only applies to survey location C. Each piece of trash that could be visually seen next to the footpath of the Migratory Bird Trail was documented by a photo (iPhone 13 Pro Max). Note that trash pollution was along the footpath, however birds (visual bird survey) were a further distance away.

Site C (Sungei Buloh Wetland Reserve) was unique from Site A (Labrador Nature Reserve) and Site B (Bukit Timah Nature Reserve) in that it consisted of a circular footpath surrounding mudflats with birds in the center, making a transect-based audio bird survey inappropriate. Therefore, C was surveyed for birds visually through photos and videos which were collected in a two-hour walk along the entire Migratory Bird Trail footpath (2152m). Segments of the survey transect were determined by using distances between pre-existing landmarks such as hides, shelters, and bridges. This method was also used to survey trash pollution at survey location C.

Surveys were conducted in a three-week span in December 2023 and January 2024. A preliminary survey was conducted at each survey location to confirm the survey methods, and then each location was surveyed according to Tabel 2. Surveys were spaced out by at least 24 hours and commenced at 9:00 AM each day.

Upon arrival and throughout surveying, any qualitative observations were noted about the environment, such as weather, sound pollution, wind, tide/water level, evidence of human interference, general ambience, etc.

### ***Species Identification***

Google Lens and the Singapore NParks Flora Fauna Web were used to identify and count the visual plant and bird species shown in the videos/photos. Audio recordings were then compared with online databases (eBird) of bird calls to identify the species and number of times each call was heard.

### ***Data Analysis***

Using Google Sheets, raw data tables were compiled based on the visual and audio data collected at the field sites/survey locations. Data was processed through Google Sheets functions: averages were calculated to have an average and standard deviation for each location per 5m, which allowed researchers to quantify species diversity and abundance and create data tables (see tables 3-7). These processed data tables then led to the creation of graphs (see figures 6-12).

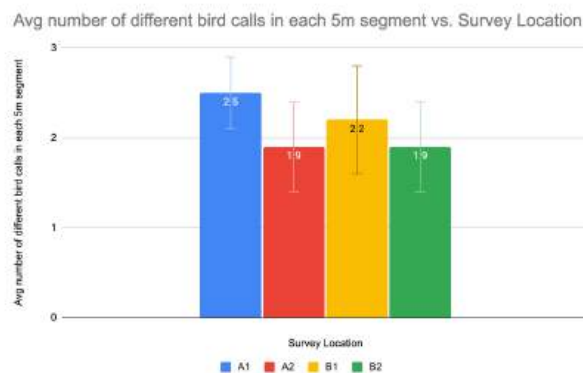
### **Results**

A total of five 40-60m transects and one full 2152m footpath was surveyed between December 2023 and January 2024.

### ***Audio Bird Survey Results***

Survey locations surveyed using the audio bird survey method (survey locations A1, A2, B1, B2) had an average number of different bird calls heard in a 5m segment ranging from 1.9-2.2 (s.d. 0.4-0.6) over the duration of surveying. Note that it is possible that during another season or time of day, results may be different.

The results on bird species diversity between field sites A and B suggest that field site A was overall more species diverse in terms of birds than field site B within the duration of surveying, and that the degree of difference in bird species diversity between 5m segments was greater in field site B than in field site A on the days where they were surveyed (figure 6).



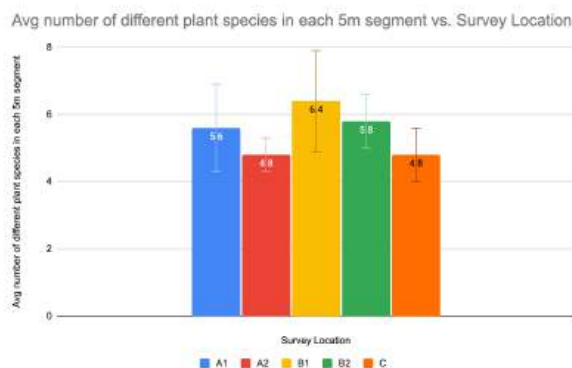
**Fig. 6:** Graph showing average number of different bird calls in each 5m segment vs survey location (A1 was 50 meters long; A2 was 40 meters long; B1 was 50 meters long; B2 was 50 meters long)

In this study, five different bird calls were heard at field site A (Labrador Nature Reserve), with 1-8 number of repeats per call, while seven different bird calls were heard at field site B, with 1-5 number of repeats per call. These different bird calls are therefore suggestive of different species.

### ***Visual Plants Survey Results***

Survey locations surveyed using the visual plants survey method (survey locations A1, A2, B1, B2, C), had an average number of different plant species in a 5m segment ranging from 4.8-6.4 (s.d. 0.5-1.5) over the duration of surveying. Note that it is possible that during another season, results may be different.

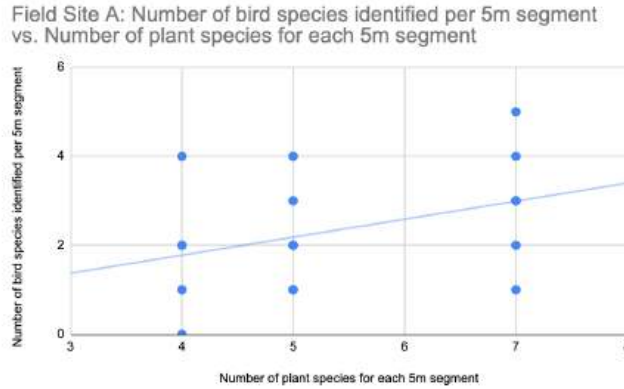
The results on plant species diversity between field sites A, B, and C suggest that field site B, followed by site sites A and C, was the most species diverse in terms of plants within the duration of surveying, and that the degree of difference in plant species diversity between 5m segments was greatest in field site B, followed by field sites A and C on the days where they were surveyed (figure 7).



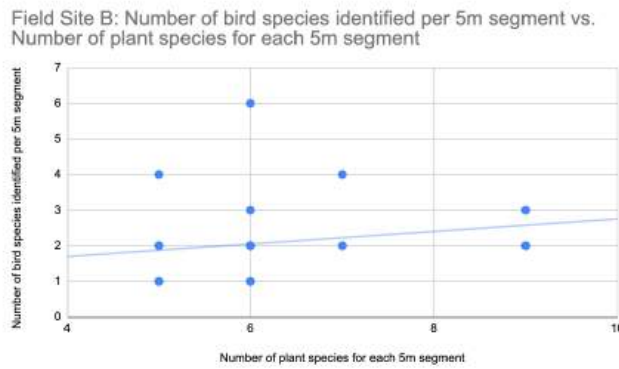
**Fig. 7:** Graph showing average number of different plant species in each 5m segment vs survey location (A1 was 50 meters long; A2 was 40 meters long; B1 was 50 meters long; B2 was 50 meters long; C was 60 meters long)

In this study, 6 plant species were found at field site A, 9 plant species were found at field site B, and 5 plant species were found at field site C.

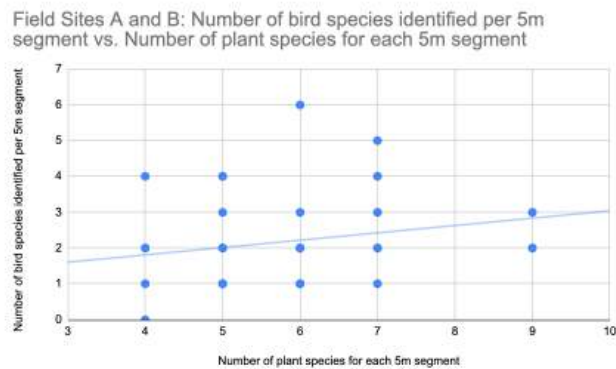
Scatter plots were used to compare and contrast relationships between the bird and plant data, as well as similarities and differences between the ecosystems (figures 8-10).



**Fig. 8:** Graph showing number of bird species identified per 5m segment vs number of plant species for each 5m segment (Field Site A)



**Fig. 9:** Graph showing number of bird species identified per 5m segment vs number of plant species for each 5m segment (Field Site B)



**Fig. 10:** Graph showing number of bird species identified per 5m segment vs number of plant species for each 5m segment (Field Sites A and B)



These graphs suggest that on the days the field sites were surveyed, there was a slight positive relationship between the number of plant species and the number of bird species identified (figures 8-10). This may suggest notable interactions and interdependence between plant and bird species.

**Visual Bird Survey Results**

At field site C, the number of different bird species per distance in each segment greatly varied, which may be correlated to the different environmental conditions of the various segments, as seen in the table below (table 2).

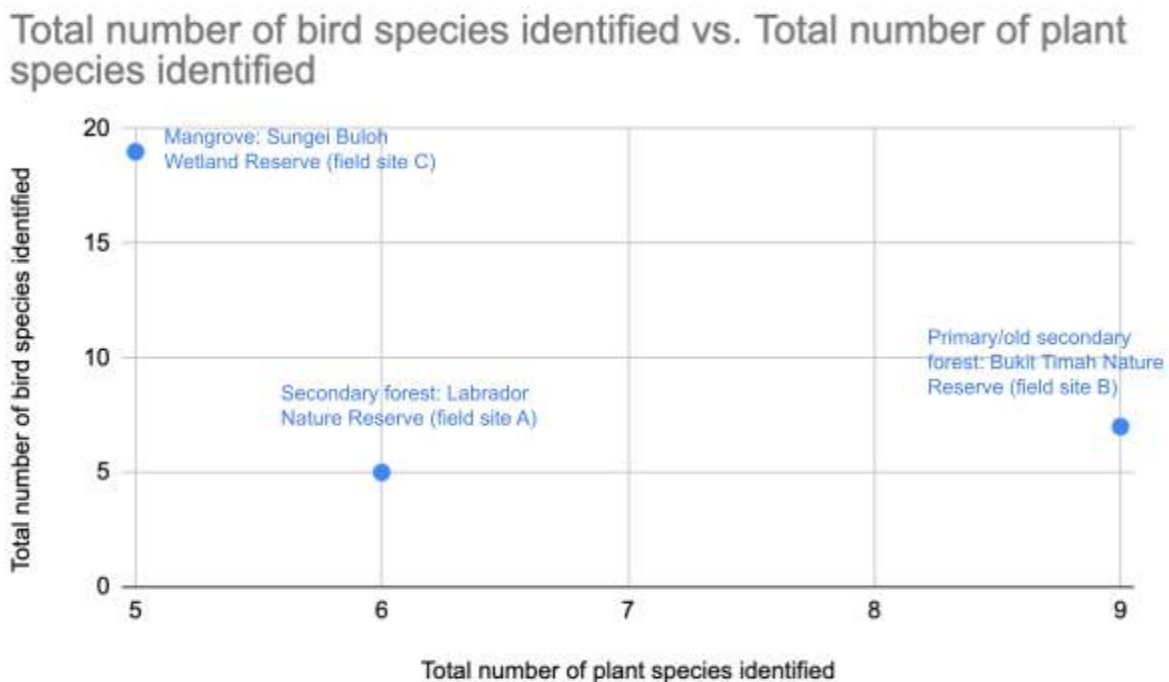
**Table 2:** Characteristics and bird species found at each segment in field site/location C

Segment along Migratory Bird Trail (m)	Segment characteristics	Number of different bird species				Avg number of different bird species per distance (n/m)	Cumulative avg number of different bird species per distance (n/m)
		Location C.1	Location C.2	Location C.3	Segment average		
0-207.27	bodies of water and trees on either side of footpath; mudflats visible on one side; water movement; quite wet	5	3	3	3.7	0.018	
207.27-549.17	bodies of water and trees on either side of footpath; mudflats visible on one side; water movement; quite wet	4	4	3	3.7	0.011	
549.17-676.38	bodies of water and trees on either side of footpath; mudflats slightly further away on one side; soil is drier, fewer mangrove trees	3	3	5	3.7	0.029	
676.38-1187.43	trees on either side of footpath; body of water on one side; mudflats barely visible and far away; soil is very dry, few mangrove trees	2	0	1	1.0	0.002	
1187.43-1480.02	bodies of water and trees on either side of footpath; mudflats slightly further away on one side; soil is drier, fewer mangrove trees and lower water level	1	2	1	1.3	0.005	
1480.02-1757.22	bodies of water and trees on either side of footpath; mudflats visible on one side; water movement; quite wet;	1	2	0	1.0	0.004	

	elevated footpath (higher water level)						
1757.22-2152.07	bodies of water and trees on either side of footpath; mudflats visible on one side; water movement; lower water level	0	1	1	0.7	0.002	

Over the three spaced visual bird surveys of field site/location C, a total of 19 bird species, both resident and migratory and including some artificially introduced species were identified. An average of 132 individuals and 12 different bird species were identified per survey, of which *Numenius phaeopus*, *Egretta garzetta*, *Tringa totanus*, and *Mycteria cinerea* occupy a large proportion of individual numbers. Bird species found also include notable migratory bird species *N. phaeopus*, *Ardea alba*, *T. totanus*, *Tringa stagnatilis*, *Tringa nebularia*, *Ardea intermedia*, and *Alcedo atthis*, which travel on the East Asian-Australasian Flyway (BirdLife International) and typically are present in Singapore during the winter months of October-February. Two resident introduced bird species, *M. cinerea* and *Mycteria leucocephala*, occupied a larger proportion of total individuals than some native species.

In order to find potential relationships between total number of plant species and total number of bird species, and compare this data between ecosystems, the researcher plotted a scatter graph (figure 11).



**Fig. 11:** Graph showing total number of bird species identified vs total number of plant species identified

From this graph, it can be seen that there may be a positive relationship between total number of plant species and total number of bird species through the Field Site A (6,5) and B (9,7) data points, however Field Site C differs substantially from the other two in that it has the lowest total number of plant species, but the highest total number of bird species (figure 11).

**Table 3:** Summary table of notable plant and bird species

Field Site	Notable species	Status
A (Labrador Nature Reserve)	<i>Gallus gallus</i>	Common
	<i>Malacocincla abbotti</i>	Representative of ecosystem type
	<i>Rhodamnia cinerea</i>	Representative of ecosystem type
	<i>Cinnamomum iners</i>	Common
	<i>Syzygium grande</i>	Representative of ecosystem type
B (Bukit Timah Nature Reserve)	<i>Pycnonotus zeylanicus</i>	Critically endangered
	<i>Irena puella</i>	Representative of ecosystem type
	<i>Dipterocarpus sublamellatus</i>	Endangered
	<i>Rhodamnia cinerea</i>	Common
C (Sungei Buloh Wetland Reserve)	<i>Numenius phaeopus</i>	Migratory/representative of ecosystem type
	<i>Ardea alba</i>	Migratory/representative of ecosystem type
	<i>Tringa totanus</i>	Migratory/representative of ecosystem type
	<i>Tringa stagnatilis</i>	Migratory/representative of ecosystem type

	<i>Tringa nebularia</i>	Migratory/representative of ecosystem type
	<i>Ardea intermedia</i>	
	<i>Bruguiera gymnorhiza</i>	Migratory/representative of ecosystem type
	<i>Rhizophora stylosa</i>	Migratory/representative of ecosystem type
		Representative of ecosystem type
		Representative of ecosystem type

### Visual Trash Pollution Survey Results

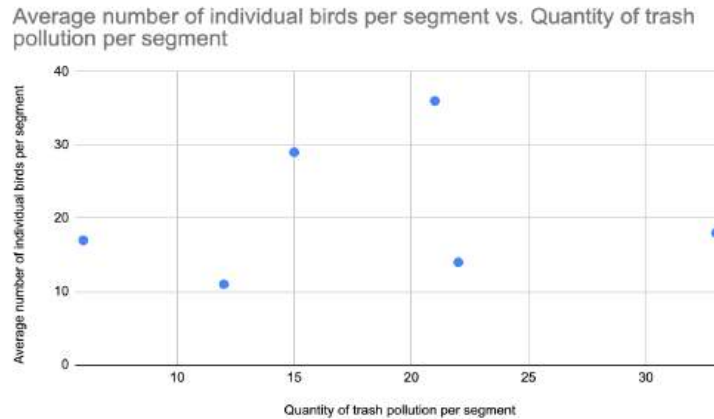
During the trash pollution survey, the researcher found 109 pieces of trash, including styrofoam boxes, bottle caps, plastic bags, and rubber items (table 4). On average, the researcher observed pieces of trash spread over the entirety of the outer-facing portion of the Migratory Bird Trail (2152m), with the largest quantity of trash found in segments 0-207.27m, 207.27-549.17m, and 1480.02-1757.22m, which were the segments that were closest to the edge of the mangrove and where water level was highest (table 2 for segment characteristics).

**Table 7:** Abundance of trash pollution types in Field site/location C

Type of Trash Pollution	Quantity	% occupied
Bottle caps	32	29.4%
Plastic Bottle	18	16.5%
Styrofoam	40	36.7%
Feed bags (polypropylene)	3	2.8%
Rubber items	4	3.7%
Plastic Bags	7	6.4%
Soft plastic wrappers (e.g. tissue packet, snacks)	5	4.6%

As birds present at the water-adjacent ecosystem at Field Site C may be in close contact with trash pollution while finding food, distribution of birds may be influenced by abundance of

trash pollution. To see if there was a relationship between abundance of trash pollution and abundance of individual birds, the researcher plotted a scatter graph (figure 12).



**Fig. 12:** Graph showing average number of individual birds per segment vs quantity of trash pollution per segment (Field Site C only)

The graph suggests that there is little relationship between trash pollution and abundance of birds at Field Site C (figure 12).

## Discussion

Over the duration of the research, a total of 31 bird species and 20 plant species were observed across the three ecosystems (primary/old secondary forest, coastal hill forest, mangrove), which display quite a high level of species diversity. Some of the bird and plant species observed also display ecological relationships with each other, such as mangrove roots (e.g. those of *Rhizophora stylosa*) providing shelter for fish that aquatic birds (e.g. *Ardea spp.*) require for nutrition.

This study also identified a number of introduced species and/or generalists across the three ecosystems, including *Gallus gallus* and *Acridotheres javanicus*. These species being spread across and rather abundant in the three surveyed ecosystems may lead to negative consequences, such as outcompeting native and more specialized species for habitat space and nutrients. Such bird and tree species tend to adapt more easily and with less resistance to ecosystems subject to human disturbance, suggesting that the ecosystems surveyed in this study have experienced notable human disturbance. Field Site B (Bukit Timah Nature Reserve) is notable in that it had the fewest number of introduced/generalist species and a higher abundance of endemic and IUCN “at risk” species out of the three field sites/ecosystems, which aligns with its categorization as a primary and old secondary rainforest.

Patterns of ecosystem-specific species were also identified, where the species that are known to be defining inhabitants of certain ecosystems and/or more threatened were only present at one field site/ecosystem. This indicates that their ecological niches are very specific, and that their specialization as opposed to the generalists requires more emphasis to be given to their

conservation and continued ecological study. These findings also support the researcher's argument that the continued protection of remaining ecosystems in Singapore is necessary in order to prevent further habitat loss for highly specialized bird and plant species.

My results of primary forest at Bukit Timah Nature Reserve being largely dominated by *Dipterocarpaceae* and *Ixonanthaceae* also corroborated Corlett's study, which found plant species from the above groups in higher proportions than other trees. Rattans (genera *Calamus*, *Daemonorops*, *Korthalsia*, *Plectocomia*) were also abundant across the primary/old secondary forest I surveyed, which was also a result of the surveying done by Corlett. However, there are very few academic studies regarding species surveys of the other two field sites/ecosystems (coastal hill and mangrove). Despite this, the results of this study have overlaps with the species found in official National Parks Board Singapore species guides for the ecosystems. This suggests that the results of this survey are in agreement with those of existing scholarly research, which serves to bolster their arguments that the ecosystems of Singapore, despite fragmentation, remain biodiverse and unique from one another.

Most of my results are in line with my hypotheses: I found greater diversity of plant and bird species at Bukit Timah Nature Reserve/Field Site B (primary/old secondary forest) than at Labrador Nature Reserve/Field Site A (secondary forest). There is evidence in the literature that this is a pattern— in their study, Turner et al. found that secondary forest contained significantly fewer tree species than primary forest (Turner et al. 537). Castelletta et al. also found that 94% of birds recently extinct in Singapore were reliant on primary and old secondary forest, indicating the higher biodiversity in primary/old secondary forest than secondary forest, noting the example of bird species *Dryocopus javensis*, *Chloropsis cyanopogon* and *C. sonnerati* being likely to go extinct due to habitat loss in upper and middle rainforest layers (Castelletta et al. 1870, 1876-1877).

A positive relationship was observed between the number of different plant species and number of different bird species identified in field sites A and B. This suggests that species richness of plants and birds influence one another, and that certain bird species may be highly dependent on certain plant species for survival and vice versa. Existing literature has also found a similar positive relationship between plant species richness and animal species richness (Castagneyrol and Jactel 2119). Such findings indicate beneficial relationships and ecological interactions between local plant and bird life in Singaporean ecosystems, which may inform future conservation efforts— propagating plants which directly benefit endemic birds and other animals may therefore be an effective method of conservation and restoring local ecosystems. Species richness can therefore be used to inform conservation efforts by functioning as a partial measurement of biodiversity (Castagneyrol and Jactel 2123).

Trash pollution included both industrial and domestic waste, which may have washed into the mangroves as a result of the tides from nearby aquaculture businesses, industrial warehouses, and residential areas. A notable incident occurred where researchers spotted a group of crows using a styrofoam takeaway container to build a nest, indicating noticeable interactions between organisms and trash pollution.

Interestingly, the graph showed little relationship between trash pollution and abundance of birds at Field Site C/Sungei Buloh Wetland Reserve (figure 12), which may indicate that birds at this site have not yet changed their distribution to avoid trash pollution. However, it should be noted that due to ocean currents and tides, the distribution of trash pollution is highly variable (Curren and Leong 4). The vast majority of birds found across survey sessions at this site were also located inside the oval footpath on the mudflats, whereas trash pollution remained on the outskirts of the footpath due to the protective area of mangrove trees. Therefore, the results did not suggest that trash pollution interfered significantly with plant and bird interactions, however more long-term research across multiple ecosystems with trash pollution is necessary to further understand this. It is definitely important to pay attention to and systematically study trash pollution, as it continues to accumulate and impact many ecosystems and organisms, particularly aquatic ones.

However, it is worth noting that as these ecosystems were surveyed using an observational method and only within a three-week period, results may differ from other times of the year or using different survey methods (mark-capture-release-recapture, wildlife camera traps, etc.). Future research could consider surveying ecosystems year-round to capture any possible shifts in species composition and trash pollution, as well as obtaining permits to engage in non-observational research as these methods may provide more accurate understanding of these ecosystems. The method used in field site C (Sungei Buloh Wetland Reserve) also differed from sites A and B due to constraints of the nature reserve, but the researcher surveyed it nonetheless as it is an ecologically significant site. If possible, further research should aim to standardize methodology across all sites/locations. This would improve comparability of data sets between the ecosystems, allowing researchers to more clearly investigate the similarities and differences between birds, plants, and trash pollution across distinct ecosystems.

Finally, this study and existing literature highlight the value of nature reserves, as they are ways to preserve remaining ecosystems and prevent them from experiencing extensive human disturbance. This is particularly evident in Singapore, where ecosystems have become heavily fragmented due to urbanization. It is therefore important to continually monitor these ecological areas of concern for issues such as invasive species and trash pollution, making a case for the continuation of nature reserves and conserving remaining biodiversity.

## Conclusions

Overall, during this study it was found that some generalist species existed across multiple ecosystems (*Rhodamnia cinerea*, *Gallus gallus*, *Cinnamomum iners*) and were relatively noticeable throughout the duration of surveying, whereas some specialist species were only observed at one ecosystem/field site (*Pycnonotus zeylanicus*, *Syzygium grande*, *Dipterocarpus sublamellatus*, *Bruguiera gymnorhiza*, *Rhizophora stylosa*, *Tringa spp.*). However, most species were not observed to inhabit multiple field sites/ecosystems during the period of surveying, suggesting that differences in abiotic and biotic factors between the three

ecosystems surveyed (primary forest, secondary forest, mangrove) have led to largely distinct plant and bird species compositions.

The ecosystems of Singapore still deserve continued scientific study of their species composition and how humans have impacted them, as they remain unique and specialized from each other. Singapore's ecosystems have become increasingly fragmented due to urbanization and continued expansion of the city-state's residential and industrial areas, which puts its variety of ecosystems under increased stress. A number of ecologically significant and globally threatened species also inhabit local ecosystems, such as *Pycnonotus zeylanicus* (critically endangered), as a specialized inhabitant of primary rainforests; and mangrove tree species, which provide invaluable coastal protection and have been restricted to the very Northwest corner of Singapore.

Understanding these ecosystems and their species would inform increased conservation efforts to protect Singapore's remaining ecosystems. This is integral towards supporting not only Singapore's native ecology, but also human society, as ecology involves both natural and human factors; ecosystems provide the foundations for the continued survival of life on Earth, and continued ecological research is needed in order to better protect them and promote their survival. Southeast Asia's ecosystems as a whole also remain biodiversity hotspots and some of the most productive biomes on Earth, but also one of the most highly threatened— the rapid urbanization of many of these countries may place more ecosystems at risk, and increased ecological surveys would be a crucial step towards communicating their importance and actioning conservation.

In order to monitor the health of remaining ecosystems in Singapore, multiple data sets must be collected across longer time frames and cover a number of different ecosystems. This would enable future research to understand not only how various species interact with each other and how they have been impacted by trash pollution, but also provide valuable insights into how changes that occur over longer periods of time (e.g. seasons) influence these relationships.



## Works Cited

- Ashton, P.S. “New *Tristaniopsis* Peter G. Wilson & J.T. Waterh. (Myrtaceae) From Borneo.” *Gardens' Bulletin Singapore*, vol. 57, no. 1, 2005, pp. 269-278, nparks.org.sg. Accessed 3 May 2024.
- BirdLife International. “EAST ASIAN-AUSTRALASIAN FLYWAY.” *BirdLife International*, 7 September 2023, <https://www.birdlife.org/news/flymap/east-asian-australasian-flyway/>. Accessed 3 May 2024.
- Castagneyrol, Bastien, and Hervé Jactel. “Unraveling plant—animal diversity relationships: a meta-regression analysis.” *Ecology*, vol. 93, no. 9, 2012, pp. 2115-2124. *JSTOR*, [https://www.jstor.org/stable/pdf/41739269.pdf?casa\\_token=MicpcgfptCQAAAAA:TENxqslRt\\_7HzdPcheLDJGucw5llmTDKQ31xsQGikcS7zUfOXExzqZqay1\\_6uWiJM8FVmcnulZV5NMDsSZJyOF1gkOdAithUQyqj0NZ7Sy\\_w\\_oVFGSPoIg](https://www.jstor.org/stable/pdf/41739269.pdf?casa_token=MicpcgfptCQAAAAA:TENxqslRt_7HzdPcheLDJGucw5llmTDKQ31xsQGikcS7zUfOXExzqZqay1_6uWiJM8FVmcnulZV5NMDsSZJyOF1gkOdAithUQyqj0NZ7Sy_w_oVFGSPoIg). Accessed 17 July 2024.
- Castelletta, Marjorie, et al. “Heavy Extinctions of Forest Avifauna in Singapore: Lessons for Biodiversity Conservation in Southeast Asia.” *Conservation Biology*, vol. 14, no. 6, 2000, pp. 1870-1880. *JSTOR*, [https://www.jstor.org/stable/pdf/2641539.pdf?casa\\_token=6q2fhEF4oQkAAAAA:VRlaw62vDeLiEZfw5lh4TW0YoBMkZNMcfX6MRpAOgxZcm8cqkbAUV3YMnofaKAYlQuOmysNhHiy7P8-AnBxrKAPGN9Zh7BT5RCv5\\_\\_zpdLKSz5HzadWiSg](https://www.jstor.org/stable/pdf/2641539.pdf?casa_token=6q2fhEF4oQkAAAAA:VRlaw62vDeLiEZfw5lh4TW0YoBMkZNMcfX6MRpAOgxZcm8cqkbAUV3YMnofaKAYlQuOmysNhHiy7P8-AnBxrKAPGN9Zh7BT5RCv5__zpdLKSz5HzadWiSg). Accessed 17 July 2024.
- Corlett, Richard T. “The Ecological Transformation of Singapore, 1819-1990.” *Journal of Biogeography*, vol. 19, no. 4, 1992, pp. 411-420. *jstor.org*, [https://www.jstor.org/stable/pdf/2845569.pdf?casa\\_token=V0AgvEPf3PQAAAAA:bW9s4BTU3TJev4sCv43q4beb1JaXBRSJ4vJvDbiKY7kl9QLUgl2aoLSU7Y2XGagktw8\\_fucXiOSvgkJViaHntKxs3WDHSeTGJ1RdtUA5mKF\\_5zi1hNw](https://www.jstor.org/stable/pdf/2845569.pdf?casa_token=V0AgvEPf3PQAAAAA:bW9s4BTU3TJev4sCv43q4beb1JaXBRSJ4vJvDbiKY7kl9QLUgl2aoLSU7Y2XGagktw8_fucXiOSvgkJViaHntKxs3WDHSeTGJ1RdtUA5mKF_5zi1hNw).
- Curren, Emily, and Sandric Chee Yew Leong. “Spatiotemporal characterisation of microplastics in the coastal regions of Singapore.” *Helicon*, vol. 9, no. e12961, 2023, pp. 1-12. *CellPress*, [https://www.cell.com/helicon/pdf/S2405-8440\(23\)00168-8.pdf](https://www.cell.com/helicon/pdf/S2405-8440(23)00168-8.pdf). Accessed 17 July 2024.
- eBird. “eBird - Discover a new world of birding.” *eBird*, n.d., <https://ebird.org/home>. Accessed 3 May 2024.
- Goh, Naomi. “Bukit Timah Nature Reserve: A Model for Conservation.” *National Parks Board (NParks)*, 2019, <https://www.nparks.gov.sg/nparksbuzz/issue-41-vol-2-2019/conservation/bukit-timah-nature-reserve>. Accessed 3 May 2024.
- Gondal, A.H., et al. “Advances in plastic pollution prevention and their fragile effects on soil, water, and air continuums.” *International Journal of Environmental Science and Technology*, vol. 20, 2023, pp. 6897–6912. *Springer*, <https://doi.org/10.1007/s13762-022-04607-9>. Accessed 17 July 2024.

- Ho, B-C, et al. "The plant diversity in Bukit Timah Nature Reserve, Singapore." *Gardens' Bulletin Singapore*, vol. 71, no. Supplement 1, 2019, pp. 1-95. *ResearchGate*, [https://www.researchgate.net/publication/333399862\\_The\\_plant\\_diversity\\_in\\_Bukit\\_Timah\\_Nature\\_Reserve\\_Singapore](https://www.researchgate.net/publication/333399862_The_plant_diversity_in_Bukit_Timah_Nature_Reserve_Singapore). Accessed 4 April 2024.
- IUCN. "The IUCN Red List of Threatened Species." *IUCN Red List of Threatened Species*, 2024, <https://www.iucnredlist.org/>. Accessed 20 July 2024.
- Lim, Haw Chuan, et al. "Undesirable aliens: factors determining the distribution of three invasive bird species in Singapore." *Journal of Tropical Ecology*, vol. 19, 2003, pp. 685-695. *JSTOR*, [https://www.jstor.org/stable/pdf/4091842.pdf?casa\\_token=IW3EVroscPcAAAAA:hzK9A7b\\_LZaTpQwi0MfUVrXMSIRehdsicLdb6GZQQOi6lRreO10x8W2JXbXPA0PnTg9NCc33umDNJHw3fSPYF4CAfFUDb4F\\_sKz8qndvo4ci1SQOpB8LCQ](https://www.jstor.org/stable/pdf/4091842.pdf?casa_token=IW3EVroscPcAAAAA:hzK9A7b_LZaTpQwi0MfUVrXMSIRehdsicLdb6GZQQOi6lRreO10x8W2JXbXPA0PnTg9NCc33umDNJHw3fSPYF4CAfFUDb4F_sKz8qndvo4ci1SQOpB8LCQ). Accessed 17 July 2024.
- Lim, K.S. "Birds of Bukit Timah Nature Reserve, Singapore." *Gardens' Bulletin Singapore*, vol. 71, no. Supplement 1, 2019, pp. 185-208.
- Low, Jeffrey, and Liang Jim Lim. "Conservation of our natural heritage: The Singapore experience." *Conference: 12th International Coral Reef Symposium*, vol. 22A, no. 1, 2012, pp. 1-4. *ResearchGate*, [https://www.researchgate.net/publication/262364395\\_Conservation\\_of\\_our\\_natural\\_heritage\\_The\\_Singapore\\_experience](https://www.researchgate.net/publication/262364395_Conservation_of_our_natural_heritage_The_Singapore_experience). Accessed 3 May 2024.
- Nai, Jacqueline. "Spot these Seven Migratory Birds at Sungei Buloh Wetland Reserve!" *National Parks Board (NParks)*, 2016, <https://www.nparks.gov.sg/nparksbuzz/issue-30-vol-3-2016/conservation/spot-these-seven-migratory-birds-at-sungei-buloh-wetland-reserve>. Accessed 3 May 2024.
- National Parks Board Singapore. "Bukit Timah Nature Reserve." *National Parks Board (NParks)*, n.d., <https://beta.nparks.gov.sg/visit/parks/park-detail/bukit-timah-nature-reserve>. Accessed 4 April 2024.
- National Parks Board Singapore. "Labrador Nature Reserve - Singapore." *National Parks Board (NParks)*, n.d., <https://beta.nparks.gov.sg/visit/parks/park-detail/labrador-nature-reserve>. Accessed 4 April 2024.
- National Parks Board Singapore. "SIXTH NATIONAL REPORT TO THE CONVENTION ON BIOLOGICAL DIVERSITY (2015-2018)." *Convention on Biological Diversity*, vol. 16, no. 6, 2020, pp. 1-135. *National Parks Board Singapore*, <https://www.nparks.gov.sg/-/media/nparks-real-content/biodiversity/international-regional-links/singapore---6th-national-report-to-the-cbd.ashx>. Accessed 2 April 2024.
- National Parks Board Singapore. "Sungei Buloh Wetland Reserve." *National Parks Board (NParks)*, n.d., <https://beta.nparks.gov.sg/visit/parks/park-detail/sungei-buloh-wetland-reserve>. Accessed 4 April 2024.

- Ng, Charlyn, and Felix Siew. “Haven for History and Natural Heritage.” *National Parks Board (NParks)*, 2022, <https://www.nparks.gov.sg/nparksbuzz/jan-issue-2022/main-feature/haven-for-history-and-natural-heritage>. Accessed 3 May 2024.
- Noreen, Annika M E, and Edward L. Webb. “High Genetic Diversity in a Potentially Vulnerable Tropical Tree Species Despite Extreme Habitat Loss.” *PLOS ONE*, vol. 8, no. 12, 2013, pp. 1-10. *researchgate.net*, 10.1371/journal.pone.0082632.
- Peh, Kelvin S.H. “Invasive species in Southeast Asia: the knowledge so far.” *Biodiversity and Conservation*, vol. 19, 2010, pp. 1083-1099. *ResearchGate*, [https://www.researchgate.net/profile/Kelvin-Peh-2/publication/224858097\\_Invasive\\_species\\_in\\_Southeast\\_Asia\\_The\\_knowledge\\_so\\_far/links/09e414fc6b4ad8364d000000/Invasive-species-in-Southeast-Asia-The-knowledge-so-far.pdf](https://www.researchgate.net/profile/Kelvin-Peh-2/publication/224858097_Invasive_species_in_Southeast_Asia_The_knowledge_so_far/links/09e414fc6b4ad8364d000000/Invasive-species-in-Southeast-Asia-The-knowledge-so-far.pdf). Accessed 17 July 2024.
- Tan, Benito C., and Koh-Siang Tan. “Invasive Alien Species in Singapore: A Review.” *Prevention and Management of Invasive Alien Species: Proceedings of a Workshop on Forging Cooperation throughout South and Southeast Asia*, The Global Invasive Species Programme, 2002, pp. 58–64, [www.gisp.org/downloadpubs/SOUTH\\_A2.PDF#page=58](http://www.gisp.org/downloadpubs/SOUTH_A2.PDF#page=58). Accessed 17 July 2024.
- Turner, I.M., et al. “Tree species richness in primary and old secondary tropical forest in Singapore.” *Biodiversity and Conservation*, vol. 6, 1997, pp. 537–543. *Springer*, <https://doi.org/10.1023/A:1018381111842>. Accessed 17 July 2024.
- Yam, Tim W., et al. “Conservation of the Native Orchids Through Seedling Culture and Reintroduction—A Singapore Experience.” *Selbyana*, vol. 26, no. 1/2, 2010, pp. 263-274. *jstor.org*, <https://www.jstor.org/stable/40792838>. Accessed 3 May 2024.
- Yee, A.T.K., et al. “The vegetation of Singapore—an updated map.” *Gardens' Bulletin Singapore*, vol. 63, no. 1, 2011, pp. 205-212. *researchgate.net*, [https://www.researchgate.net/publication/280601159\\_The\\_vegetation\\_of\\_Singapore\\_-\\_An\\_updated\\_map](https://www.researchgate.net/publication/280601159_The_vegetation_of_Singapore_-_An_updated_map).

# Lightweight Cryptography as A Tool to Meet Singapore's Carbon Emission Targets

## By Ananya Kharbanda

### Abstract

Cryptographic practices play a vital role in securing digital communication and financial transactions, but their energy-intensive nature poses significant environmental challenges. The increasing energy demands of cryptographic algorithms and protocols have created a growing tension between robust security measures and environmental sustainability, particularly in the context of massive data centers and emerging technologies like quantum computing and the Internet of Things (IoT). Singapore is a hub of business and trading, and therefore requires robust cryptographic systems. This paper explores various mitigation efforts, including the development of energy-efficient cryptographic algorithms, hardware acceleration, server virtualization, and the integration of renewable energy sources in data centers. It also discusses Singapore's Smart Nation initiative, focusing on the country's efforts to integrate renewable energy into its digital infrastructure, enhancing sustainability while maintaining financial and technological competitiveness. In conclusion, the study emphasizes the need for a multifaceted approach that balances security, efficiency, and sustainability in cryptographic practices, addressing pertinent environmental concerns while also meeting the needs of national security and economic growth.

### Introduction

The growth (McCurley) of the internet and the increasing reliance on secure communication protocols have led to a rise in the energy consumption associated with encryption processes. The issue of energy consumption for encryption has only grown ever since it was first brought up in the early 2000s. As the tech industry begins to take its first steps to reduce its ecological impact, it becomes prudent to explore the evolution of encryption methods and their implications on climate change.

In the early stages of digital communication, encryption was primarily employed by governments and large organizations to secure sensitive information (Berret). Traditional cryptographic techniques, such as symmetric-key cryptography, were computationally less intensive, and the environmental impact was relatively modest. However, with the advent of the internet and the widespread adoption of digital communication in the late 20th century, the demand for more robust encryption to protect sensitive personal information has grown exponentially (Abdalla et al.). Common types of encrypted data that people may encounter include encrypted messages in communication apps like WhatsApp, encrypted files on cloud storage platforms such as Google Drive, and encrypted payment information during online transactions using services like PayPal or credit card portals.

One of the major changes in commonly used encryption is the shift towards public-key cryptography, particularly with the introduction of the RSA algorithm (Milanov). Public-key cryptography, while providing enhanced security, has introduced higher computational

requirements. This led to an increase in the processing power and energy consumption associated with encryption practices, as complex mathematical operations became integral to securing digital communication. As internet (Wendl et al.) usage expanded globally, so did the implementation of secure communication protocols like SSL/TLS for web browsing and email (Das and Samdaria). These protocols rely on cryptographic algorithms to ensure the confidentiality and integrity of data in transit. The adoption of encryption became widespread, driven not only by security concerns but also by privacy considerations.

One of the latest large-scale developments in the internet is the rise of cloud computing. Cloud services, which host vast amounts of data for individuals and organizations, rely on encryption to protect sensitive information. The constant need for data encryption and decryption in the cloud contributes to the overall energy consumption of data centers.

While encryption is crucial for safeguarding sensitive information, it requires energy-intensive computations, contributing to the carbon footprint associated with digital technologies. This environmental impact is a growing concern as questions about climate change and sustainability have become more prominent globally (Abbass et al.). Efforts have been made to address the environmental implications of encryption practices. Researchers and technologists are exploring more energy-efficient cryptographic algorithms, optimizing existing protocols, and promoting the use of renewable energy sources in data centers (Manganelli et al.). Striking a balance between maintaining robust security measures and mitigating the environmental impact of encryption remains an ongoing challenge as society continues to grapple with the implications of digitalization on climate change.

Currently, the majority of data centers powering these computations rely on traditional energy sources, further exacerbating environmental concerns. This trend contradicts the objectives outlined in the Paris Agreement, which aims to reduce greenhouse gas emissions and limit climate change (UNFCCC). The growing energy consumption from cryptographic practices highlights the need for sustainable approaches within digital technologies. Efforts to mitigate this impact have focused on the optimization of data center operations to minimize energy usage, and the promotion of renewable energy sources like solar and wind power to power data centers. However, the development of more energy-efficient cryptographic algorithms and protocols must also be a part of these efforts. Addressing the environmental implications of cryptographic practices requires collaboration among various stakeholders, including researchers, industry leaders, and policymakers. Finding a balance between maintaining robust security measures and reducing the environmental footprint of encryption processes is crucial to aligning with the goals of the Paris Agreement and promoting sustainability in digital infrastructures.

This paper takes a deeper look at the need for more efficient and less computationally intensive cryptography for the world. Beginning with the evolution of more-intensive cryptographic standards, it will then look at trends driving the proliferation of these standards and their consequences for environmental sustainability. It will conclude with a look at some efforts to mitigate the impact of highly intensive cryptographic standards, and what the

implications, opportunities and challenges these hold for the world in general and Singapore in particular.

### **Evolution of Cryptography (Less-Intensive to More-Intensive)**

In the early stages of digital communication, cryptographic techniques were primarily focused on securing sensitive information using symmetric-key cryptography. This approach involves using a single shared key for both encryption and decryption, making it relatively less computationally intensive (Sharma et al.).

Symmetric-key cryptography, while effective for securing data, had limitations in terms of key management and secure key exchange over insecure channels. As digital communication expanded and the internet became more pervasive, the need for stronger encryption mechanisms became apparent. This led to the development and adoption of asymmetric encryption algorithms, such as the RSA algorithm and Elliptic Curve Cryptography (ECC) (Mahto and Yadav).

Public-key cryptography revolutionized the field by introducing the concept of key pairs – a public key for encryption and a private key for decryption (Yilek et al.). This approach addressed the key management challenges of symmetric-key cryptography and enabled secure key exchange protocols like Diffie-Hellman key exchange (Li).

The RSA algorithm, named after its inventors Ron Rivest, Adi Shamir, and Leonard Adleman, relies on the computational complexity of factoring large prime numbers (Thabit et al.). While RSA provided enhanced security, it also introduced higher computational requirements due to its reliance on modular exponentiation operations. This shift towards asymmetric encryption marked a milestone in the evolution of cryptography but also contributed to increased energy consumption in encryption processes.

Similarly, Elliptic Curve Cryptography (ECC) gained popularity for its ability to provide strong security with relatively smaller key sizes compared to RSA (Ullah et al.). ECC relies on mathematical properties of elliptic curves to achieve secure encryption and digital signatures. However, ECC also requires computational-intensive operations, especially in key generation and elliptic curve scalar multiplication.

As digital communication continued to evolve, cryptographic techniques evolved further to address emerging security challenges. Advanced encryption standards like AES (Advanced Encryption Standard) replaced older symmetric-key algorithms like DES (Data Encryption Standard) and Triple DES due to their stronger security properties and efficiency in hardware and software implementations (Sood and Kaur).

The transition from less-intensive symmetric-key cryptography to more-intensive asymmetric encryption and modern encryption standards reflects the growing complexity of security requirements in the digital age. While these advancements provide enhanced security, they also contribute to the overall energy consumption of cryptographic processes. However, asymmetric cryptography requires significantly greater computation than symmetric-key cryptography.

Consider a 1000 Unicode character email, approximately 2000 bytes, involves distinct computational processes and efficiencies depending on the cryptographic method employed. Encrypting it via symmetric key cryptography, which uses the same key for both encryption and decryption, is generally efficient and operates linearly with the length of the plaintext. This method involves generating a single key, encrypting each block of plaintext, and combining the encrypted blocks to form the ciphertext. AES, a widely used symmetric algorithm, supports key sizes of 128, 192, or 256 bits and encrypts data in 128-bit blocks through a series of transformation rounds. These transformations, optimized for performance, are accelerated by modern CPU instructions like AES-NI, making AES particularly suitable for encrypting large amounts of data efficiently.

In contrast, RSA, an asymmetric cryptographic algorithm, uses a pair of public and private keys generated from two large prime numbers. While RSA is highly secure, it is computationally intensive due to the resource-heavy process of modular exponentiation. This makes RSA less ideal for bulk data encryption. For a 1000-character email, using RSA to encrypt each chunk of plaintext would be inefficient compared to AES. Therefore, a practical approach involves using RSA to securely exchange an AES key, which then encrypts the email content. This combination leverages the strengths of both algorithms: the efficiency of AES for data encryption and the security of RSA for key management.

Research has explored the computational requirements and energy consumption of various cryptographic techniques, highlighting the trade-offs between security and energy efficiency (Thabit et al.). These studies underscore the importance of balancing security needs with energy-conscious cryptographic implementations to mitigate the environmental impact of encryption practices.

### **Increased Adoption of Cryptographic Techniques**

The increased adoption of cryptographic techniques across digital applications has been fueled by the exponential increase in digital communication and the need to protect sensitive information from unauthorized access both on specific devices and in transit (Ullah et al.; Ketha). Online communication platforms, such as messaging apps, email services, and social media portals handle vast amounts of sensitive information, including personal conversations, financial transactions, and confidential documents (Naeem). Encryption ensures that this data remains secure and confidential, even if intercepted by malicious actors.

In fact, messaging apps like WhatsApp and Signal and email services like ProtonMail tout the privacy and security of their end-to-end encryption as a differentiator from their competitors (Rösler et al.; Kobeissi)

Cloud computing has also played a significant role in the increased adoption of encryption. With the shift towards cloud-based storage and computing services, organizations and individuals rely on encryption to protect data stored in the cloud. Encrypted files on cloud storage platforms, such as Google Drive and Dropbox, ensure that data remains confidential and inaccessible to unauthorized users (Daryabar et al.).

Furthermore, the rise of e-commerce and online payment systems has driven the adoption of encryption in financial transactions. Services like PayPal, credit card portals, and cryptocurrency exchanges use encryption protocols to secure payment information and prevent fraudulent activities (Adewole et al.).

Researchers have found a significant correlation between the volume of encrypted data and the energy-intensive computations required for encryption and decryption (Zhou et al.). In the context of Singapore, a vast amount of data flows into and out of the country due to its status as a major global financial hub and data center destination; Singapore is home to 100 data centers, 1,195 cloud service providers, and 22 network fabrics (Interesse).

A typical server farm processing similar volumes of data would employ a combination of symmetric key encryption for bulk data and asymmetric encryption for key exchanges. Advanced Encryption Standard (AES) can encrypt data at a rate of approximately 1 gigabyte per second using hardware acceleration (AES-NI) on modern processors (Intel, 2021). Encrypting 1 petabyte (1 million gigabytes) of data, therefore, would require around 1 million processor seconds, or approximately 278 processor hours, assuming continuous operation.

The energy consumption of such operations is considerable. A study by the U.S. Department of Energy's Lawrence Berkeley National Laboratory reported that data centers in the United States consumed about 70 billion kWh in 2014, with significant portions attributed to computational tasks, including encryption. Given Singapore's efforts towards energy efficiency, with data centers consuming an estimated 7% of the country's total electricity in 2020, the energy demands for encrypting large data volumes remain substantial.

Comparatively, large-scale data centers like Google's, which manage extensive data volumes, consumed about 12.4 terawatt-hours (12.4 billion kWh) of energy globally in 2020 (Google, 2021). The energy demands for cryptographic operations highlight the need for optimization in encryption technologies and infrastructure to mitigate the substantial energy impacts. This is particularly crucial in data-heavy environments like Singapore, where data centers are pivotal to the digital economy.

These examples underscore the importance of continuing innovation in cryptographic processes and data center efficiency to manage the energy implications of securing large-scale data flows.

As the demand for encryption continues to grow with increased digitalization, finding energy-efficient cryptographic solutions becomes imperative to mitigate their environmental impact (Kohli et al.).

### **Tension Between Security and Sustainability**

Cryptographic practices face a fundamental tension between maintaining robust security measures and minimizing environmental impact. Encryption is crucial for safeguarding sensitive information, but the algorithms and protocols that enable strong encryption often require significant computational resources, leading to increased energy consumption and carbon emissions. This challenge becomes even more pronounced with the advent of quantum



computing, which threatens to break long-established cryptographic methods, necessitating a transition to quantum-resistant systems. However, this transition is complex and requires careful consideration of both security and sustainability.

The rise of the Internet of Things (IoT) amplifies this tension, as the demand for lightweight cryptography that can withstand quantum computing attacks increases. This situation raises critical questions about how users will adapt to quantum technology and its impact on mobile internet security, a key aspect of modern cybersecurity.

To address these challenges, the deployment of post-quantum cryptographic techniques and privacy-enhancing technologies is essential. These solutions are crucial for future security, especially in light of evolving regulations and the growing data collection by IoT devices.

In conclusion, the tension between security and sustainability in cryptographic practices requires a comprehensive approach that balances the need for robust security with the environmental impacts of encryption, while also preparing for the challenges posed by quantum computing.

### **Mitigation efforts**

Cryptographic operations are fundamental to securing digital communication and transactions, yet their energy-intensive nature raises environmental concerns. Mitigation efforts encompass various strategies aimed at reducing the environmental footprint while maintaining security and operational efficiency (Abdullah and Lim).

Efforts to develop energy-efficient cryptographic algorithms are underway, focusing on reducing computational requirements without compromising security. The NIST Lightweight Cryptography Standardization Process has led to the adoption of algorithms like Simon and Speck, designed for resource-constrained environments such as IoT devices (Gookyi et al.; Hou et al.). These algorithms prioritize minimal energy consumption while maintaining robust security measures, making them suitable for energy-efficient cryptographic operations.

Algorithmic optimizations play a crucial role in enhancing energy efficiency. Techniques such as algorithmic pruning, where redundant computations are minimized, and parallel execution of cryptographic tasks on multicore processors, contribute to reduced energy consumption. Additionally, hardware acceleration, such as specialized cryptographic coprocessors, further improves energy efficiency by offloading cryptographic computations from general-purpose processors (Carreira-Perpinán and Idelbayev; Nagendra and Sekhar; Busby et al.). Singapore is embracing the use of energy-efficient cryptographic algorithms to minimize computational overhead while maintaining robust security (Sharon). This is particularly crucial for the financial sector, where high-frequency and algorithmic trading demand substantial computational resources. Optimizing these algorithms reduces energy consumption and enhances the overall efficiency of financial transactions.

Sustainable practices encompass a holistic approach that integrates energy efficiency, environmental responsibility, and long-term sustainability goals. Energy-efficient protocols, such as Transport Layer Security (TLS) and Secure Shell (SSH), are critical components of

cryptographic operations. Efforts to optimize these protocols for energy efficiency focus on reducing computational overhead and resource utilization (Sikeridis et al.). For example, Cloudflare's implementation of TLS 1.3 prioritizes minimal energy consumption in cryptographic processes while maintaining robust security measures, contributing to sustainable cryptographic practices across the internet.

Effective management of cryptographic keys and certificates is essential for optimizing resource usage and minimizing energy consumption. Automated key management systems, certificate lifecycle management tools, and key rotation strategies streamline cryptographic processes while reducing energy waste. Efficient lifecycle management practices contribute to energy savings and operational efficiency in cryptographic operations (Leng et al.).

Optimizing data center operations is another key aspect of mitigating the environmental impact of cryptographic operations. Data centers are significant consumers of energy, including in cryptographic processes (Huang et al.). Advanced cooling techniques, like free cooling and liquid immersion cooling, contribute significantly to energy efficiency in data centers. For example, Google's data centers leverage free cooling systems, where ambient air or water is used for cooling instead of traditional mechanical refrigeration, reducing energy expenditure for cooling operations (Gong et al.). Overall, by reducing the energy load, the efficiency of high energy cooling systems is enhanced.

Server virtualization is another strategy employed to reduce energy consumption in data centers (Satra et al.). By consolidating multiple virtual machines onto fewer physical servers, server virtualization reduces idle hardware and optimizes resource utilization. This consolidation not only reduces energy consumption but also simplifies management and maintenance tasks, leading to overall operational efficiency gains. Furthermore, workload optimization and load balancing techniques help distribute computational tasks evenly across data center resources, preventing resource bottlenecks and reducing energy waste. Dynamic workload management systems adjust resource allocations in real time based on demand, optimizing energy usage without compromising performance (MirhoseiniNejad et al.). Singapore's data centers are also moving towards server virtualization and workload optimization strategies to improve resource utilization and reduce energy waste.

The integration of renewable energy sources into data centers is a pivotal step towards sustainable cryptographic practices. Renewable energy, including solar, wind, hydroelectric, and geothermal power, offers a clean and sustainable alternative to fossil fuel-based electricity generation, significantly reducing the carbon footprint of data centers and cryptographic operations (Rahman et al.). Major technology companies have made substantial commitments to renewable energy integration in their data centers. For instance, Microsoft's Azure data centers are powered by 100% renewable energy, with investments in solar and wind projects to offset energy consumption (Acun et al.). By sourcing renewable energy certificates and engaging in power purchase agreements (PPAs) for renewable energy, Microsoft ensures that its data center operations are environmentally responsible (Hundt et al.). This shift not only reduces the carbon footprint of data centers but also supports Singapore's aim to become a green technology leader.

Similarly, Amazon Web Services (AWS) has launched renewable energy projects globally, including wind and solar farms, to power its data centers sustainably (Xu et al.). AWS's commitment to achieving net-zero carbon emissions includes renewable energy integration in cryptographic operations, contributing to overall sustainability goals.

These case studies highlight the tangible operational efficiencies that result from adopting sustainable practices in cryptographic operations. By leveraging energy-efficient algorithms, optimizing data center operations, integrating renewable energy sources, and adhering to regulatory standards, organizations can further mitigate the environmental impact of cryptographic processes while ensuring robust security and operational reliability.

The adoption of green data center principles promotes sustainability in cryptographic operations by integrating energy-efficient infrastructure design, renewable energy sources, and waste heat reuse. Green data center initiatives aim to maximize resource utilization, minimize energy waste, and reduce carbon emissions in cryptographic processes. Collaborative efforts between data center operators, environmental organizations, and regulatory bodies drive the adoption of sustainable practices in data center operations, benefiting cryptographic activities as well.

The approaches of these global giants align with Singapore's initiatives to integrate renewable energy and mitigate the environmental impact of industrial processes. Major technology companies in Singapore are committing to sourcing their energy from renewable sources.

In conclusion, mitigating the environmental impact of cryptographic operations requires a multifaceted approach that integrates energy-efficient algorithms, data center optimizations, renewable energy integration, sustainable practices, and regulatory compliance. Industry initiatives, case studies, and technological advancements demonstrate the feasibility of achieving energy-efficient and environmentally responsible cryptographic operations. By prioritizing sustainability in cryptographic practices, organizations can reduce their carbon footprint, contribute to environmental conservation, and align with global efforts towards a greener future.

### **Efficient cryptography and the Singapore Smart Nation Initiative**

The integration of renewable energy into data center operations in Singapore is a strategic move that aligns with the nation's ambition to be a global financial and technological hub. Beyond carbon reduction, this shift enhances the resilience and sustainability of the country's digital infrastructure, reducing reliance on fossil fuels, improving energy security, and supporting community development through local job creation. These efforts are in line with Singapore's Smart Nation initiative, which aims to leverage advanced technologies to drive economic growth and improve the quality of life for its citizens.

Singapore's commitment to renewable energy integration positions it as an economic and environmental leader, attracting global financial institutions and tech companies drawn by the city's dedication to sustainability and innovation. This influx of companies would bolster Singapore's GDP and strengthen its reputation as a pioneer in green technology. The transition to

renewable energy is expected to spur significant advancements in green technologies within Singapore, including cutting-edge cooling systems, energy-efficient cryptographic algorithms, and urban-specific renewable energy solutions. These innovations could be commercialized and exported, establishing Singapore as a key player in the global green tech market.

As Singapore integrates renewable energy into data centers, the government is likely to implement progressive policies to encourage energy-efficient practices. These could include tax incentives, grants for green tech development, and stringent environmental standards, promoting sustainability and setting a global benchmark. This shift towards sustainable data centers will also necessitate a workforce skilled in both IT and renewable energy technologies, prompting Singapore's educational institutions to offer specialized programs in green technology and sustainable computing.

In the financial sector, the efficient and secure operation of cryptographic algorithms is crucial for supporting trading activities, which are vital to Singapore's economy. The energy demands of data centers hosting financial platforms and trading algorithms are substantial, reflecting the critical role of digital technologies in financial transactions. A deeper analysis could quantify the proportion of Singapore's GDP attributable to financial services and trading activities reliant on data centers and cryptographic algorithms, highlighting the energy intensity of these sectors and their implications for Singapore's energy landscape.

The growth of internet-based trading platforms further amplifies the demand for energy-efficient cryptographic solutions, as the energy footprint of these operations becomes more pronounced. Singapore's strategic focus on integrating renewable energy into data center operations supports the country's Smart Nation vision, enhancing its resilience, attracting global investments, and setting new standards for green technology. This proactive approach contributes to sustainable economic growth, technological advancement, and the global transition towards a more sustainable and resilient future.

## Works Cited

- Abbass, Kashif, et al. "A Review of the Global Climate Change Impacts, Adaptation, and Sustainable Mitigation Measures." *Environmental Science and Pollution Research*, vol. 29, no. 28, 2022, pp. 42539–59.
- Abdalla, Michel, et al. "Robust Encryption." *Journal of Cryptology*, vol. 31, no. 2, 2018, pp. 307–50.
- Abdullah, Nurul, and Amirah Lim. "The Incorporating Sustainable and Green IT Practices in Modern IT Service Operations for an Environmentally Conscious Future." *Journal of Sustainable Technologies and Infrastructure Planning*, vol. 7, no. 3, 2023, pp. 17–47.
- Acun, Bilge, et al. "Carbon Explorer: A Holistic Framework for Designing Carbon Aware Datacenters." *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2*, 2023, pp. 118–32.
- Adewole, Kayode, et al. "Application of Cryptocurrencies Using Blockchain for E-Commerce Online Payment." *Blockchain for Cybersecurity and Privacy*, CRC Press, 2020, pp. 263–305.
- Berret, Charles. *The Cultural Contradictions of Cryptography*. 2019. Columbia University, PhD Thesis.
- Busby, James A., et al. "The IBM 4769 Cryptographic Coprocessor." *IBM Journal of Research and Development*, vol. 64, no. 5/6, 2020, pp. 3–1.
- Carreira-Perpinán, Miguel A., and Yerlan Idelbayev. "'Learning-Compression' Algorithms for Neural Net Pruning." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 8532–41.
- Daryabar, Farid, et al. "Forensic Investigation of OneDrive, Box, GoogleDrive and Dropbox Applications on Android and iOS Devices." *Australian Journal of Forensic Sciences*, vol. 48, no. 6, 2016, pp. 615–42.
- Das, Manik Lal, and Navkar Samdaria. "On the Security of SSL/TLS-Enabled Applications." *Applied Computing and Informatics*, vol. 10, no. 1–2, 2014, pp. 68–81.
- Gong, Yuexuan, et al. "Advancements on Mechanically Driven Two-Phase Cooling Loop Systems for Data Center Free Cooling." *International Journal of Refrigeration*, vol. 138, 2022, pp. 84–96.
- Gookyi, Dennis Agyemanh Nana, et al. "NIST Lightweight Cryptography Standardization Process: Classification of Second Round Candidates, Open Challenges, and Recommendations." *Journal of Information Processing Systems*, vol. 17, no. 2, 2021, pp. 253–70.
- Hou, ZeZhou, et al. "Improve Neural Distinguishers of Simon and Speck." *Security and Communication Networks*, vol. 2021, no. 1, 2021, p. 9288229.
- Huang, Pei, et al. "A Review of Data Centers as Prosumers in District Energy Systems: Renewable Energy Integration and Waste Heat Reuse for District Heating." *Applied Energy*, vol. 258, 2020, p. 114109.

- Hundt, Steffen, et al. "Power Purchase Agreements and Financing Renewables: An Interdependency." *Journal of Structured Finance*, vol. 27, no. 1, 2021, pp. 35–50.
- Interesse, Giulia. *Singapore's Data Center Sector: Regulations, Incentives, and Investment Prospects*. 1 Sept. 2023, <https://www.aseanbriefing.com/news/singapores-data-center-sector-regulations-incentives-and-investment-prospects/>.
- Ketha, Anirudh. "The Evolution of Cryptography and a Contextual Analysis of the Major Modern Schemes." *NHSJS Reports*, 2024.
- Kobeissi, Nadim. "An Analysis of the Protonmail Cryptographic Architecture." *Cryptology ePrint Archive*, 2018.
- Kohli, Varun, et al. "An Analysis of Energy Consumption and Carbon Footprints of Cryptocurrencies and Possible Solutions." *Digital Communications and Networks*, vol. 9, no. 1, 2023, pp. 79–89.
- Leng, Jiewu, et al. "Blockchain-Empowered Sustainable Manufacturing and Product Lifecycle Management in Industry 4.0: A Survey." *Renewable and Sustainable Energy Reviews*, vol. 132, 2020, p. 110112.
- Li, Nan. "Research on Diffie-Hellman Key Exchange Protocol." *2010 2nd International Conference on Computer Engineering and Technology*, vol. 4, 2010, pp. V4-634-V4-637, <https://doi.org/10.1109/ICCET.2010.5485276>.
- Mahto, Dindayal, and Dilip Kumar Yadav. "Performance Analysis of RSA and Elliptic Curve Cryptography." *Int. J. Netw. Secur.*, vol. 20, no. 4, 2018, pp. 625–35.
- Manganelli, Matteo, et al. "Strategies for Improving the Sustainability of Data Centers via Energy Mix, Energy Conservation, and Circular Energy." *Sustainability*, vol. 13, no. 11, 2021, p. 6114.
- McCurley, Kevin S. "Cryptography and the Internet: Lessons and Challenges." *Advances in Cryptology — ASIACRYPT '96*, edited by Kwangjo Kim and Tsutomu Matsumoto, Springer Berlin Heidelberg, 1996, pp. 50–56.
- Milanov, Evgeny. "The RSA Algorithm." *RSA Laboratories*, 2009, pp. 1–11.
- MirhoseiniNejad, SeyedMorteza, et al. "Joint Data Center Cooling and Workload Management: A Thermal-Aware Approach." *Future Generation Computer Systems*, vol. 104, 2020, pp. 174–86.
- Naeem, Samreen. "Network Security and Cryptography Challenges and Trends on Recent Technologies." *Journal of Applied and Emerging Sciences*, vol. 13, no. 1, 2023, pp. 01–08.
- Nagendra, M., and M. Chandra Sekhar. "Performance Improvement of Advanced Encryption Algorithm Using Parallel Computation." *International Journal of Software Engineering and Its Applications*, vol. 8, no. 2, 2014, pp. 287–96.
- Rahman, Abidur, et al. "Environmental Impact of Renewable Energy Source Based Electrical Power Plants: Solar, Wind, Hydroelectric, Biomass, Geothermal, Tidal, Ocean, and Osmotic." *Renewable and Sustainable Energy Reviews*, vol. 161, 2022, p. 112279.

- Rösler, Paul, et al. "More Is Less: On the End-to-End Security of Group Chats in Signal, Whatsapp, and Threema." *2018 IEEE European Symposium on Security and Privacy (EuroS&P)*, IEEE, 2018, pp. 415–29.
- Satra, Ramdan, et al. "Comparison of Server Technologies Using Kernel-Based Virtual Machine and Container Virtualization." *AIP Conference Proceedings*, vol. 2595, no. 1, AIP Publishing, 2023.
- Sharma, Sonali, et al. "A Study of Post Quantum Cryptographic Security Model Using Symmetric Key Algorithm." *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, no. 5s, Apr. 2023, pp. 181–98.
- Sharon, Alita. "Singapore: AI and Quantum Transforming Financial Services." *OpenGov*, 9 July 2024, <https://opengovasia.com/2024/07/09/singapore-ai-and-quantum-transforming-financial-services/>.
- Sikeridis, Dimitrios, et al. "Assessing the Overhead of Post-Quantum Cryptography in TLS 1.3 and SSH." *Proceedings of the 16th International Conference on Emerging Networking Experiments and Technologies*, 2020, pp. 149–56.
- Sood, Roopali, and Harpreet Kaur. "A Literature Review on Rsa, Des and Aes Encryption Algorithms." *Emerging Trends in Engineering and Management*, 2023, pp. 57–63.
- Thabit, Fursan, et al. "Data Security Techniques in Cloud Computing Based on Machine Learning Algorithms and Cryptographic Algorithms: Lightweight Algorithms and Genetics Algorithms." *Concurrency and Computation: Practice and Experience*, vol. 35, no. 21, 2023, p. e7691.
- Ullah, Shamsheer, et al. "Elliptic Curve Cryptography; Applications, Challenges, Recent Advances, and Future Trends: A Comprehensive Survey." *Computer Science Review*, vol. 47, 2023, p. 100530.
- UNFCCC, United Nations Framework Convention on Climate Change. *The Paris Agreement*. <https://unfccc.int/process-and-meetings/the-paris-agreement>.
- Wendl, Moritz, et al. "The Environmental Impact of Cryptocurrencies Using Proof of Work and Proof of Stake Consensus Algorithms: A Systematic Review." *Journal of Environmental Management*, vol. 326, no. Pt A, Jan. 2023, p. 116530, <https://doi.org/10.1016/j.jenvman.2022.116530>.
- Xu, Minxian, et al. "A Self-Adaptive Approach for Managing Applications and Harnessing Renewable Energy for Sustainable Cloud Computing." *IEEE Transactions on Sustainable Computing*, vol. 6, no. 4, 2020, pp. 544–58.
- Yilek, Scott, et al. "When Private Keys Are Public: Results from the 2008 Debian OpenSSL Vulnerability." *Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement*, 2009, pp. 15–27.
- Zhou, Boyou, et al. "High-Performance Low-Energy Implementation of Cryptographic Algorithms on a Programmable SoC for IoT Devices." *2017 IEEE High Performance Extreme Computing Conference (HPEC)*, IEEE, 2017, pp. 1–6.

# The Impact of Cryptocurrencies and Central Bank Digital Currencies on Financial Stability: A Cross-Country Analysis By Mingrui Deng<sup>67</sup>

## Abstract

This paper examines the rise of cryptocurrencies and the development of Central Bank Digital Currencies (CBDCs), analyzing their distinct impacts on monetary policy, financial stability, and international trade. Through a comprehensive analysis, this study reveals that while cryptocurrencies like Bitcoin and Ethereum provide decentralized and borderless transaction capabilities, they significantly challenge traditional monetary control and regulatory frameworks. Conversely, CBDCs offer a promising digital evolution of sovereign currencies, designed to enhance payment systems, promote financial inclusion, and mitigate the risks posed by cryptocurrencies. The findings highlight the nuanced roles of cryptocurrencies and CBDCs in shaping the future of global financial systems, offering new insights into the potential benefits and challenges of digital currency adoption.

**Keywords:** Cryptocurrency; central bank digital currencies, financial stability; monetary policy; digital finance

## 1. Introduction

Cryptocurrencies like Bitcoin and Ethereum have been a subject of extensive research. Berentsen and Schär (2018) discuss the decentralized nature of cryptocurrencies and how they pose challenges to traditional monetary policy frameworks by operating outside government control. This decentralization can undermine central banks' ability to regulate money supply and maintain economic stability (Berentsen, A., & Schär, F. 2018). Similarly, Liu and Tsyvinski (2018) analyze the investment characteristics of cryptocurrencies, highlighting their high volatility and speculative nature, which complicate their integration into established financial systems (Liu, Y., & Tsyvinski, A. 2018).

On the other hand, CBDCs are being developed by central banks to enhance payment efficiencies and retain monetary control. Adrian and Mancini-Griffoli (2019) argue that CBDCs can improve financial inclusion and payment security, providing central banks with new tools to counter the influence of decentralized cryptocurrencies. This could potentially strengthen the effectiveness of monetary policy and enhance financial stability (Adrian, T., & Mancini-Griffoli, T. (2019)). He et al. (2017) present a comprehensive framework for the design and implementation of CBDCs, emphasizing their potential to streamline monetary operations and improve the transmission of monetary policy (Adrian, T., & Mancini-Griffoli, T. (2019)). The impact of digital currencies on financial stability is a critical area of concern. Garratt and Martin (2021) highlight the systemic risks posed by cryptocurrencies, such as market volatility and the potential for cyber-attacks, which can destabilize financial systems. Conversely, CBDCs are viewed as stabilizing instruments. Auer and Böhme (2021) discuss how CBDCs can mitigate

---

<sup>67</sup> E-mail: jerrydengdmr@gmail.com



risks associated with private digital currencies by providing a stable, government-backed alternative (Auer, R., & Böhme, R. 2021).

Digital currencies also have significant implications for international trade. Avdokushin and Zhukovskii (2019) explore how cryptocurrencies can reduce transaction costs and increase the efficiency of cross-border payments, potentially transforming global trade dynamics. (World.digital.currency) However, the lack of regulation and potential for misuse pose significant challenges. In comparison, CBDCs, due to their centralized nature, are more secured and regulated means of facilitating international trade, though they require extensive international cooperation and standardization (Adrian & Mancini-Griffoli, 2019). The comparative analysis of cryptocurrencies and CBDCs reveals distinct advantages and challenges between two of them. On the one hand cryptocurrencies offer innovation and financial inclusion but pose regulatory and stability risks. CBDCs, on the other hand, promise enhanced control and stability but require careful design and opportunity cost to avoid disrupting existing financial systems (Chiu & Koepl, 2017; He et al., 2017).

The rest of the paper begins with a data and methodology section, where it describes the sources of data, the indices used, and the rationale for employing the Tobit regression model. The results section presents the findings from the analysis, including descriptive statistics, regression results, and robustness checks, with a focus on interpreting the implications of cryptocurrency adoption for financial stability. Finally, the conclusion summarizes the key insights, discusses potential limitations, and suggests avenues for future research.

## 2. Data and Methodology

In this study, we investigate the relationship between cryptocurrency scores and financial stability indices across various countries. Understanding this relationship is crucial as the rise of cryptocurrencies has significant implications for financial stability, monetary policy, and economic development. By analyzing the correlations between these variables, we aim to provide insights into whether cryptocurrencies contribute to or detract from financial stability.

### 2.1. Methodology and Expected Outcomes

To analyze the data, we will employ statistical methods to examine the correlations between the cryptocurrency score and the financial stability indices. The analysis will help determine whether there is a positive, negative, or no correlation between the variables.

- **Positive Correlation:** If we find a positive correlation, it would suggest that higher cryptocurrency adoption is associated with greater financial stability. This outcome could indicate that cryptocurrencies are being integrated into financial systems in a way that supports stability.
- **Negative Correlation:** A negative correlation would imply that higher cryptocurrency adoption is linked to lower financial stability. This could suggest that cryptocurrencies introduce risks and vulnerabilities to financial systems.

- **No Correlation:** If there is no significant correlation, it would indicate that cryptocurrency adoption neither enhances nor undermines financial stability, suggesting that other factors might be more influential in determining financial stability.

By understanding these relationships, policymakers and financial regulators can better assess the impact of cryptocurrency adoption on financial stability and develop strategies to mitigate potential risks while leveraging the benefits of digital currencies. (n.d.,Salgam G)

## 2.2 Strengths and Weaknesses of the Data Source

One of the strengths of our data is its comprehensiveness and recency. The 2023 wave of the Financial Stability Database and the Global Cryptocurrency Adoption Index provide up-to-date observations, which is crucial for analyzing current trends and impacts. Compared to other data sources in the literature, our dataset includes more recent observations, allowing us to capture the latest developments in cryptocurrency adoption and financial stability.

The data sources used in this study are particularly well-suited for our research objectives. The Financial Stability Database is recognized for its rigorous data collection methods, which enhances the reliability of our results. However, there are potential limitations to consider. For instance, certain populations might be overrepresented or underrepresented in the data, leading to possible selection bias. Additionally, changes in data collection methods over time could introduce inconsistencies.

By understanding these relationships and considering the strengths and weaknesses of our data, policymakers and financial regulators can better assess the impact of cryptocurrency adoption on financial stability and develop strategies to mitigate potential risks while leveraging the benefits of digital currencies.

## 2.3 Methodology

### 2.3.1 Overview

This study explores the relationship between cryptocurrency adoption and financial stability indices across a sample of countries. Given the rapid rise of cryptocurrencies and their potential to disrupt traditional financial systems, this analysis seeks to uncover whether these digital assets contribute to or undermine financial stability.

### 2.3.2 Data and Indices Used

The dataset comprises several key indices reflecting various aspects of financial stability, alongside a cryptocurrency score that captures the extent of cryptocurrency adoption within each country.

- **Cryptocurrency Score:** This score is derived from the Global Cryptocurrency Adoption Index, which ranks countries based on factors such as transaction volume, peer-to-peer exchange usage, and the overall penetration of cryptocurrency in the economy. The score is standardized to facilitate comparison across countries.

- **Financial Stability Indices:** Given that financial stability does not have a single direct measure, the study utilizes several indices to provide a comprehensive view:
  - **Banking Stability Index (BSI):** Measures the health and stability of the banking sector.
  - **Insurance Sector Stability Index (ISSI):** Assesses the stability of the insurance sector.
  - **Market Stability Index (MSI):** Evaluates the stability of financial markets.
  - **Credit Market Stability Index (CMSI):** Focuses on the stability of the credit market.
  - **Foreign Exchange Market Stability Index (FEMSI):** Assesses the stability of the foreign exchange market.
  -

### 2.3.3 Data Sources and Sample Size

The data spans from 2018 to 2023, providing a recent and comprehensive view of the evolving landscape of cryptocurrency adoption and financial stability. The primary data sources include:

- **Financial Stability Database (2023 wave):** Provides the financial stability indices for a wide range of countries.
- **Global Cryptocurrency Adoption Index (2023):** Offers a detailed assessment of cryptocurrency usage across different nations.

The study includes data from 70 countries, selected based on the availability of comprehensive financial stability data and cryptocurrency adoption scores. This sample size is sufficient to capture diverse economic contexts and levels of cryptocurrency integration, enhancing the robustness of the findings.

### 2.3.4 Selection of the Tobit Model

The Tobit model is employed in this analysis due to the censored nature of the financial stability indices. These indices have natural bounds, often with lower limits at zero or upper limits at a certain maximum value. The Tobit model is particularly suited for this scenario because it can handle cases where the dependent variable is censored, ensuring that the estimator is unbiased and consistent.

### 2.3.5 Strengths of the Tobit Model:

- **Handling Censored Data:** It accounts for the fact that the financial stability indices cannot take on values outside their defined range.
- **Efficiency:** Provides efficient estimates in cases where the ordinary least squares (OLS) method might fail due to the censored nature of the data.

### 2.3.6 Limitations:

- **Assumption of Normality:** The Tobit model assumes that the error terms are normally distributed, which may not always hold in practice.
- **Sensitivity to Specification:** The results can be sensitive to the choice of censoring points and model specification.

### 2.3.7 Analysis and Expected Outcomes

The study employs the Tobit regression model to analyze the correlations between the cryptocurrency score and the various financial stability indices. The analysis seeks to determine whether the relationship is positive, negative, or neutral.

- **Positive Correlation:** Suggests that higher cryptocurrency adoption is associated with enhanced financial stability, potentially indicating successful integration of digital assets into traditional financial systems.
- **Negative Correlation:** Implies that increased cryptocurrency usage may introduce risks, such as volatility or regulatory challenges, that could destabilize financial systems.
- **No Correlation:** Indicates that other factors might be more critical in determining financial stability, and cryptocurrency adoption does not play a significant role.

### 2.4 Estimation Method and techniques:

In this study, we use Tobit regression models to estimate the parameters and examine the relationship between cryptocurrency scores and various financial stability indices. The Tobit model is appropriate for this analysis due to the nature of our dependent variables, which are censored and cannot take on values below or above certain thresholds.

#### 2.4.1 Notation and Model Specification

Let  $y_i^*$  represent the latent dependent variable, which in this case is the financial stability index. The observed variable  $y_i$  is censored and defined as:

$$y_i = y_i^* \text{ if } L \leq y_i^* \leq U$$

$$y_i = L \text{ if } y_i^* \leq L$$

$$y_i = U \text{ if } y_i^* \geq U$$

where  $L$  and  $U$  are the lower and upper bounds, respectively. The Tobit model is then specified as:

$$y_i = \beta_0 + \beta_1 gdp_{percap}_i + \beta_2 cryptocurrency\_score_{c_i} + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2)$$

where  $gdppercapita_i$  is the GDP per capita, and  $cryptocurrency\_score_{c_i}$  is the cryptocurrency adoption score.

#### 2.4.2 Estimation Technique

We use the maximum likelihood estimation (MLE) technique to estimate the parameters of the Tobit model. MLE is suitable for Tobit models as it handles the censored nature of the dependent variable effectively. The likelihood function for the Tobit model combines the probabilities of the censored and uncensored observations, and the parameters are estimated by maximizing this likelihood function.

#### 2.4.3 Controlling for Effects and Corrections

Throughout the estimation process, we address potential econometric issues such as multicollinearity, heteroskedasticity, and autocorrelation:

1. **Multicollinearity:** We check for multicollinearity by examining the Variance Inflation Factors (VIFs) for our independent variables. Variables with high VIFs are scrutinized, and if necessary, we consider excluding or transforming them.
2. **Heteroskedasticity:** To correct for heteroskedasticity, we employ robust standard errors. The use of robust standard errors helps to ensure that the inference drawn from our model is reliable even in the presence of heteroskedasticity.
3. **Autocorrelation:** While our cross-sectional data minimizes the risk of autocorrelation, we still check for its presence using the Durbin-Watson statistic and, if needed, adjust our model accordingly.

#### 2.5 Application and Justification

The Tobit regression models were applied to various dependent variables, including the Financial Stability Index (FSI), Banking Stability Index (BSI), and others. The results of these regressions provide insights into how cryptocurrency adoption correlates with different aspects of financial stability:

1. **Financial Stability Index (FSI):** The Tobit model results indicate the relationship between cryptocurrency scores and the overall stability of the financial system.
2. **Banking Stability Index (BSI):** This regression helps to understand the impact of cryptocurrency adoption on the stability of the banking sector.
3. **Market Stability Index (MSI):** This model assesses the effect on financial market stability.

Each of these models includes GDP per capita as a control variable to account for economic development differences across countries.

By utilizing the Tobit regression model and addressing potential econometric issues, this study aims to provide robust and insightful results on the impact of cryptocurrency adoption on financial stability.

### 3. Results

The results section presents the findings from the analysis of the relationship between cryptocurrency scores and financial stability indices. The purpose of this section is to test the hypotheses and provide empirical evidence on how cryptocurrency adoption affects various aspects of financial stability.

#### 3.1 Descriptive Statistics:

**Table 1. Descriptive Summary Statistics**

Variable	Mean	Standard Deviation	Min	Max
Cryptocurrency score	0.62	1.47	0.00	6.94
GDP per capita (in USD)	15735.77	19301.42	219.50	98280.40
Banking Stability Index	2.86	0.34	2.03	3.29
Market Stability Index	2.83	0.33	2.09	3.29
Credit Market Stability Index	3.12	0.14	0.14	3.38
Foreign Exchange Market Stability Index	3.01	0.17	2.58	3.38

Table 1 presents the descriptive statistics for the key variables used in the analysis. The mean cryptocurrency score is 0.62, reflecting a relatively low level of adoption across the sampled countries. The wide range in GDP per capita highlights the economic diversity among the countries included in the study.

#### 3.2 Regression Analysis Results:

**Table 2. Tobit Regression Results for Market Capitalization**

	Coefficient	Std.Error	t-value	P> t
GDP per Capita	0.0018954	0.0008283	2.29	0.025
Cryptocurrency Score	42.11326	139.6009	0.30	0.764
Constant	26.3985	33.169	0.80	0.429

Table 2 illustrates the tobit regression results of market capitalization.

- **GDP per Capita:** The positive and significant coefficients for GDP per capita in the market capitalization and credit market stability indices suggest that economic development positively influences financial stability and market growth. For instance, a unit increase in GDP per capita is associated with a 0.0019 increase in market capitalization, which indicates that wealthier countries generally have more robust financial markets.
- **Cryptocurrency Score:** The coefficients for the cryptocurrency score are positive across all models but lack statistical significance, except in the credit market stability index, where the p-value is close to the threshold for significance (0.056). This suggests that while cryptocurrency adoption may be starting to impact certain aspects of financial stability, particularly in credit markets, the effect is still limited or emerging.

**Table 3. Tobit Regression Results for Banking Stability Index**

	Coefficient	Std.Error	t-value	P> t
GDP per Capita	0.0000189	0.0000132	1.43	0.161
Cryptocurrency Score	1.338593	1.059173	1.26	0.214
Constant	3.058698	0.1343272	22.77	0.000

Table 3 presents the tobit regression results of the banking stability index.

- **GDP per Capita:** The coefficient for GDP per capita is very small (0.0000189) and not statistically significant (p-value = 0.161), indicating a weak and statistically insignificant relationship between economic development and banking stability within the sampled countries.
- **Cryptocurrency Score:** The coefficient for the cryptocurrency score (1.338593) suggests a positive relationship between cryptocurrency adoption and banking stability, but this result is not statistically significant (p-value = 0.214). Therefore, no strong conclusions can be drawn about the effect of cryptocurrencies on banking stability based on this model.

**Table 4. Tobit Regression Results for Credit Market Stability Index**

	Coefficient	Std.Error	t-value	P> t
GDP per Capita	0.001696	0.000559	3.03	0.004
Cryptocurrency Score	274.1759	139.8235	1.96	0.056
Constant	32.62287	15.84448	2.06	0.045

Table 4 shows the tobit regression results of the credit market stability index.

### 3.3 Addressing Potential Endogeneity

One concern in this analysis is the possibility of endogeneity, where variables such as GDP per capita and cryptocurrency adoption could be influenced by unobserved factors or could themselves influence each other. For instance, more economically stable countries might be more inclined to adopt cryptocurrencies, rather than cryptocurrency adoption driving financial stability.

To address potential endogeneity, future research could apply instrumental variable (IV) techniques, where variables that affect cryptocurrency adoption but not financial stability directly (e.g., technological infrastructure or internet penetration rates) are used to isolate the effect of cryptocurrency on financial stability.

#### 3.1.4 Correlation Analysis:

**Table 5. Correlation Analysis**

	GDP per Capita	Cryptocurrency Score
Financial Stability Index	0.2217	0.0526
GDP per Capita	1.0000	-0.1576
Cryptocurrency Score	-0.1576	1.0000

Table 5 provides the correlation matrix for the main variables. The weak positive correlation between GDP per capita and the financial stability index suggests a modest relationship where wealthier countries tend to have more stable financial systems. The correlation between cryptocurrency scores and financial stability is minimal, indicating that cryptocurrency adoption is not yet a significant determinant of financial stability.

### 3.4 Robustness Checks



Robustness checks were conducted to ensure the reliability of the results. These included testing for multicollinearity using VIFs, which indicated no serious issues. Additionally, heteroskedasticity was addressed by using robust standard errors. The main findings remained consistent across these checks.

#### **4. Conclusion**

This study provides valuable insights into the relationship between cryptocurrency adoption and financial stability. The results show that while GDP per capita is positively related to financial stability and market capitalization, cryptocurrency adoption currently does not have a statistically significant impact on financial stability metrics.

These findings align with literature like Berentsen and Schär (2018) and Liu and Tsyvinski (2018), which highlight the challenges cryptocurrencies pose without necessarily destabilizing financial systems. However, the near-significant relationship between cryptocurrency adoption and credit market stability suggests that cryptocurrencies may start to impact financial systems more as adoption grows, challenging some concerns about their destabilizing effects.

The strong link between GDP per capita and financial stability highlights the importance of economic development for stable financial systems. While the current lack of significant impact from cryptocurrencies suggests no immediate need for stringent regulations, regulators should monitor their growth, particularly in credit markets, and ensure preparedness for potential risks. Future research should revisit the impact of cryptocurrencies on financial stability as adoption increases, focusing on specific financial markets. Investigating the potential endogeneity between cryptocurrency adoption and financial stability, possibly using instrumental variables, could provide clearer insights. Additionally, exploring the role of Central Bank Digital Currencies (CBDCs) compared to decentralized cryptocurrencies could offer valuable perspectives on the future of global financial systems.

### Works Cited

- Adrian, T., & Mancini-Griffoli, T. (2019). The Rise of Digital Money. *International Monetary Fund (IMF) Finance & Development*. [Link](#).
- Auer, R., & Böhme, R. (2021). CBDCs: Costs and Benefits. *Bank for International Settlements (BIS) Working Papers*. [Link](#).
- Avdokushin, E., & Zhukovskii, V. (2019). Digital Currencies and International Trade. *Journal of International Commerce and Economics*. [Link](#).
- Berentsen, A., & Schär, F. (2018). Cryptocurrencies and Monetary Policy. *Federal Reserve Bank of St. Louis Review*. [Link](#).
- Chiu, J., & Koepl, T. V. (2017). The Economics of Cryptocurrencies – Bitcoin and Beyond. *Bank of Canada Staff Working Paper*. [Link](#).
- Garratt, R., & Martin, A. (2021). Financial Stability Implications of Digital Currencies. *Federal Reserve Bank of New York Staff Report*. [Link](#).
- He, D., Leckow, R., Haksar, V., et al. (2017). Central Bank Digital Currencies: A Framework for Assessing Why and How. *IMF Staff Discussion Note*. [Link](#).
- Liu, Y., & Tsyvinski, A. (2018). Cryptocurrencies: Stylized Facts on a New Investable Instrument. *NBER Working Paper Series*. [Link](#).

# How Convolutional Neural Networks Can Help Predict NFT Prices by Analyzing Images

## By Shunsuke Nakatani

### Abstract

This research explores the use of Convolutional Neural Networks (CNNs) to predict the market prices of Non-Fungible Tokens (NFTs) by analyzing their visual features. While traditional price prediction models rely heavily on historical price data and financial indicators, this study focuses on the unique artistic and visual aspects of NFTs, which often play a crucial role in determining their value. Using a dataset from the Moonbirds NFT collection, a CNN model was developed to extract image-based features such as colors, textures, and shapes, and predict corresponding market prices. The results reveal that while the CNN model could identify significant visual patterns, it often underestimated prices and produced a wide range of errors. This indicates the need for further refinement of the model and the incorporation of additional factors, such as metadata and market trends, to improve accuracy. Despite its limitations, this research provides a novel approach to NFT price prediction and offers insights for future work in blending visual data with traditional financial models to enhance market forecasting tools for investors.

### 1. Introduction

Non-fungible tokens (NFTs) represent a revolutionary way to prove ownership of digital assets (Sharma 2023). Unlike cryptocurrencies (The Investopedia Team 2024), which are fungible and identical to one another, NFTs are unique and indivisible, making them akin to digital collectibles or trading cards. Each NFT can represent anything from digital art to virtual real estate, and their value is often driven by their rarity, demand, and the perceived utility or access they provide, such as exclusive memberships or privileges. The NFT market has seen explosive growth, with some projects fetching extraordinary prices due to their unique characteristics and the benefits they offer to owners.

However, predicting the future value of NFTs remains a significant challenge. The rarity of an NFT and the specific attributes of its associated image play crucial roles in determining its market value. The sheer diversity and volume of NFT images make it nearly impossible for human evaluators to accurately assess the potential market price of each individual token. This is where machine learning, and more specifically, Convolutional Neural Networks (CNNs), comes into play (IBM 2024).

CNNs, a class of deep learning models particularly adept at image analysis, offer a promising solution to this problem. By analyzing the visual features of NFT images, CNNs can identify patterns and characteristics that correlate with higher market values. While previous research has explored the prediction of NFT prices based on historical price data, these approaches often overlook the significant influence of the visual appeal and artistic quality of the NFTs themselves (Ghosh et al. 2023).

This research aims to bridge this gap by developing models that predict NFT prices through image analysis using CNNs. By focusing on the visual elements of NFTs, we hope to provide a more comprehensive and accurate prediction model. Ultimately, our goal is to reduce the number of investment failures in the NFT market by equipping investors with better tools to evaluate the potential value of NFTs based on their images.

## **2. Background**

Research on forecasting NFT prices is still in its early stages. One significant approach is presented by Ghosh, I., Alfaro-Cortés, E., Gámez, M., and García-Rubio, N (Ghosh et al. 2023), which predicts NFT and DeFi prices using machine learning models. This study combines technical indicators, macroeconomic data, and sentiment analysis to create accurate price predictions, particularly during periods of financial instability. The models used, such as Gradient Boosting and Random Forest, have demonstrated strong performance in capturing trends based on past price movements and external factors. While this approach effectively incorporates financial and economic indicators, it does not consider the unique artistic and visual aspects of NFTs, which often play a significant role in their valuation.

Additionally, Baals, L. J., Lucey, B. M., Vigne, S. A., and Long's research offers a systematic projection of future research directions for the NFT market (Baals et al. 2022). This work emphasizes the need for further investigation into unexplored areas, highlighting the importance of integrating more diverse data sources and methodologies to fully understand the NFT market dynamics. While insightful, this research primarily focuses on market trends and does not delve into the potential of image-based analysis for NFTs.

Despite these advances, one key limitation remains: neither of these approaches fully addresses the influence of the art itself, which is often a major factor in determining an NFT's price. Unlike traditional financial assets, NFTs are largely valued based on the perceived quality and uniqueness of the artwork, and predicting their price without accounting for visual elements can lead to significant gaps in accuracy.

This gap in existing research motivates the need to explore how the art aspect of NFTs can be incorporated into price-prediction models. Specifically, convolutional neural networks (CNNs) have shown potential in analyzing visual data and could play a crucial role in forecasting NFT prices by interpreting the artistic features that drive value. By combining CNNs with traditional financial models, my research aims to provide a more comprehensive and accurate approach to NFT price prediction, addressing the limitations of prior studies.

## **3. Dataset**

For this research, the dataset used includes both numerical and vision data. Specifically, the dataset provides the price at which NFTs were sold, along with links to the corresponding

NFT images. Initially, this dataset can be classified as a language dataset, as it includes text-based information such as the image URLs. However, after downloading the images from these links, the dataset also becomes a vision dataset.

The dataset consists of NFT data from five different projects: Bored Ape Yacht Club (Amure 2024), Mutant Ape Yacht Club (Daly 2024), The Otherdeeds (“Otherdeed NFT Traits | Otherside Wiki”), Azuki (Daly 2024), and Moonbirds (Georgiev 2022). Given that each project varies significantly in terms of price range, utility, and popularity, the research focuses on one project as a proof of concept—Moonbirds. This decision helps isolate the impact of image-based features on NFT pricing without introducing variability from differing utilities across projects.

The dataset contained over 10 columns of information, such as block number, transaction hash, and marketplace, which were unnecessary for predicting NFT prices. Therefore, the dataset was simplified by retaining only the essential features: `usd_price` (the price of the NFT in USD), `name` (the name of the NFT), `image_url` (the link to the NFT’s image), and `collection_name` (the name of the NFT project). These features were relevant to understanding the relationship between visual aspects of the NFTs and their market price.

In total, the dataset consists of 2,000 samples from the Moonbirds collection. The data was split into two sets: 80% (1,600 samples) for training and 20% (400 samples) for evaluation. This split allowed us to train our model on a large portion of the data while keeping a separate set for evaluating its performance.

## **4. Methodology**

### **4.1. Modeling**

A Convolutional Neural Network (CNN) was built to predict NFT prices by analyzing image features. The architecture consists of two convolutional layers with 32 and 64 filters, respectively, followed by max-pooling layers to reduce the spatial dimensions. After extracting key image features such as colors, textures, and shapes, the output was flattened into a one-dimensional vector. This was then passed through fully connected layers, leading to a final prediction for the NFT price.

For training, 80% of the 2,000 NFT images were used. The model trained for 10 epochs, processing mini-batches of data to refine its predictions. Once trained, the model was saved for future use in predicting NFT prices based on image data.

### **4.2. Evaluation**

In the evaluation stage, the model's performance was tested using 20% of the dataset, consisting of 400 NFT images. These images were not seen by the model during training,

making them ideal for assessing how well the model generalizes to new data. The model's predictions were compared against the actual NFT prices, providing a measure of its accuracy.

During evaluation, the model was set to inference mode (`model.eval()`) to prevent further training. Predictions were made without adjusting the model's parameters. The predicted prices were then collected and compared with the true prices from the dataset. This comparison allowed for an assessment of how close the model's predictions were to the actual NFT prices, giving insight into the model's overall accuracy and performance.

## 5. Results

The performance of the Convolutional Neural Network (CNN) model in predicting NFT prices was evaluated by comparing its predicted values with the actual prices. Below are the key visual analyses from the evaluation phase.

### 5.1. Histogram of Differences Between Predictions and Actual Prices

The histogram (Fig. 1) plots the frequency of differences between predicted and actual NFT prices. The x-axis represents the differences between the model's predictions and the true values, while the y-axis shows how often these differences occurred. From the histogram, it's evident that there are substantial discrepancies between the predicted and actual prices, with many predictions significantly lower than the actual values. This pattern indicates that the model frequently underestimates NFT prices, which suggests a potential bias in the model's predictions that needs to be addressed. Additionally, the broad distribution of prediction errors highlights the model's inconsistency in predicting NFT prices accurately.

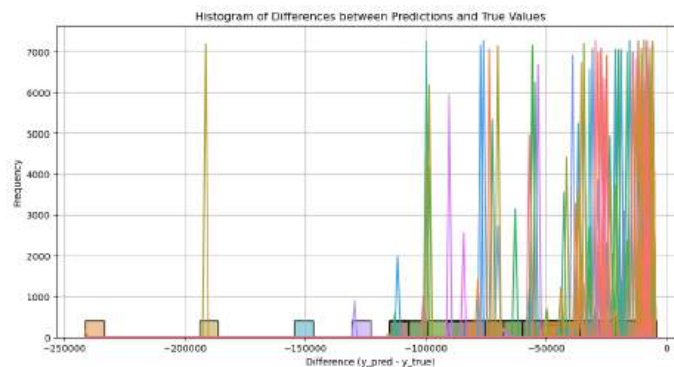


Figure1: Histogram Plot of Differences Between Predictions and Actual Prices

### 5.2. Box Plot of Prediction Differences

A more detailed analysis is provided by the box plot (Fig. 2), which summarizes the distribution of prediction errors. The box represents the interquartile range (middle 50%) of the errors, while the orange line inside the box denotes the median error. The whiskers extend to

show the overall range of errors, with several data points lying far from the whiskers, representing extreme outliers. The presence of these large prediction errors and outliers further underscores the model's lack of precision, as well as a significant variance in its predictions. The clustering of errors below the median line supports the earlier observation that the model tends to underestimate NFT prices. These results suggest that while the model captures some patterns in the data, its predictions are not sufficiently reliable for practical use without further refinement.

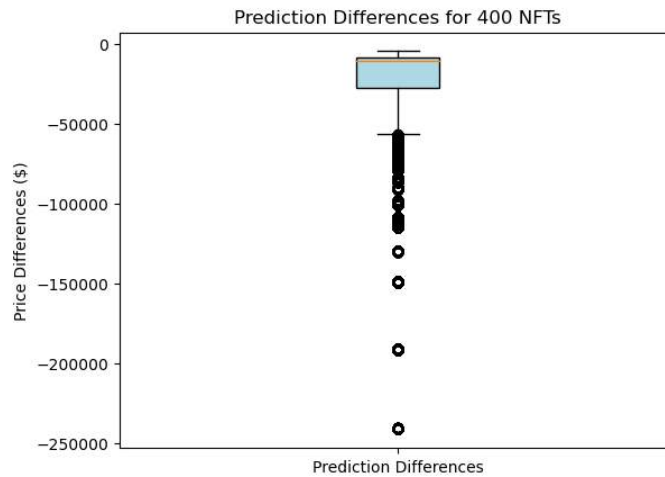


Figure 2: Box Plot of Differences Between Predictions and Actual Prices

### 5.3. Q-Q Plot of Prediction Differences

Lastly, the Q-Q plot (Fig. 3) compares the distribution of prediction errors to a normal distribution. Ideally, the points should align along the red diagonal line if the errors were normally distributed. However, in our case, many points deviate significantly from this line, particularly at the tails, indicating that the prediction errors are not normally distributed. This further demonstrates that the model has difficulty generalizing across the entire price range, leading to high variance in its predictions, especially for extreme cases.

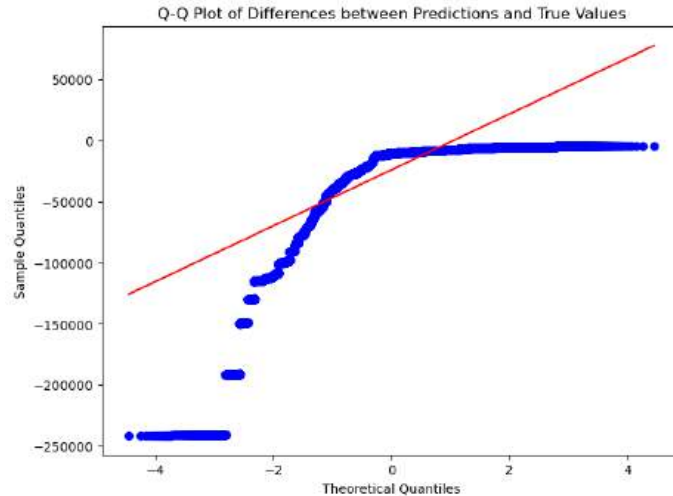


Figure 3: Q-Q Plot of Differences Between Predictions and Actual Prices

## 6. Discussion

The evaluation results show that while the CNN model can extract meaningful features from NFT images, it struggles with accurately predicting their market prices. The model frequently underestimates prices, as demonstrated by both the histogram and box plot analyses, and exhibits large prediction errors with several outliers. The Q-Q plot confirms that the model's prediction errors deviate from a normal distribution, particularly at the extremes.

These issues could stem from various factors, including limitations in the model architecture, insufficient training data, or complexity in the NFT market that the current model does not account for. To improve model performance, further optimization steps could include:

- **Data Augmentation:** Expanding the dataset or augmenting the existing data with variations of the images could help the model learn more robust features.
- **Advanced Architectures:** Implementing deeper architectures like ResNet or experimenting with transfer learning might improve the model's ability to capture more complex patterns in the data.
- **Fine-Tuning Hyperparameters:** Adjusting hyperparameters such as learning rate, batch size, and the number of epochs could lead to better convergence and accuracy.
- **Price-Related Features:** Incorporating non-image features, such as metadata about NFTs (e.g., artist reputation, rarity) could provide additional context for the price prediction.

Ultimately, while the CNN model shows some promise in predicting NFT prices based on image data, it requires significant improvements to achieve consistent and reliable results.



## **7. Conclusion**

In conclusion, this research highlights the potential of using Convolutional Neural Networks (CNNs) to predict the prices of NFTs by analyzing the visual characteristics of their images. The model successfully demonstrated that image-based features such as colors, textures, and shapes play a significant role in price determination, offering an innovative approach to understanding the NFT market. However, the results also exposed several challenges. The model tended to underestimate prices and produced a wide range of prediction errors, including notable outliers. This inconsistency suggests that while visual features are important, they alone may not provide a complete picture of an NFT's market value.

To enhance the model's accuracy, future research should consider integrating additional factors, such as NFT metadata (e.g., rarity, artist reputation, or utility) and exploring more advanced architectures like ResNet or transfer learning to capture complex patterns more effectively. Data augmentation could also help improve generalization by allowing the model to learn from a broader set of image variations. Lastly, fine-tuning hyperparameters such as learning rates and batch sizes could lead to better model performance. Despite the challenges, this research provides valuable insights into the role of image analysis in predicting NFT prices and opens avenues for further exploration in combining visual and non-visual features for more accurate market forecasting.

## **8. Acknowledgement**

I would like to express my sincere gratitude to Joe Xiao for his invaluable guidance and support throughout my research journey. His extensive expertise in AI and machine learning, combined with his passion for mentoring students, has greatly enriched my understanding of the field. I am particularly thankful for his encouragement and insight, which have inspired me to pursue independent research and strive for excellence in my work.

## Works Cited

- Sharma, Rakesh. "Non-Fungible Token Definition: Understanding NFTs." *Investopedia*, 6 Apr. 2023, [www.investopedia.com/non-fungible-tokens-nft-5115211](http://www.investopedia.com/non-fungible-tokens-nft-5115211). Accessed 12 Aug. 2024.
- The Investopedia Team. "Cryptocurrency Explained with Pros and Cons for Investment." Investopedia, 15 June 2024, [www.investopedia.com/terms/c/cryptocurrency.asp](http://www.investopedia.com/terms/c/cryptocurrency.asp). Accessed 12 Aug. 2024.
- Ghosh, Indranil, et al. "Prediction and Interpretation of Daily NFT and DeFi Prices Dynamics: Inspection through Ensemble Machine Learning & XAI." *International Review of Financial Analysis*, Feb. 2023, p. 102558, <https://doi.org/10.1016/j.irfa.2023.102558>. Accessed 12 Aug. 2024.
- IBM. "What Are Convolutional Neural Networks? | IBM." [www.ibm.com](http://www.ibm.com), IBM, 2024, [www.ibm.com/topics/convolutional-neural-networks](http://www.ibm.com/topics/convolutional-neural-networks). Accessed 13 Aug. 2024.
- Baals, Lennart John, et al. "Towards a Research Agenda on the Financial Economics of NFT's." SSRN Electronic Journal, Mar. 2022, <https://doi.org/10.2139/ssrn.4070710>.
- Amure, Tobi. "Bored Ape Yacht Club." Investopedia, 25 July 2024, [www.investopedia.com/bored-ape-yacht-club-6747275](http://www.investopedia.com/bored-ape-yacht-club-6747275). Accessed 13 Aug. 2024.
- Daly, Lyle. "What Is Mutant Ape Yacht Club? | the Motley Fool." The Motley Fool, 26 Mar. 2024, [www.fool.com/terms/m/mutant-ape-yacht-club](http://www.fool.com/terms/m/mutant-ape-yacht-club). Accessed 13 Aug. 2024.
- "Otherdeed NFT Traits | Otherside Wiki." Otherside-Wiki.xyz, [www.otherside-wiki.xyz/otherdeed](http://www.otherside-wiki.xyz/otherdeed). Accessed 14 Aug. 2024.
- Daly, Lyle. "What Is Azuki NFT? | the Motley Fool." The Motley Fool, 20 Mar. 2024, [www.fool.com/terms/a/azuki/](http://www.fool.com/terms/a/azuki/). Accessed 14 Aug. 2024.
- Georgiev, George. "What Are Moonbirds? Guide to the NFT Collection." CryptoPotato, 22 Apr. 2022, [cryptopotato.com/what-is-moonbirds-nft](http://cryptopotato.com/what-is-moonbirds-nft). Accessed 14 Aug. 2024.

# **The “Keys” to Success: Exploring the Impact of Music on the Abilities of Novice Surgeons**

## **By Rithika Jamisetty**

### **Abstract**

Music therapy is one of the many environmental stimuli present in the operating room (OR); not only is it a widespread practice because of its potential to alleviate patient anxiety during a surgical procedure, but it has become a normalized industry practice because of its perceived benefit to a surgical team’s focus. However, previous literature on the topic focused on the effects of music in the OR on the abilities of experienced surgeons rather than the less-equipped surgical trainees present in the OR. This study uses a results-based convergent synthesis design, wherein a meta-analysis of the time efficiency of several experimental trials was conducted in tandem with a thematic synthesis of the mental perceptions of novice surgeons to integrate and understand relevant aspects of the trainee surgeons’ learning process and cognitive performance. The quantitative analysis in this study used Hedge’s *g* effect size coefficient to standardize the mean differences in operating time between an intervention group (novice surgeons listening to music) and a control group. Ultimately, a medium effect size of 0.49, denoting moderate practical significance, demonstrated that the time to complete a surgical task is shorter in the control group. The qualitative synthesis found that the most relevant factors among previous studies determining a novice surgeon’s perception of music in the OR were cognitive capacity and experience level. It is thus affirmed that novice surgeons are more vulnerable to cognitive overload because of the increased number of intraoperative stressors; novices also lack the experience required to automate surgical tasks, further increasing the mental resources needed to complete a surgical task accurately. Highlighting the adverse impacts a supposedly beneficial interventional measure has on a vulnerable cohort in the OR clears a path for better stress-management practices to inform future training programs.

### **Introduction**

Digital media, once a hot commodity, is now available to a vast audience at the click of a button (Sanseverino et al.). Attractive to young users, music especially has been integrated more commonly than ever before into even the most mundane tasks - including studying. It is estimated that 75% of students listen to music while studying, justifying this habit with a perceived benefit in mood and creativity (Whittinghill et al.). Consequently, whether or not a student can efficiently learn academic material while listening to music has become a question frequently debated within the scientific community.

Not only has pre-existing research aimed to understand the role of music in the context of grade-school attendees, but there is also a growing body of literature that addresses the incorporation of music in professional circumstances, too. Music therapy, defined as the integration of varying types of music experience to nurture both mental and physical health, has been proposed as a possible solution to rising rates of stress in the work environment; specifically, in the fields of medical care, education, and customer service, music was seen to

decrease employee fatigue while increasing work efficiency (Mao). However, particularly vulnerable professions are those in the healthcare and educational sectors, especially in a post-pandemic era that continues to see heightened strain on caretaking professions (Abramson). While music has already been evaluated within the productive environments of both K-12 children and experienced professionals, the nexus between the two has not been explored - that is, novice surgeons (also referred to as surgical trainees throughout this paper), who are intending surgeons already actively in the professional world, but still in their foundational, educational years. This paper aims to address this population specifically, using a mixed-methods approach to address the topic more holistically. In the end, it is concluded that music is not beneficial for this cohort; music detracts overall efficiency in procedures due to novice surgeons' limited mental capacity and limited application experience, differentiating their abilities from experts.

## **Literature Review**

### **Music in Education**

The Mozart effect is popularly cited in various experimental studies done in the classroom setting; this highly disputed theory essentially suggests that listening to classical Mozart compositions results in temporary improvements in spatial reasoning and awareness (Rauscher et al.). Since, decades of research have been done to substantiate this effect, but no definitive conclusion has been reached. Among current competing explanations for the Mozart effect, the Yerkes-Dodson Law suggests that task performance is optimized under a moderate level of stress; however, when this threshold is either not met or exceeded, task performance is negatively affected. Therefore, reaching this intermediate level of stress maximizes cognitive function and productivity, which could be accomplished through a musical approach (May).

### **Music in Healthcare**

The addition of music therapy in the operating room (OR) has long been normalized, as 53% to 72% of surgical procedures worldwide incorporate music to calm both the surgical team and the patient (Sherer). As a result, several studies have aimed to understand the perceptions of musical intervention within a surgical team; one such study by Anantha Narayanan on the staff at a New Zealand hospital finds that surgical teams, on average, have a positive perception of the inclusion of music during procedures. From an online survey, it was found that only 6% of respondents thought playing music in the OR was inappropriate. Understandably, music was found to be preferred in familiar and lower-stakes procedures, as 84% of respondents said music is distracting in a crisis. Surgical team members incorporate music intending to improve calmness, mood, and general performance; still, music presents yet another cognitive weight in the OR (Narayanan and Gray).

### **Intraoperative Stressors in the OR**

The OR proves to be a site of particularly intense pressure for the medical staff because of the high degree of liability, the necessity of heightened precaution, and the rather unpredictable circumstances. As a result, researchers have aimed to scrutinize all existing external stimuli present in the OR to optimize the environment and achieve the most positive surgical outcome possible. One study by Ahreum Lee, a postdoctoral researcher at the University of Eastern Finland, categorized external stressors into seven main themes. This paper will generally focus on one: environmental (i.e. the elements that facilitate the smooth-running operation process). Lee's study noted that the most prevalent subtype studied to date within the environmental category was "acoustic", or the presence of music (Lee et al.). Moreover, a study by Victor Fu similarly categorizes certain stressors found within the OR. His study focused on the relationship between patient outcomes and auditory distractors and evaluated the extent to which background music and excess noise in the OR affect the surgical team. Fu finds that environmental noise is a stress inducer and it can have unintended detrimental effects on concentration, vital communication, and staff health (Fu et al.). His conclusion about the well-being of professionals being negatively affected by 'unwanted sound' is corroborated in a study by Cara Rogers, who added that it can hurt task productivity. Rogers finds that such inhibitors to performance in one's career can ultimately lead to increased rates of burnout, thus playing a role in overall job satisfaction and career longevity for many in stressful fields (Marie Rogers et al.).

### **Distraction vs. Solution**

There is a lack of consensus on the effects of music in the OR. On the one hand, researchers who believe in the stress-reducing potential of music find that the induced calmness from this intervention leads to improved performance, such as hand-eye coordination (Kacem et al.). Furthermore, music with varied tempos has been found to induce differences in operating times. Indeed, one study found that there was a 6.8% reduction in operating time with slower music and a 10% reduction in operating time with fast-paced music (Co et al.). To the opposite effect, researchers focused on possible hindering distractions in the OR find that, specifically, medical residents accumulate a greater number of errors with an increase in stress as a result of the addition of added stimuli (Tjønnås et al.). In all, it is imperative to further assess whether or not playing music enhances the atmosphere in the OR; fostering a more productive working environment can ensure greater work satisfaction from surgical professionals and even improve procedural success.

### **Effect of Music on Surgical Trainees' Performance**

Much of the existing research focuses on surgical experts' attitudes toward including music in the OR. There is not enough information about the perception of music by the less experienced cohort, surgical trainees, who are newly acquiring hands-on surgical experience. It is known that intricate procedures increase stress to a level that eventually surpasses the trainees' stress-coping ability. This can impact the trainee's ability to perform the surgical task at hand in a

safe manner (Carr et al.). One study by RWJ Mcleod finds that this cohort is especially vulnerable to stressors. The study found that background music at a high volume more significantly affected trainees than experts; Mcleod concluded that this would be an aspect of surgery that trainees must simply become acclimated to during their training (Mcleod et al.).

While this study does address the trainee population directly, an evaluation solely focused on the the impact of music on surgical trainees' performance and mental perception specifically has not been thoroughly fleshed out. This gap in the current literature is uniquely important because existing studies have primarily focused on either experienced surgeons or general healthcare settings, without addressing the specific needs and challenges faced by surgical trainees. Trainees are at a critical stage of skill development, where factors like stress, focus, and mental well-being significantly influence their learning curve and long-term proficiency. An evaluation dedicated solely to how music affects this group can provide targeted insights into optimizing their performance and mental state, contributing to more effective training environments and potentially better patient outcomes. Thus, this paper aims to address the question: What effect does the presence of music in the operating room have on the surgical performance of novice surgeons undergoing training, and what are the underlying factors influencing this outcome?

## **Method**

### **Overview**

To understand the role music plays in the abilities of trainee surgeons, a mixed-methods results-based convergent synthesis study was conducted. An article that presented recommendations for various mixed-methods designs within the medical field was reviewed to determine which synthesis model would best fit the question. The article found that papers analyzing the quantitative and qualitative elements separately but simultaneously should opt for a convergent synthesis model; this approach allows for the thorough evaluation of multiple perspectives independently, lending a more robust and nuanced conclusion.

In this study, the convergent synthesis design is necessary because it allows for the capturing of both measurable performance outcomes (quantitative) and the personal experiences or perceptions of the trainees (qualitative). Trainee surgeons' performance can be objectively evaluated through metrics like task completion times or accuracy rates, while their mental perceptions of music's impact—such as changes in stress levels or focus—require subjective, thematic analysis. The convergent design enables these two types of data to be integrated, giving a fuller understanding of music's dual role in affecting both technical performance and mental well-being. This methodology mirrors practices commonly used in clinical science, where quantitative meta-analyses of experimental trials are paired with qualitative thematic syntheses, ultimately yielding a more holistic evaluation of complex phenomena; this format was employed in the subsequent research process (Hong et al.).

Importantly, one foundational paper within the clinical science discipline by James Watson was consulted throughout the data collection process (Watson et al.). Similar to the goals of this paper, Watson used two distinct data types to assess clinical effectiveness and participant perceptions of a certain intervention outlined in his research question. This study mimics Watson's method protocol; a quantitative meta-analysis was conducted to pool treatment effect sizes to quantify the existence of an effect while a qualitative systematic synthesis was used to gain insight into mental/psychological factors from multiple perspectives. The variable chosen to measure performance numerically was the time it takes to complete a surgical task, labeled in many studies as time to completion (abbreviated as TTC). It is important to note that TTC was specifically chosen as the dependent variable for the quantitative studies because, generally, increased operative time is associated with an increased risk for surgical site infection and complications, which can be detrimental for both healthcare economics and the post-surgery quality of life for the patient. Therefore, it is in the best interest of surgeons to increase surgical efficiency and reduce operative time while still maintaining accuracy (Jackson et al.; Cheng et al.).

### **Eligibility Criteria**

First, to outline the requirements for the pool of studies that would be used for both the qualitative and quantitative phases, a systematic search approach using the PICO (patient, intervention, comparison, outcome) model was used. Such a structure is used in clinical research to narrow the search for relevant evidence under a targeted research question, as only studies that fit all four parameters were considered for further data analysis. For the participants, doctors currently in a surgical training program were heavily focused on, called trainee/novice surgeons, as this study targets those in pursuit of a career in surgery but still in the early stages of their career. Next, the intervention outlined was the playing of music in the operating theater of any genre (classical, soft rock, hip-hop, jazz, etc.). Third, the group of comparison was a control group of novice surgeons that were not listening to music; the participants in the control performed the same task/procedure without any musical intervention but still with other naturally occurring OR noises (from equipment or conversation). Lastly, the outcome required to be in the study being assessed was either TTC or a narrative assessment of mental impacts/perceptions of music by novice surgeons in the OR for the separate quantitative and qualitative aspects, respectively. Using the PICO model as a guiding structure to set highly specific inclusion and exclusion criteria reduced the risk of bias in the selection of sources that would have otherwise been inherent in a vague approach conducted by a single researcher.

### **Search Strategy and Study Selection**

After outlining the boundaries of the search, studies were collected to be used in further data analysis. A literature search was conducted from December 15, 2023, to January 16, 2024, using 3 main databases: PubMed, National Library of Medicine, and ScienceDirect. As unlimited access to all possible databases and every paper within the aforementioned areas was not

possible, login credentials were obtained from the University of Texas at Dallas to gain access to a larger number of studies. The following search terms were used in all three databases: («Surg» OR «Surgeons» OR «Medical» OR «Surgeon») AND «Music» AND («Novice» OR «Resident» OR «Student» OR «Trainee» OR «Expert»).

To further narrow the pool of studies for the data analysis phase, a three-step filter was implemented in the search algorithm that was used across all search engines: 1) papers only originally published in English to limit misinterpretation of the study's results, 2) studies dated from 2010 to present to ensure utilization of modern technology and practices, and 3) full texts were made available to extract the data necessary for later analysis.

### **Source Collection**

The next step was to organize the studies in a way that would allow for an efficient and accurate analysis. To uphold the method design, the qualitative and quantitative assessments were conducted separately; the two types of sources were kept in two different Google folders before carrying out their respective procedures. Preliminarily, basic study characteristics were recorded including the author, database, year published, language of origin, sample size, study design, and main findings. Details of the specific procedures or patient medical history were not included as this study is more focused on the ability of the surgeons to successfully perform in the presence of music rather than the context in which the procedure is occurring. Therefore, the collected data was not distinguished based on the types of surgical procedures or tasks when compiling surgeon performance outcomes.

This step introduces an important limitation of the methodology. By grouping several surgery types into one synthesis, the inherent nuances of the surgical procedures, such as operative requirements, performance standards, and technical demands, may be lost. Thus, the true relationship between OR interventions, their impact on trainees, and surgical outcomes may not represent all surgery types equally. However, a statistical strategy was used in the aforementioned study conducted by James Watson to account for such variation. Here, Watson recommended using a minimum of five studies when pooling data to ensure the statistical power (the probability that a test of significance will correctly identify an existing effect) of the results from the combined studies will be maximized, and thus yield a more accurate conclusion resistant to inter-study differences (Watson et al.).

### **Data Analysis**

#### **Quantitative Analysis**

Starting with the quantitative phase, the means and the standard deviations were recorded for both the control and experimental groups for five studies. To analyze their clinical significance within the context of the overall research question, the What Works Clearinghouse (WWC) Procedures provided by the U.S. Department of Education were used to characterize the intervention; the protocols outlined in this handbook provide standardized guidelines for



synthesis research methods and relevant mathematical interpretations (“What Works Clearinghouse Procedures and Standards.”). An effect size for each study was formulated using the Hedge’s g calculation provided by the WWC guidelines as seen in equation 1 and the associated pooled standard deviation  $SD_{*pooled}$  and small-sample bias-correction term  $\omega$  shown in equation 2 and equation 3, respectively.

$$(1) \quad Hedges\ g = \omega \frac{M1 - M2}{SD_{*pooled}}$$

$$(2) \quad SD * pooled = \sqrt{\frac{(n1-1)SD1^2 + (n2-1)SD2^2}{n1+n2-2}}$$

$$(3) \quad \omega = 1 - \frac{3}{4(n1+n2-2)-1}$$

Hedge’s g was used to calculate the standardized mean difference between the control and experimental groups for the continuous quantitative variable, time to completion. Hedge’s g is particularly suited for smaller sample sizes because it corrects for biases that occur with small groups, offering a more accurate estimate of the effect size compared to other coefficients, like Cohen’s d. In smaller sample sizes, Cohen’s d tends to overestimate effect sizes, while Hedge’s g adjusts for this by incorporating a correction factor that accounts for the sample size, leading to a less biased and more reliable estimate of the true effect. The results are presented in Table 1, where Hedge’s g allows for the comparison of mean values across varying sample sizes through a process called “weighting.” This means that studies with larger sample sizes contribute more to the overall effect size, while smaller studies are appropriately weighted to prevent them from skewing the findings. This weighting process ensures that the average effect size is more representative of the population, particularly in cases where sample sizes vary significantly across studies.

**Table 1**

Calculated effect sizes and associated standard deviations for control and intervention groups in the quantitative analysis using equations 1 and 2, respectively.

Study	Intervention Group (sample 1)			Control Group (sample 2)			Effect Size Estimate	
	$\bar{X}$ (sec.)	SD	n	$\bar{X}$ (sec.)	SD	n	Hedge's g	SE (g)
Shover, 2021	2.4	0.3	170	2.2	0.3	170	0.667	0.1111929006
Urps, 2023	61.4	8.45	36	59.5	10.8	36	0.1959	0.2337264232
Nees, 2021	157	38.5	16	120	17.3	16	1.239	0.376256616
Fu, 2020	206	255	42	211	207	42	0.021	0.2162218506

Muhammed, 2019	19.3	3.28	5	22	4.26	5	0.71	0.5889731801
----------------	------	------	---	----	------	---	------	--------------

For the standard error of the Hedge’s  $g$  effect size ( $SE(g)$ ) for each study, the WWC-provided formula shown in equation 4 was used to summarize the amount of overall variability in a particular dataset.

$$(4) \quad SE(g) = \omega \sqrt{\frac{n_1+n_2}{n_1n_2} + \frac{g^2}{2(n_1+n_2)}}$$

### Qualitative Analysis

To code for the qualitative portion of the studies, which included the more narrative reviews, a thematic synthesis was conducted. This part of the study very closely aligns with the coding process of a meta-analysis by Ahreum Lee, following a three-step plan: 1) dive into the results section of the paper to record conclusive statements of the paper, noting down their most significant findings, 2) engage in line-by-line coding to find all instances of possible factors relating to the effect of music on surgical performance among trainees, and 3) organize the initial codes into broader categories/themes utilizing their similarities and differences to reduce redundancy (Lee et al.).

The main difference between this paper and Lee’s meta-analysis is the number of researchers involved. The aforementioned paper had multiple researchers carry out different phases, and even had a fourth step to discuss and ensure consensus for coding decisions. Although this study did not integrate various perspectives, a smaller sample size of eleven studies was used for the qualitative portion of evaluation, making the findings more concise and relevant to this particular cohort of individuals in the healthcare field.

## Results

### Quantitative Studies

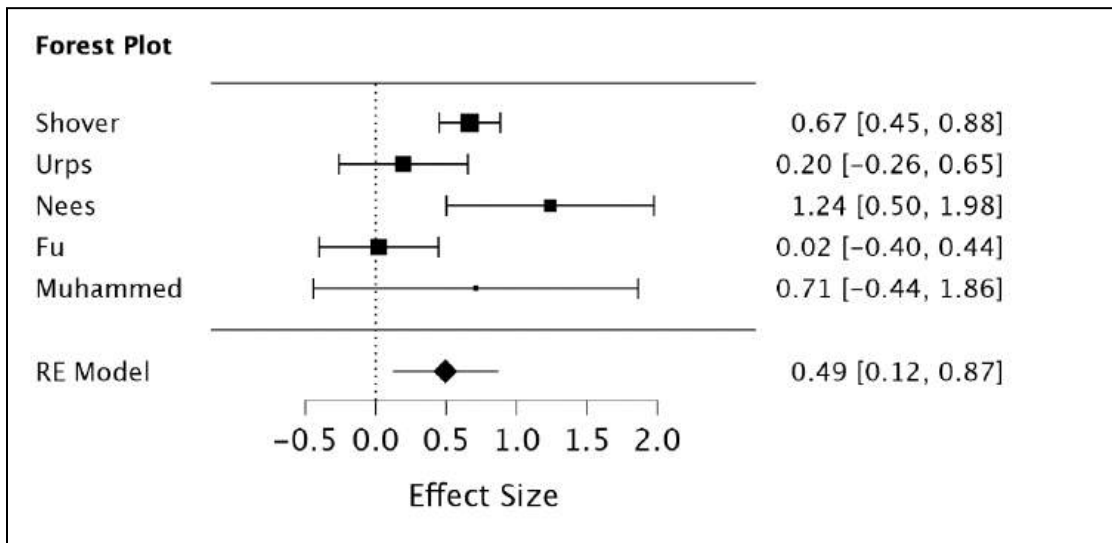
After analyzing the quantitative data using Hedge’s  $g$  equation and the associated standard error formula, a statistical analysis software, JASP Version 3.0, was used to visually represent the findings in a forest plot (“JASP 4.0.”).

A forest plot is a visual tool used in meta-analyses to display the results of multiple studies on the same topic. In Figure 1, each study’s effect size (a measure of the study’s outcome) is represented by a square. The size of the shape reflects how much that study contributes to the overall result—the larger the shape, the more weight it has, usually because it has a larger sample size or more precise data.

The horizontal lines extending from each shape represent the 95% confidence interval (CI). This interval shows a range of values that is likely to contain the true effect size. A narrow confidence interval suggests the study’s results are more precise, while a wider interval suggests

more uncertainty. If the confidence interval crosses the "line of no effect" (often represented as a vertical line down the center of the plot), it means that the study's result could potentially show no effect at all.

At the bottom of the forest plot, the black diamond represents the overall or "pooled" effect size from combining all the studies. The width of the diamond indicates the confidence interval for this combined result. This diamond is the most important part of the plot because it shows the final conclusion of the meta-analysis. If the diamond falls entirely on one side of the line of no effect, it means that the combined evidence points to a clear effect. If it crosses the line, the overall effect might not be statistically significant.



**Figure 1**

Forest plot of Meta-Analysis Using Hedge's  $g$  for the Numerical Effect Sizes

In the end, the point estimate of the average effect size, denoted by the black diamond at the bottom of the figure, is 0.49, with an associated confidence interval of (0.12 to 0.87); this value represents a medium effect size, meaning the finding has notable practical significance. To interpret the effect size, Cohen's established rule of thumb for evaluating effect sizes was used: the threshold for a small effect size is  $g \leq 0.2$ , medium effect size is then at  $0.2 \leq g \leq 0.5$ , and finally, large effect sizes are measured at  $g \geq 0.8$  (Cohen). Such a distinction is necessary to contextualize results as well as provide an understanding of the practical significance of an effect.

Moreover, the confidence interval of the average effect size value also lends significance. Essentially, the dotted vertical line represents the "line of null effect" on the forest plot; if an effect size extends over this boundary, there is a possibility that the true difference between the means of the control and experimental group is 0, presenting a null finding. In other words, there could be no significant relationship between an imposed intervention and a certain outcome. However, because the summative effect size value in this study does not cross the line of null

effect, it is possible to conclude that the findings of this paper do have practical significance and do present the existence of an effect - that is, that music does influence the time it takes to complete a surgical task (TTC). To understand if the influence is beneficial or hindering in nature, the plot needs to be understood within the context of the research question. Because the outcome of interest in the quantitative study (TTC) was not directly desirable (it is not being sought to be increased in a surgical procedure), a positive effect size to the right of the line of null effect favors the control group. Thus, TTC was found to be shorter in the control group of novice surgeons who were not listening to music while operating. The research article “Seeing the Forest by Looking at the Trees: How to Interpret a Meta-Analysis Forest Plot” accurately informed the interpretation of the quantitative findings from this visual (Dettori et al.).

Furthermore, the meta-analysis module available within JASP provided essential descriptive statistics, such as the overall p-value, which is the probability of obtaining a result as or more extreme under the assumption of no effect, and the  $I^2$  statistics, which is the percentage of inter-study variation that is attributed to heterogeneity rather than chance. The p-value of a study (presented in Table 2) and the heterogeneity estimates (shown in Table 3) are typically reported in papers following a systematic review and meta-analysis protocol (McKechnie et al.). Using a level of significance at the 0.05 level, a conventional standard, the resulting values prove to be statistically significant as  $0.010 < 0.05$ ; this means that the probability of attaining such a result under the assumption that there is no true relationship between music and TTC is incredibly low.

The heterogeneity assessments are provided as a way of ensuring that the appropriate meta-analysis model was used; generally, if  $I^2 > 50\%$ , a fixed-effects model should not be used as there are said to be considerable variances between the included studies (Schroll et al.). Using this, it can confirm that the right meta-analysis tool was used to calculate the findings, using the Hedge’s model.

**Table 2**  
Generated Meta-Analysis Coefficients Using JASP 3.0 Software

	Estimate	Standard Error	z	p	Standard Error	
					Lower	Upper
Intercept	0.495	0.193	2.567	0.010	0.117	0.872

**Table 3**  
Generated Residual Heterogeneity Estimates of Meta-analysis Using JASP 3.0 Software

		95% Confidence Interval
Estimate		

		Lower	Upper
$\tau$	0.109	0.006	1.795
$\tau^2$	0.330	0.080	1.340
$I^2$ (%)	66.820	10.468	97.071
$H^2$	3.014	1.117	34.146

### Qualitative Studies

Conversely, for the qualitative studies, precise measurements were not relied on to come to a definitive conclusion; rather, a line-by-line coding process was used to group existing, more narrative data into emerging themes. From the eleven final studies that were included, seven final themes were formulated (“limited capacity,” “stress impairs,” “preparation,” “music type,” “importance,” “experience,” and “intraoperative stressors”). The codes and their respective categories are denoted in Table 4, as well as the number of times each code was found within the qualitative data from the compiled studies. Because the focus of the qualitative study was to identify the psychosocial impacts music can cause, only thematic groups that fit this description were included. As a result, the two themes that appeared at both the highest frequency and with topical relevance were “limited capacity” and “experience”, containing nine and five codes within their categories, respectively. Thus, it was concluded that the most relevant factors impacting a novice surgeon's ability while listening to music in the OR were their ability to adjust to a variety of unpredictable stimuli (limited capacity) as well as their need to expend a greater amount of cognitive resources on a given surgical task (experience level).

**Table 4**

Qualitative Codes for Factors Affecting Novice Surgeons’ Learning Ability in the OR

Theme	Code	Code Frequency	Theme Frequency
Limited capacity	Mental Load	1	9
	Easy to mentally overwhelm	1	
	Low-Stress training benefits	2	
	Reduced stress progress	2	

	Clinical reasoning impaired	1	
	Stress Impairs Performance	1	
	Noise hinders concentration	1	
Preparation	Novice Accommodations	1	6
	Stress-coping mechanisms	2	
	Simulation-Based Training	2	
	Training Recommendations	2	
Music type	Music Preference	1	1
Importance	Reluctance to Admit Stress	1	1
Experience	Automation	2	5
	Experience Advantage	2	
	Novice vs. Expert	1	
Intraoperative stressors	Multiple Stimuli	2	5
	Stimuli Recognition	1	
	Task prioritization	1	

**Discussion**

## Key Findings

A mixed-methods study was conducted through a results-based convergent design to explore the research question: What effect does the presence of music in the operating room have on the surgical performance of novice surgeons undergoing training, and what are the underlying factors influencing this outcome? In the end, data from five quantitative studies of randomized experimental trials yielded an effect size of medium significance favoring the control conditions. On the other hand, the data from eleven qualitative studies expands on the previously mentioned trend by revealing two contributory psychological factors: limited mental capacity and lack of prolonged familiarization/experience.

Generally, an excess of noise in one's environment has been found to detract from their overall psychomotor performance (the speed and efficiency of cognitive processing) (Szalma and Hancock). The results of this study corroborate this overarching trend; after interpreting the average effect size value and its associated confidence interval, p-value, and statistics, it was concluded that the quantitative result demonstrated that the addition of music into the already bustling environment of the operating room (OR) posed a hindrance to the performance of novice surgeons, attributed to an increased procedural time. Juxtaposing this conclusion were the results of a recent study with similar aims that found that trainees listening to music in the OR were able to complete surgical tasks faster due to improvements in dexterity, allowing them to handle the equipment with greater efficiency (Hadjicosta and Halim). However, the focus on equipment efficiency within this study fundamentally fails to address how accurately they are used in the procedure; therefore, this study is limited in its generalizability to the overall abilities of novice surgeons.

The qualitative findings of this study were ultimately two-pronged, where the top thematic groups mentioning a particular mental factor at the highest frequency were prioritized. Firstly, the limited mental capacity of novice surgeons to handle unpredictable external stressors will affect their ability to quickly grasp and master highly technical surgical tasks. Notably, two of the studies finding observations within this theme mentioned the same concept that the volume and the complexity of work presented in the OR can lead to cognitive overload for trainees just beginning to attain experience in the field; multiple, independent stressors in isolation pose a challenge and contribute to a higher perceived task load. In summary, the environment itself hinders the performance of trainees, who, as a result, make a greater number of surgical errors or experience raised stress levels during a procedure (Poolton et al.; McNeer et al.). This is also reflected in the quantitative component of my study, as an effort intended to create a more relaxing environment through music, an additional stressor, had the unintended consequence of decreasing efficiency among trainee surgeons.

Second, the experience level of the surgeon determines how likely they are to handle greater exposure to intraoperative stimuli. Importantly, two studies mentioned related findings within this theme regarding the concept of automation: as more experience is acquired, less active mental processing is required when completing a routine task, and it can thus be done

automatically (Ghazanfar et al.; Hsu et al.). Comparatively, a previous systematic review evaluating the length of the learning curve among trainees, found that as surgeons attained more experience, they were able to improve accuracy through repeated trials despite an increase in the number of stressors faced in the OR. This demonstrates that the development of automation of surgical tasks is achieved simultaneously with a greater tolerance for a higher number of stressors, but until then, novice surgeons are still particularly vulnerable to excess distractions/stimuli (Miskovic et al.).

Overall, both aspects of the analysis lent one, complementary conclusion: music should be ruled out as an effective strategy to elevate the surgical atmosphere due to its negative effects on surgical performance and its detraction from essential cognitive resources within novices. A definitive stance is crucial, as a previous study addressing the vulnerability of trainee surgeons finds that changes to the work environment itself can manifest itself in practical benefits that address burnout (Johnson et al.).

### **Study Limitations**

Importantly, because secondary data sources were relied on throughout the paper, the participant pool was widely varied to maximize the amount of data available for analysis; as a result, variances in intangible variables, such as participant experiences, could not be accounted for. Because this paper aims to evaluate the preparedness of novice surgeons, it is important to realize that preparedness can be affected by differences in the efficacy of residency/training programs across regions and institutions. This could introduce confounding variables into the study, possibly affecting the ability to establish direct causation between the intervention of music and certain responses observed from novice surgeons. Nevertheless, individual participants' mental aptitude is not something that can be measured objectively, but a mixed-methods approach allowed for the minimization of the subjective nature that would otherwise arise from purely qualitative reports.

### **Applications and Future Directions**

Practically, this study offers insights into how stress in high-stakes medical situations, such as surgeries, can be better managed to improve performance and patient outcomes. Since robotically assisted procedures are not widely available and are only used in specific cases, most surgeries are still performed directly by surgeons, who face immense pressure to deliver consistent, high-quality care. The study suggests that music may not be an effective stress-reduction tool in the operating room. Instead, greater emphasis should be placed on developing stress-management practices, particularly for novice surgeons who are more vulnerable to stress.

To improve surgical training environments, several recommendations can be considered. First, integrating formal stress-management training into medical education could better prepare young surgeons for the realities of the operating room. This might include mindfulness techniques, mental resilience training, or simulation-based exercises that replicate high-pressure



situations in a controlled, lower-risk environment. Offering consistent mentorship and peer support can also create a more open culture where novice surgeons feel comfortable discussing stress and seeking help without fear of judgment.

Furthermore, training programs should focus on progressive skill-building, where trainees are gradually exposed to increasing levels of complexity and pressure. By fostering a more structured learning environment with clear goals and support systems, surgical trainees can build confidence and reduce anxiety over time. Finally, normalizing feedback and open discussions about the challenges in training can help address dissatisfaction and prevent burnout, creating a healthier, more supportive training culture. This holistic approach will ensure that the next generation of surgeons is better equipped to manage stress, leading to improved patient care and reduced burnout.

Not only does this study provide further evidence for the effectiveness of interventional measures in the OR, but it also raises new needs for the exploration of this topic. Future research should use controlled experimentation to understand how the brain receives different genres of music. Specifically, quantitative indicators of brain activity should be relied on to closely monitor cognitive functions in correspondence to music in the OR; for example, Sonal Arora suggested in her article that various chemical responses could be recorded as measurements of stress instead of relying on self-reported indicators, which may be inaccurate due to the reluctance within the community to admit personal burnout (Arora et al.). Additionally, there is a need for longitudinal qualitative research that investigates the long-term relationship between burnout rates and the implementation of interventional measures like music.

Zooming out, this study also aims to contribute to the broader question: How can music affect learning and overall productivity? While this conversation has taken its roots in the educational field, this study's findings still hold relevance; they suggest that in cognitively demanding settings, such as those with a high level of liability or intense technicality, music can hinder performance and the ability to acquire/apply new skills. Thus, the results provide a new understanding of the unique role of music in its various applications.

## Works Cited

- Sanseverino, Domenico et al. "'Don't Stop the Music,' Please: The Relationship between Music Use at Work, Satisfaction, and Performance." *Behavioral sciences (Basel, Switzerland)* vol. 13,1 15. 24 Dec. 2022, doi:10.3390/bs13010015
- Whittinghill, Joshuah et al. "Don't Stop the Music: An Exploration of First-Generation College Students' Use of Music during Pre-Pandemic and Pandemic Times." *Journal of Higher Education Theory and Practice*, Journal of Higher Education Theory and Practice, 29 Dec. 2021, doi.org/10.33423/jhetc.v21i15.4895.
- Mao, Nan. "The Role of Music Therapy in the Emotional Regulation and Psychological Stress Relief of Employees in the Workplace." *Journal of Healthcare Engineering*, U.S. National Library of Medicine, 29 Jan. 2022, www.ncbi.nlm.nih.gov/pmc/articles/PMC8817864/.
- Abramson, Ashley. "Burnout and Stress Are Everywhere." *2022 TRENDS-REPORT*, American Psychological Association, 1 Jan. 2022, www.apa.org/monitor/2022/01/special-burnout-stress.
- Rauscher, F H et al. "Music and spatial task performance." *Nature* vol. 365,6447 (1993): 611. doi:10.1038/365611a0
- May, Cindi. "Does Music Boost Your Cognitive Performance?" *Scientific American*, Scientific American, 20 Feb. 2024, www.scientificamerican.com/article/does-music-boost-your-cognitive-performance/.
- Sherer, Erin. "The Choices and Challenges Associated with Music in the Operating Room." *American Association of Surgical Physician Assistants* , Sutureline, 2015, www.aaspa.com/the-choices-and-challenges-associated-with-music-in-the-operating-room.
- Narayanan, Anantha, and Andrew R Gray. "First, do no harmony: an examination of attitudes to music played in operating theatres." *The New Zealand medical journal* vol. 131,1480 68-74. 17 Aug. 2018
- Fu, Victor X., et al. "The Perception and Attitude toward Noise and Music in the Operating Room: A Systematic Review." *Journal of Surgical Research*, Academic Press, 4 Mar. 2021, www.sciencedirect.com/science/article/pii/S0022480421000743.
- Marie Rogers, Cara, et al. "Effect of Various OR Noise on Fine Motor Skills, Cognition, and Mood." *Surgery Research and Practice*, 31 July 2019. *Gale Academic OneFile*, dx.doi.org/10.1155/2019/5372174. Accessed Mar. 2024.
- Kacem, I et al. "Effects of music therapy on occupational stress and burn-out risk of operating room staff." *The Libyan journal of medicine* vol. 15,1 (2020): 1768024. doi:10.1080/19932820.2020.1768024
- Co, Michael et al. "Effect of Background Music in the Operating Room on Surgical Outcomes: A Prospective Single-Blinded Case-Control Study." *Journal of the American College of Surgeons* vol. 235,3 (2022): 447-453. doi:10.1097/XCS.0000000000000279

- Tjønnås, M.S., Guzmán-García, C., Sánchez-González, P. et al. Stress in surgical educational environments: a systematic review. *BMC Med Educ* 22, 791 (2022).  
<https://doi.org/10.1186/s12909-022-03841-6>
- Carr, Shane et al. “Determining the Effect of External Stressors and Cognitive Distraction on Microsurgical Skills and Performance.” *Frontiers in surgery* vol. 6 77. 22 Jan. 2020, doi:10.3389/fsurg.2019.00077
- McLeod, RWJ et al. “The impact of noise in the operating theatre: a review of the evidence.” *Annals of the Royal College of Surgeons of England* vol. 103,2 (2021): 83-87. doi:10.1308/rcsann.2020.7001
- Hong, Q.N., Pluye, P., Bujold, M. et al. Convergent and sequential synthesis designs: implications for conducting and reporting systematic reviews of qualitative and quantitative evidence. *Syst Rev* 6, 61 (2017). <https://doi.org/10.1186/s13643-017-0454-2>
- Watson, James A et al. “Pain Neuroscience Education for Adults With Chronic Musculoskeletal Pain: A Mixed-Methods Systematic Review and Meta-Analysis.” *The journal of pain* vol. 20,10 (2019): 1140.e1-1140.e22. doi:10.1016/j.jpain.2019.02.011
- Jackson, Timothy D et al. “Does speed matter? The impact of operative time on outcome in laparoscopic surgery.” *Surgical endoscopy* vol. 25,7 (2011): 2288-95. doi:10.1007/s00464-010-1550-8
- Cheng, Hang et al. “Prolonged Operative Duration Increases Risk of Surgical Site Infections: A Systematic Review.” *Surgical infections* vol. 18,6 (2017): 722-735. doi:10.1089/sur.2017.089
- “What Works Clearinghouse Procedures and Standards.”. *WHAT WORKS CLEARINGHOUSE PROCEDURES AND STANDARDS HANDBOOK, VERSION 5.0.*, Institute of Education Sciences, Aug. 2022, [ies.ed.gov/ncee/wwc/Docs/referenceresources/Final\\_WWC-HandbookVer5.0-0-508.pdf](https://ies.ed.gov/ncee/wwc/Docs/referenceresources/Final_WWC-HandbookVer5.0-0-508.pdf).
- “JASP 4.0.” *JASP*, The JASP Team, 2018, [jasp-stats.org](http://jasp-stats.org).
- Cohen, Jacob. “Statistical Power Analysis - Jacob Cohen, 1992.” *Statistical Power Analysis*, Current Directions in Psychological Science, June 1992, [journals.sagepub.com/doi/10.1111/1467-8721.ep10768783](https://journals.sagepub.com/doi/10.1111/1467-8721.ep10768783).
- Dettoni, Joseph R et al. “Seeing the Forest by Looking at the Trees: How to Interpret a Meta-Analysis Forest Plot.” *Global spine journal* vol. 11,4 (2021): 614-616. doi:10.1177/21925682211003889
- McKechnie, Tyler et al. “Enhanced Recovery After Surgery (ERAS) protocols following emergency intra-abdominal surgery: A systematic review and meta-analysis protocol.” *PloS one* vol. 18,9 e0291140. 8 Sep. 2023, doi:10.1371/journal.pone.0291140
- Schroll, Jeppe B et al. “Dealing with substantial heterogeneity in Cochrane reviews. Cross-sectional study.” *BMC medical research methodology* vol. 11 22. 24 Feb. 2011, doi:10.1186/1471-2288-11-22
- Szalma, James L, and Peter A Hancock. “Noise effects on human performance: a meta-analytic synthesis.” *Psychological bulletin* vol. 137,4 (2011): 682-707. doi:10.1037/a0023987

- Hadjicosta, Olympia, and Jonathan Halim. "The Effect of Music on Laparoscopic Skills in Novices: A Randomised Controlled Trial." *SURGICAL EDUCATION*, Journal of the American College of Surgeons, Oct. 2020, pubmed.ncbi.nlm.nih.gov/21707130/.
- Poolton, Jamie M et al. "A comparison of evaluation, time pressure, and multitasking as stressors of psychomotor operative performance." *Surgery* vol. 149,6 (2011): 776-82. doi:10.1016/j.surg.2010.12.005
- McNeer, Richard R et al. "Intraoperative Noise Increases Perceived Task Load and Fatigue in Anesthesiology Residents: A Simulation-Based Study." *Anesthesia and analgesia* vol. 122,2 (2016): 512-25. doi:10.1213/ANE.0000000000001067
- Ghazanfar, Mudassar, et al. "The Effect of Divided Attention on Novices and Experts in Laparoscopic Task Performance." *Surgical Endoscopy*, U.S. National Library of Medicine, Mar. 2015, pubmed.ncbi.nlm.nih.gov/25030475/.
- Hsu, K E et al. "Experienced surgeons can do more than one thing at a time: effect of distraction on performance of a simple laparoscopic and cognitive task by experienced and novice surgeons." *Surgical endoscopy* vol. 22,1 (2008): 196-201. doi:10.1007/s00464-007-9452-0
- Miskovic, Danilo et al. "Learning curve and case selection in laparoscopic colorectal surgery: systematic review and international multicenter analysis of 4852 cases." *Diseases of the colon and rectum* vol. 55,12 (2012): 1300-10. doi:10.1097/DCR.0b013e31826ab4dd
- Johnson, Judith et al. "Burnout in Surgical Trainees: a Narrative Review of Trends, Contributors, Consequences and Possible Interventions." *The Indian journal of surgery* vol. 84,Suppl 1 (2022): 35-44. doi:10.1007/s12262-021-03047-y
- Arora, Sonal et al. "Stress impairs psychomotor performance in novice laparoscopic surgeons." *Surgical endoscopy* vol. 24,10 (2010): 2588-93. doi:10.1007/s00464-010-1013-2
- Lee, Ahreum et al. "Unpacking the Broad Landscape of Intraoperative Stressors for Clinical Personnel: A Mixed-Methods Systematic Review." *Journal of multidisciplinary healthcare* vol. 16 1953-1977. 17 Jul. 2023, doi:10.2147/JMDH.S401325

# **Contaminated Waters: Unveiling the Health Risks of Islamabad's Water Quality Crisis By Abeeha Atif**

## **Abstract**

This research investigates the degradation of water quality in Islamabad and its direct impact on public health. Water samples collected from key locations, such as Rawal Dam and different Islamabad neighborhoods, were analyzed for contamination levels, including turbidity, coliform bacteria, and harmful trace elements. Findings revealed that several areas have water quality far below acceptable standards, leading to widespread health issues such as hepatitis E. This study aims to document and assess water quality in Islamabad. It also identifies significant sources of pollution, including poor waste management and industrial discharge, and suggests solutions to improve water safety in Islamabad.

## **Introduction**

Pakistan, a nation marked by its complex geopolitical and environmental landscape, faces significant challenges regarding water resources. Since gaining independence in 1947, the country has struggled to manage its water supply effectively. One pivotal agreement governing water resources in Pakistan is the Indus Water Treaty, signed with India in 1960. This treaty, facilitated by the World Bank, aimed to equitably divide the use of the Indus River system between the two countries, ensuring a stable water supply. However, while the treaty has successfully prevented conflict over water resources, it has yet to address the internal water quality issues plaguing Pakistan, particularly in urban centers.

Islamabad, the capital city, is a microcosm of the broader national water crisis. The city's primary water sources include groundwater from tubewells and surface water from Rawal Dam. Unfortunately, both sources have experienced significant contamination due to industrial runoff, agricultural activities, and ineffective waste management systems. The deterioration of water quality has led to severe public health consequences, including a notable rise in waterborne diseases like hepatitis E, particularly among marginalized communities.

According to a report by the Pakistan Council of Research in Water Resources (PCRWR), in 2021 approximately 80% of Pakistan's water supply is estimated to be contaminated, exacerbating health crises across the nation (PCRWR, 2021). This statistic highlights the urgency of addressing water quality issues, especially in rapidly urbanizing cities like Islamabad, where over 2 million residents depend on these compromised water sources (Pakistan Bureau of Statistics, 2021). However, according to the PCRWR report, while water quality issues remain a significant challenge in Pakistan today, considerable progress has been made in improving access to clean water over the years. The report highlights noticeable improvements in water treatment and infrastructure between the early 2000s and more recent years, reflecting efforts to reduce contamination and provide safer drinking water. This progress demonstrates a positive shift, though ongoing challenges still persist, particularly in urban areas like Islamabad.

The situation in Islamabad is particularly concerning. The rapid urbanization of the city, coupled with a lack of adequate infrastructure to manage waste and treat water, has resulted in

serious public health risks. Waterborne diseases are highly prevalent, and residents are often left with little choice but to use contaminated water sources. The World Health Organization (WHO) reports that unsafe drinking water contributes to the deaths of over 500,000 children under five annually worldwide, a statistic that highlights the urgent need for improved water quality management (WHO, 2022).

This research aims to investigate the factors contributing to the deterioration of water quality in Islamabad, analyze water samples for pollutants, and explore the broader implications for public health. Understanding these factors is crucial for developing effective interventions to safeguard the health of Islamabad's residents.

#### Methodology

The methodology employed in this research involved the systematic collection and analysis of water samples from twelve different locations across Islamabad. These samples were obtained from both surface water (e.g., Rawal Dam) and groundwater (e.g., tubewells in various neighborhoods). The sample collection process adhered to the guidelines set forth by the National Water Quality Monitoring Program, ensuring proper handling to avoid contamination. Each sample was meticulously labeled with its source information, GPS coordinates, and collection time to maintain an organized profile for testing. Please see the supplementary information to see the form utilized during sampling events.

Sample collection was conducted following stringent protocols on 11th July, with 12 samples taken as they are representative of Islamabad's key water sources. This number aligns with the methodology used by PCRWR, which has conducted similar studies with a comparable sample size to effectively capture the city's water quality variations. For surface water, samples were taken from Rawal Dam, with upstream, midstream, and downstream points observed for contamination levels. Tubewell water samples were drawn from key locations, such as G-5 Government Hostel and G-11/1 near Masjid, where tubewell water serves as a primary source. Groundwater samples were allowed to run for ten minutes before sampling, ensuring an accurate representation of the water quality.

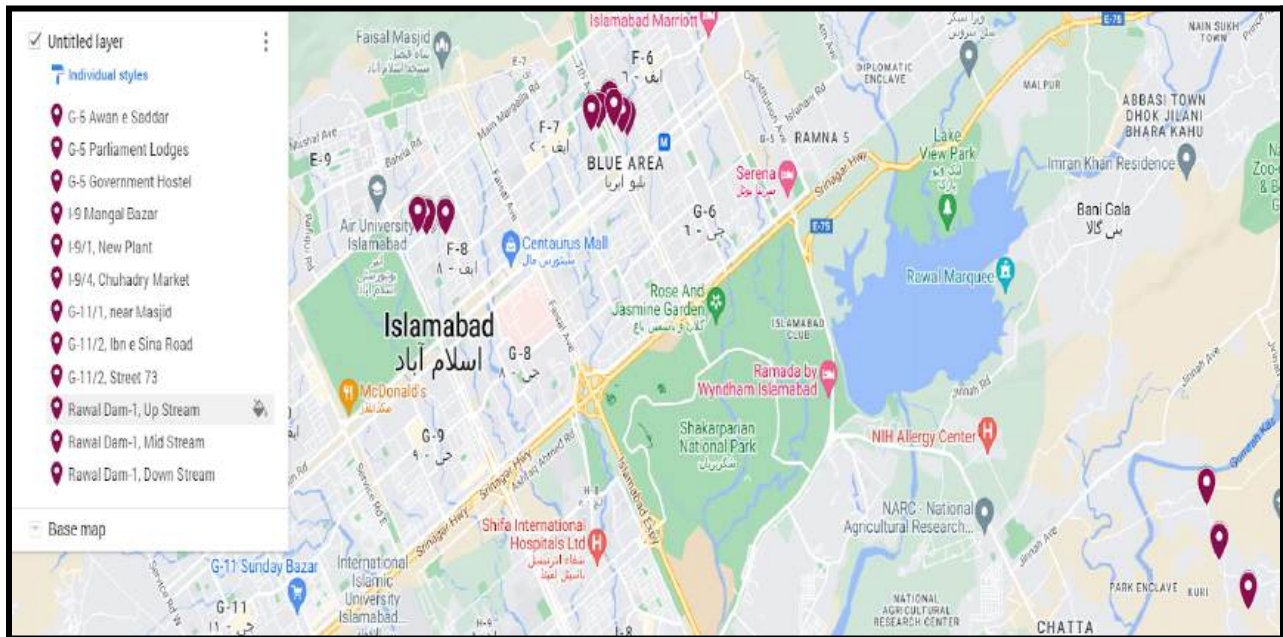
Testing parameters included microbiological analysis for total coliform, fecal coliform, and *E. coli* detection to assess bacterial contamination; trace elements focusing on potentially harmful elements like selenium; and physiochemical parameters testing for pH, turbidity, and chlorine levels. The collection methods utilized were in line with PCRWR recommendations, particularly regarding contamination prevention measures, such as avoiding rusted taps and sanitizing with chlorine.

In addition to these parameters, the methodology also included the use of standard laboratory techniques for analysis. For microbiological testing, samples were cultured on selective media to identify and quantify the presence of coliform bacteria. Physiochemical tests were conducted using calibrated equipment to ensure accuracy in measurements. All analyses were carried out at the PCRWR (Pakistan Council of Research in Water Resources) Laboratory in Islamabad, ensuring compliance with established protocols. The results were then compared

against national and international standards for drinking water quality to assess their safety for human consumption

Furthermore, qualitative data were gathered through surveys and interviews with residents regarding their perceptions of water quality and health impacts. This information provided valuable context to the quantitative findings and helped identify specific areas where community engagement is necessary. Residents expressed concern over the availability of clean drinking water and the prevalence of waterborne diseases, further emphasizing the need for targeted interventions.

The sampling locations can be seen in the map provided below:



## Results

The analysis of water samples revealed significant contamination across the tested locations. In Rawal Dam, turbidity levels exceeded the acceptable limit of 5 Nephelometric Turbidity Units (NTU), with readings of 74 NTU upstream and 46 NTU downstream. Turbidity, a measure of water clarity affected by the presence of suspended particles, can hinder the effectiveness of disinfection processes, making water treatment more challenging and increasing health risks (WHO, 2017).

Total coliform levels were alarmingly high across all sampling points, with the highest concentration found upstream, where coliform levels reached 74 MPN/100 mL—well above acceptable standards. Most Probable Number (MPN) is a statistical estimate of the number of organisms per unit volume of water. The unit MPN/100 mL represents the density of organisms in 100 milliliters of water. The presence of fecal coliform, which was similarly concerning, indicates inadequate sanitation practices in nearby residential areas. The presence of fecal

coliform in drinking water is a strong indicator of contamination with pathogens that can cause severe gastrointestinal diseases (CDC, 2020).

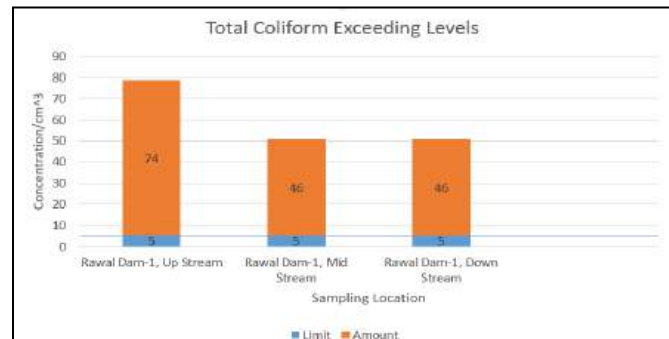
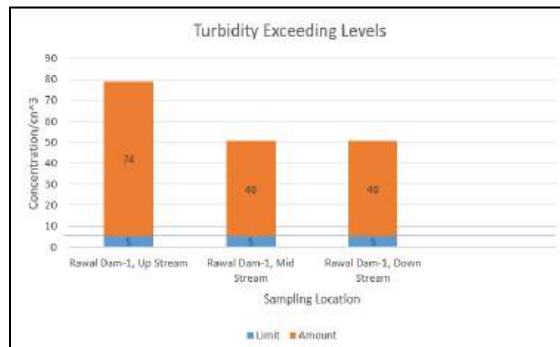
In the groundwater samples taken from G-5 Government Hostel and G-11/1, selenium concentrations were significantly above the safe threshold of 0.12 parts per billion (ppb), with measurements reaching as high as 12.92 ppb. This poses a serious health risk, as prolonged exposure to selenium can lead to chronic health issues, including neurological damage and reproductive problems (WHO, 2019). The unit ppb refers to parts per billion, indicating the concentration of a substance per billion units of volume, typically used to express extremely low concentrations in water.

The results also highlighted differences in water quality between the sampled locations. While some areas, such as G-5 Government Hostel, exhibited high levels of microbiological contamination, others demonstrated more favorable results, albeit still below safe drinking water standards. For instance, groundwater from some tubewells showed lower coliform counts, suggesting varying levels of contamination based on location and proximity to waste sources. This disparity underscores the need for localized assessments and targeted interventions.

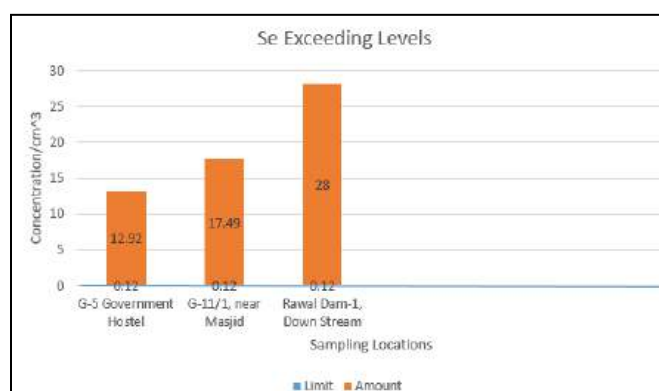
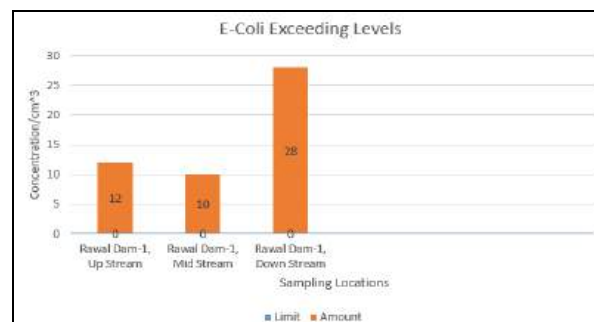
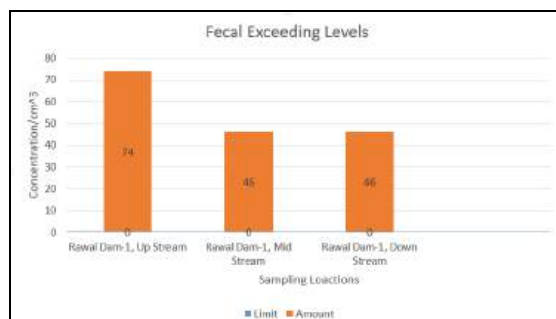
The sampling also indicated variations in the effectiveness of local water treatment facilities. In some areas, water was treated to acceptable standards; however, distribution issues led to contamination post-treatment. Reports from the Pakistan Council of Research in Water Resources (PCRWR) highlight the need for regular maintenance and monitoring of water distribution systems to prevent post-treatment contamination (PCRWR, 2021).

Additionally, the presence of heavy metals in some samples raises further concerns. Lead and arsenic levels were detected in certain areas, although they did not exceed the current regulatory limits set by national and international standards. However, even small amounts of these heavy metals, below enforceable standards, can still pose significant health risks over time. Chronic exposure to lead, even at low levels, has been linked to developmental delays, neurological disorders, and cognitive impairments, particularly in children (WHO, 2018). Similarly, long-term exposure to arsenic, even at concentrations below regulatory limits, can increase the risk of cancers and cardiovascular diseases. These metals are often introduced through industrial activities and agricultural runoff, making vulnerable populations, such as children and the elderly, particularly susceptible to their harmful effects.

The following tables show the parameters that exceeded the amounts required for drinkable water:







## Discussion

The results of this study indicate that the water quality in Islamabad is being compromised by a combination of factors, including industrial runoff, poor waste management, and inadequate water treatment infrastructure. The high levels of turbidity and bacterial contamination in Rawal Dam suggest that surface water sources are particularly vulnerable to pollution, exacerbated by urban expansion and industrial activity in surrounding areas.

The health impacts of water contamination in Islamabad are profound, with diseases like hepatitis E being prevalent among populations reliant on untreated or poorly treated water sources. Poor sanitation practices, combined with limited access to treated water, have created conditions where waterborne diseases can thrive, particularly in informal settlements and densely populated regions. According to the WHO, unsafe drinking water is linked to various diseases, including cholera, dysentery, and typhoid fever, which continue to pose significant public health challenges in Pakistan (WHO, 2022).

Moreover, the public health implications of these findings cannot be overstated. The presence of coliform and fecal contamination in water sources poses immediate health risks, leading to gastrointestinal illnesses and other waterborne diseases. Vulnerable populations, including children and the elderly, are particularly at risk. A comprehensive approach to improving water quality in Islamabad is crucial for safeguarding public health and enhancing the quality of life for residents.

## **Socio-Economic Implications**

Water quality issues not only pose health risks but also have broader socio-economic implications. The economic burden of waterborne diseases on families and the healthcare system can be substantial. Households may spend a significant portion of their income on medical treatment for water-related illnesses, diverting funds from essential needs such as education and food. Moreover, the loss of productivity due to illness can impact local economies, perpetuating cycles of poverty. Studies have shown inadequate water quality and sanitation contribute to lower educational attainment and reduced economic productivity (UNICEF, 2021).

The burden of water quality issues is not evenly distributed across populations. Vulnerable communities, particularly those in low-income areas, are disproportionately affected. This disparity exacerbates existing inequalities and limits opportunities for social and economic mobility. Addressing these disparities requires targeted interventions that focus on the most affected communities, ensuring they have access to safe drinking water and sanitation facilities.

## **Potential Solutions**

Potential solutions to these issues must consider both immediate and long-term approaches. Immediate actions could include the installation of water filtration systems in high-risk areas and enhanced public awareness campaigns regarding safe water practices. Long-term solutions may involve improving waste management infrastructure and implementing stricter regulations on industrial discharges to protect water sources. Community engagement is critical; residents should be involved in water management strategies to ensure that interventions meet their needs.

Investment in infrastructure is essential. Upgrading sewage treatment facilities and expanding access to clean water in underserved communities will help reduce contamination levels. Furthermore, public-private partnerships can play a pivotal role in financing and implementing these initiatives, ensuring sustainable water management practices are in place. A comparative analysis of successful water management practices in other countries can also provide valuable insights for Pakistan.

To improve the overall quality of water in Islamabad, it is vital to implement a multi-faceted approach that involves collaboration among government agencies, local communities, and non-governmental organizations (NGOs). Establishing a centralized water quality monitoring system can help track contamination levels and ensure timely interventions. Additionally, training programs for local communities on safe water handling practices can empower residents to take control of their water supply.

## **Limitations**

While this study provides valuable insights into the water quality issues in Islamabad, several limitations should be acknowledged. One key limitation is the relatively small sample size of 12 water sources, which, while representative of major water points, may not fully capture the variability across the city's diverse geographic and socio-economic regions. A larger

sample size, including additional sampling from peripheral areas and over multiple seasons, would provide a more comprehensive understanding of the water contamination patterns. Additionally, it would have been beneficial to conduct a more detailed analysis of industrial discharge points in the vicinity of water sources to better understand the direct impact of industrial pollution.

Another limitation lies in the reliance on spot samples, which provide a snapshot of the water quality at a given time but may not reflect daily or seasonal variations in contamination levels. Continuous monitoring systems could offer more comprehensive data on water quality over time. Furthermore, expanding the scope to include socio-economic factors influencing water access and contamination levels could make the study stronger. For instance, assessing the financial and technical capacities of water treatment facilities would provide a clearer picture of the infrastructure gaps that need to be addressed.

Lastly, while this study focused on chemical and microbiological contaminants, future studies could also explore emerging contaminants such as pharmaceuticals and personal care products, which are increasingly found in water supplies but remain under-researched in the context of Pakistan. Addressing these limitations would contribute to a more thorough understanding of Islamabad's water quality issues and help refine interventions to ensure long-term public health improvements.

## Conclusion

This study underscores the urgent need for action to improve water quality in Islamabad. While industrial growth and urbanization contribute significantly to the problem, systemic failures in water management and infrastructure maintenance must also be addressed. Comprehensive strategies, including improved water treatment facilities, stricter pollution controls, and public awareness campaigns, are essential for mitigating the ongoing contamination crisis and protecting public health.

In conclusion, local authorities and government agencies must take decisive action to address the water quality issues facing Islamabad. By prioritizing investments in infrastructure and implementing effective regulatory measures, it is possible to create a healthier environment for residents and ensure access to safe drinking water for all. Future research should focus on evaluating the effectiveness of proposed interventions and exploring innovative technologies for water treatment and management.

## Supplementary Information

- [Raw Water Analysis Data from Lab](#)
- [Water Collection Sampling Profile](#)

## Acknowledgments

I would like to express my sincere gratitude to Polygence for providing me with the platform and resources to pursue this research. I am also deeply thankful to the Pakistan Council of Research in Water Resources (PCRWR) for their invaluable guidance and laboratory support throughout this study. Most importantly, I am profoundly grateful to my mentor, John Paul

Angulo, MPH, whose expertise, encouragement, and unwavering support were instrumental in the completion of this paper. His experience brought invaluable insight that greatly enriched this research.

### Works Cited

1. World Bank. "Indus Waters Treaty: History, Provisions, and Impact." World Bank.
2. Pakistan Council of Research in Water Resources (PCRWR). "Drinking Water Quality in Pakistan 2021." PCRWR.
3. Abdullah, Subhan. "The Water Crisis in Islamabad: Causes, Consequences, and Solutions." Earth.Org. Earth.Org.
4. World Health Organization (WHO). "Drinking-Water Quality." WHO, 2022. WHO.
5. Centers for Disease Control and Prevention (CDC). "Water-Related Diseases." CDC, 2020. CDC.
6. World Health Organization (WHO). "Selenium in Drinking-Water." WHO, 2019. WHO.
7. Pakistan Bureau of Statistics. "Population and Housing Census 2021." PBS.
8. United Nations Children's Fund (UNICEF). "Water, Sanitation and Hygiene (WASH)." UNICEF, 2021. UNICEF.
9. Zafar, M. et al. "Impact of Water Quality on Health and Well-Being in Pakistan." *Journal of Environmental Management*, vol. 300, 2021, pp. 113512. ScienceDirect.

## **Music as a Tool in the World of Streaming Platforms By Michael L Wan**

### **ABSTRACT**

This is an ethnomusicological analysis of how music on streaming platforms has impacted listeners' relationship with music and playlists' value to people. It is inspired by the author's personal experience as a prolific music playlists curator with over 4 million plays on an internet music platform. The paper reviews the literature on the role of music, explores the author's journey of incubating influential music playlists using autoethnography, and demonstrates the author's reflections on music's role as a tool for individuals' distinct purposes across different contexts. Both qualitative and quantitative data collection methods are used in this research. The author concludes by examining the versatile impacts of music under the support of streaming platforms.

### **INTRODUCTION**

The number of music streaming platform subscribers has surged dramatically over the past decade, with nearly 616.2 million people using streaming platforms nowadays (Leu, 2024). In 2023, global streaming revenues hit a record high of 19.3 billion U.S. dollars, more than seven times the 2.6 billion recorded in 2015. Streaming now represents over 67 percent of the total revenue from recorded music worldwide (Statista Research Department, 2023). Music streaming platforms not only rapidly ascend in the music industry, but also widely affect individuals' lives. Countless teenagers and adults including me nowadays listen to music almost every day. Frequently, people put their headphones on across various contexts such as working, studying, commuting and other settings. Music is already greatly incorporated into our daily activities while streaming services are making it much easier for people to access music. In particular, the selection and curation of music playlist via streaming platforms represent one of the primary ways people interact with music online. Some platform users attentively select music or playlists with specific functions to support their ongoing activities, such as doing homework or working out, or even to influence their current emotional states. Additionally, some users become curators who develop playlists to generate a new type of digital capital.

This article not only discusses listeners' common uses of music as a tool to address the common needs individuals have in the rapidly evolving society, but also explores people's utilization of music to generate or transform digital capital on streaming platforms. It first presents a literature review on how digital streaming has affected listeners' relationships with music. Then it explains the methodology and the findings, which includes a personal narrative of the author's music playlist curation journey using autoethnography, a survey analysis from 130 respondents, and an interview analysis with a dedicated follower of the author's music playlists. The article concludes with an examination of people's active selection of music to adjust their personal states and music's value as digital capital on streaming platforms.

### **THEORETICAL BACKGROUND: THE ROLE OF MUSIC**

The rise of streaming platforms offers people a more accessible and available way of

engaging with music, which in turn enhances the influence of music in people's daily lives, especially as a background accompaniment to other activities. As musicologist Lee Marshall writes, "In the vast majority of cases in which music is heard, music is not the central focus of attention. Listening to music is most often a complementary activity – people put on music in the background while they are doing something else such as exercising, doing housework and socializing" (Marshall, 2019).

Music itself has become an effective technique which positively impacts people's emotions or their social identities, even as a background supplement. Music could change listeners' emotional status, which has motivated many music streaming platforms including Spotify, Tencent QQ Music, NetEase Cloud Music, among many others, to provide a myriad of music playlists featuring various themes online. For example, the Tencent QQ Music playlist "100% Happiness: Nothing Will Always be Bad" aims to foster an optimistic mood "through providing the rainbow after storms" (Zhang, 2024). Likewise, "Casual Rap to Drive Away Depression" promises people that the playlist will "comfort their souls through Chinese hip-hop" (Niananna, 2024). The projected changes are not only emotional. "Essential Songs for Spinning Bikes" promotes the hopeful idea that people will "lose 10 pounds in a month" (Mengmengdeyediannvwang, 2024).

To understand how streaming platforms bring those possible effects of music upon individuals, Eric Drott suggests to conceive music in terms of technology "not simply as something produced or circulated by means of technology, but as itself a technology" (Drott, 2019). Similarly, music sociologist Tia DeNora writes that music is a technology which serves as "a cultural resource that actors may mobilize for their on-going work of self-construction" (DeNora, 1999). From DeNora's perspective, music acts as a constructor of an individual's emotional status, identity, conduct, social contact, etc. Instead of being passive receptors of music, people proactively select specific music to transform their current emotional states into a preferred alternative.

Music has further expanded its role in recent years as a tool to accumulate "digital capital," "a set of internalized abilities and aptitudes (digital competencies) as well as 'externalized resources' (digital technology) that can be historically accumulated and transferred from one arena to another" (Ragnedda, 2019). Since the 2000s and 2010s, streaming platforms and social media have developed and dramatically expanded. New opportunities appeared for both creators and listeners within the music industry, as numerous ways of promotion and unlimited features to entertain audiences arose across platforms and media (Fountain, 2021). Music playlists are one of the main features of such platforms, which not only promote music they contain but also provide listeners with on-demand streaming. Streaming playlists are selections of recordings sequenced in a particular order, usually compiled and arranged in accordance with a genre, mood, artist, or popularity. Music playlist curators have garnered increased importance as more people are involved in musical discovery processes, especially due to its promotional value (Pellitero, 2018). Based on other scholars' analyses and the author's own experience as an independent playlist curator, listeners' interaction with playlists can result in considerable value

either psychological or economic.

## METHODS, FINDINGS AND ANALYSIS

### **From Zero to 4 Million Plays: My Experience as a Music Playlist Curator**

“It was so funny that after I shared this playlist with my desk mate. He got totally lost in the music and had a great time while writing an essay!”

“What a great playlist! It’s the first playlist which I’ve listened to from the beginning to end, and I really love every song it includes!”

“There’re so many songs that I used to listen to. It seems like I’ve retrieved a treasure from the past!”

These are just a few of the many comments I have received as a music playlists curator. During the past five years, I’ve been curating and sharing playlists which I tailor around various music genres or characteristics on the Tencent QQ Music platform, one of the most popular music streaming platforms in China with over 220 million monthly active users and a total of 800 million users (Tencent Cloud, 2024). My playlists encompass a variety of themes, such as Cantonese Classics, Disney movie soundtracks, and K-pop. Several of my playlists have received considerable attention, with my Electronic Dance Music (EDM) playlist gaining particular popularity of around 3.19 million plays.

With great passion for music since childhood, I originally started off developing playlists with the simplest intention to share the music I love with others in 2019. The playlists were also made accessible to the platform community users, as I was keen to have my musical taste recognized by others and connect with likeminded listeners. I also began to improve my playlists based on listeners’ feedback and publish them on Tencent QQ Music. Based on my observations and analysis of people’s music preferences, both online and in real life, I created my first electronic dance music playlist, whose title, “Ultimate EDM | Sound Lightning Shattering the Dynamic Illusion,” conveyed intensity and power for Tencent QQ Music Platform users. It contained music from a variety of high-quality EDM producers including the Chainsmokers, Martin Garrix, Vicetone, Alan Walker, and others. In addition, I created a playlist which was aimed to generate happy moods in listeners – “Western Pop Music ~ Put Sunshine in Your Pocket” – which embodied various upbeat music from Taylor Swift, Charlie Puth, Sabrina Carpenter, and other western pop singers. Moreover, I curated a playlist of movie soundtracks entitled “Epic Films' Powerful Scores” that contained numerous music soundtracks from great movies like Star Wars, The Lord of the Rings, Harry Potter, etc. Ranging from EDM to Western pop and from tunes that boost happiness to tracks that soothe depression, my original playlists were aimed to provide music collections tailored to diverse listeners’ preferences. I also made adjustments to the playlists based on the music preferences expressed by users on streaming platforms and feedback from previous listeners, ensuring that my selections meet the needs of an even broader audience. My efforts have finally paid off. Since 2019, my playlists have cumulatively reached over 4 million plays and more than 30,000 likes, along with hundreds of positive comments and messages.



### Quantitative and Qualitative Analysis

Such a desirable outcome sparked my curiosity about the factors that motivate people's engagement with music. To gain a holistic understanding in a structural way, I conducted a survey on WeChat which was accessible to individuals of all age groups and successfully collected 130 responses. The majority of the survey respondents fall within the age range of 16 and 50, aligning with the general demographic that constitutes the largest audience for online streaming platforms.

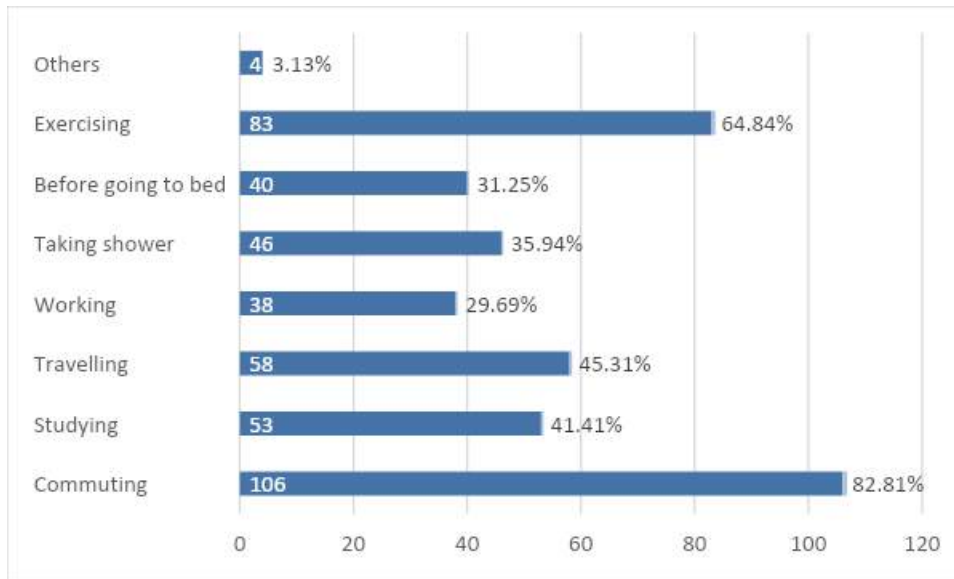


Figure 1 – Contexts in which people listen to music

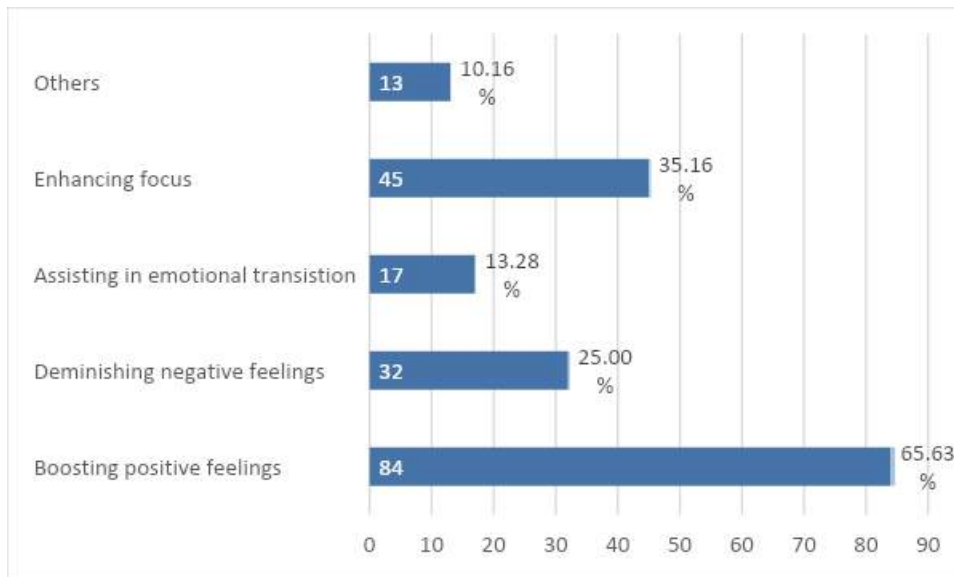


Figure 2 – Reasons people listen to music

Figure 1 above shows that a majority of respondents listen to music while commuting and engaging in sports, with 82.81% and 64.84% of the survey respondents doing so, respectively. Studying and travelling are also major contexts in which music supplements individuals' ongoing activities, with over 40% of the respondents indicating they use music in these settings. However, the role of music extends beyond the aforementioned contexts and is integrated into a variety of other activities, such as working, bathing, and winding down before bedtime, etc. It's evident that listening to music nowadays is not just a standalone activity but often serves as an accompaniment to other tasks, which aligns with Lee Marshall's viewpoint that music is not just a central focus, but exists as an accompaniment or background sound in a variety of activities. In fact, Spotify has launched a new strategy which emphasizes music's supplemental role to other activities, emphasized by a new slogan, "the right music for every moment" (Marshall, 2019).

As shown in Figure 2, 65.63% of the survey respondents use music to boost positive emotions, 25.00% employ it to alleviate negative emotions, and 35.16% utilize it to improve focus attention. Music plays multiple roles, particularly on streaming platforms, where it enhances enjoyment, manages negative emotions, and offers emotional support. This versatility illustrates the diverse ways that music can impact individuals across various contexts and emotional needs.

In addition to the standard questionnaires, I conducted a formal interview with a teenage girl named Isabella from a middle school to understand how her use of music achieves the desired effects. She listens to music on Tencent QQ Music platform and has a particular fondness for EDM. She shared, "There are indeed moments of passion and unique experiences in various situations. Listening to electronic music while studying can significantly improve my efficiency, and at other times, it even offers a boost of positive energy." While it might seem counterintuitive for a genre characterized by intense rhythms and electronic sound effects to facilitate studying, it is important to recognize that individual contexts significantly influence how people engage with music. Individual listeners have also endowed music to suit their own needs, environments, and meanings, thus becoming active initiators in shaping their musical experiences. Specifically, reflexivity constantly shapes the way people interact with music. This is particularly evident within reflexive communities where individuals place themselves into various contexts and continuously contribute their reinventions to the context. Reflexivity is a critical element in an individual's wielding of music, and its resulting product will be cultural and abstract (Kahn-Harris, 2004). In other words, people place themselves into the context of music and then reinterpret the context with their own experience. Listeners are inclined to attach their own experiences, memories, and emotions to the music they listen to. The respondent Isabella is one example of ascribing different functions to music while in different contexts.

Music also serves as an emblem of identity and a source of belonging. Scholars O'Hera and Brown (2006) state that "music choice is tied up with the formation of identity and membership of different groups. Often youth subcultures identify themselves using music as a way of forming and establishing their identity". Isabella also mentioned that she believes EDM is a niche genre

that only a small percentage of people listen to, and she constantly posts her favorite electronic music on social media and enjoys sharing it with others. Isabella said, “I wanted to connect with individuals who share similar music preferences.” Through streaming platforms or social media, she endeavors to find like-minded peers and communities who share her passion for EDM. Social Identity Theory suggests that individuals define their own identities with regard to social groups and that such identifications work to protect and bolster self-identity (Islam, 2014). Similarly, other scholars’ findings have suggested that those who share our musical taste are likely to be regarded as in-group members (Lonsdale, 2020). Isabella’s music sharing activity could be seen as a way of self-expression and an attempt to reinforce her definition of herself, as she stated in the interview, “I wish to share my music preference with the entire world.” Hence, sharing of music taste helps form bonds among the community of EDM listeners and also aids the search for a sense of belonging.

### **Digital Capital and its Ecosystem**

Music fans accumulate “digital capital” not only as skills in combining the production of content and the circulation of meaning by curating music websites, but also as abilities in managing a set of strategies as a result of those processes to occupy certain positions within fields (Zevolli, 2017). Music playlist curators are similar to the aforementioned music websites creators. They have the ability to build audiences and garner downloads, “clicks” and “likes” through their curation efforts. By leveraging streaming services to circulate their playlists, they can elevate their statuses on digital platforms. This enhanced visibility may then be translated into economic capital through advertising opportunities. As an officially recognized playlist curator on Tencent QQ Music platform in China, I have had the privilege of seeing my playlist promoted and enjoyed by millions of plays. In particular, the peak daily plays on my playlists have reached almost 15,000 while monthly plays have soared to around 240,000. These substantially daily and monthly play counts on music streaming platforms have become the capital that attracts independent musicians to seek collaboration with me. They reached out to me with the intention of featuring their original tracks in my playlist and also expressed their willingness to offer financial incentives for inclusion. My digital capital, which encompasses the expertise in curating playlists and the popularity I have gained on Tencent QQ Music Platform, has been gradually accumulated through my engagement with music and playlists curation. The digital capital can be transformed into economic benefits.

To ensure quality control, I would review the music submitted by independent producers before making a decision on collaboration. The play counts, likes, and the number of fans of the producers should be taken into consideration. In particular, the comments left by listeners under their tracks are one of the most crucial elements in assessing the quality and popularity of those producers’ music. I once had to reject an independent EDM producer’s request of collaboration after discovering comments accusing his alleged original music of plagiarism. Moreover, I would only consider including music which corresponds to the characteristics of my playlist. For example, I wouldn’t incorporate wistful or sentimental tracks into my EDM playlist, as it is

designed to evoke the feeling of excitement and passion. To safeguard my digital capital including the quality of my playlists and the accumulation of my followers or listeners, I maintain a very rigorous and prudent review process for independent producer's music.

Nonetheless, digital capital can be transferred between playlist curators and music producers or brands. Through adding top-quality music into outstanding playlists, curators can improve the overall quality of playlists, and in turn, music producers can benefit from the increased promotion of their music. In other words, the exchange of digital capital gradually forms between playlist creators and music producers, where large-scale flows of audience and enduring listener predilection contribute to high-quality music, and vice versa. The ecosystem within music streaming platforms is thus formed, constituted by reciprocating parties, such as playlist curators and music producers, who possess digital capital which is beneficial to all of them. Particularly, only with the combination of music and streaming platforms, can this reciprocation succeed. Streaming platforms like Tencent QQ Music provide people with essential online conditions to utilize music as a technique for digital capital accumulation, either through playlists curation or music production.

## CONCLUSION

People now use music as a tool to accomplish specific purposes across various contexts in their daily lives. The expanded function of music, which includes enhancing mood, increasing work efficiency, or simply creating a background sonic supplement, is partly a consequence of the widespread adoption of streaming platforms. The convenience and accessibility of music brought by streaming services have enabled the mode of "ubiquitous listening" characterized by Kassabian (2013): "a mode of listening dissociated from specific generic characteristics of the music," through which people inattentively listen to music as a background accompaniment to their routines and activities (Marshall, 2019). In addition, music and playlists can also serve as a media for accumulating digital capital, thereby enhancing music's economic value. Playlist curators' competence in creating or managing playlists, as well as their fan base or overall flow of listeners, stems from the synergy between music and streaming services.

The role of music and the ways it functions are continuously expanding in the era of streaming platforms. Sometimes, the actual role of music may align with people's expectations. Individuals' perceptions of the same piece of music can vary significantly due to their unique personal contexts. With the prevalence of streaming services, listening to music can now be incorporated into almost every setting. People have the ability and tendency to bestow new functions to music as they attach their own needs and interoperations upon it. Music can serve as a support to an individual's social life, a source of economic profit, a mood enhancer, or any other tools, depending on people's intentions and needs. Hence, people have become proactive wielders of music as a tool, with its function largely depending on their subjective desires.

## Work Cited

- DeNora, Tia. "Music as a Technology of The Self." University of Exeter, UK, 1999. Drott, Eric. "Music in the Work of Social Reproduction." University of Texas at Austin, 2019, pp. 11-13.
- Fountain, Tony M. "The Evolution of The Music Industry — And What It Means for Marketing Yourself as A Musician." *Forbes*, 13 Sept. 2021.  
<https://www.forbes.com/councils/forbesbusinesscouncil/2021/09/13/the-evolution-of-the-music-industry-and-what-it-means-for-marketing-yourself-as-a-musician/>. Accessed October 14, 2024.
- Hracs, Brian J., Michael Seman, and Tarek E. Virani, eds. *The Production and Consumption of Music in the Digital Age*. Routledge, 2017.
- Islam, Gazi. "Social Identity Theory." ResearchGate, 2014.  
[https://www.researchgate.net/publication/281208338\\_Social\\_Identity\\_Theory](https://www.researchgate.net/publication/281208338_Social_Identity_Theory). Accessed October 14, 2024.
- Kahn-Harris, Keith. "The 'Failure' of Youth Culture. Reflexivity, Music and Politics in the Blackmetal Scene." *European Journal of Cultural Studies*, 2004.  
[https://www.researchgate.net/publication/27225629\\_The\\_'Failure'\\_of\\_Youth\\_Culture\\_Reflexivity\\_Music\\_and\\_Politics\\_in\\_the\\_Black\\_Metal\\_Scene](https://www.researchgate.net/publication/27225629_The_'Failure'_of_Youth_Culture_Reflexivity_Music_and_Politics_in_the_Black_Metal_Scene). Accessed October 14, 2024.
- Kassabian, Anahid. "Ubiquitous Listening: Affect, Attention, and Distributed Subjectivity." University of California Press, 2013.  
<https://doi.org/10.1525/california/9780520275157.001.0001>. Accessed October 14, 2024.
- Leu, Patrick. "Music Streaming Worldwide - Statistics & Facts." Statista.  
<https://www.statista.com/topics/6408/music-streaming/#topicOverview>. Accessed October 14, 2024.
- Lonsdale, Adam J. "Musical Taste, In-group Favoritism, and Social Identity Theory: Re-testing the Predictions of the Self-esteem Hypothesis." Vol. 49, No. 4, 2020.
- Marshall, Lee. "Do People Value Recorded Music?" Vol. 13 No. 2, University of Bristol, UK, 2019, pp. 151-152.
- Mengmengdeyediannvwang. "Essential Songs for Spinning Bikes." QQ Music,  
<https://y.qq.com/n/ryqq/playlist/2374456197>. Accessed October 14, 2024.
- Niananna. "Casual Rap to Drive Away Depression." QQ Music,  
<https://y.qq.com/n/ryqq/playlist/6927234768>. Accessed October 14, 2024.
- O'Hera, Kenton, and Barry Brown. "Consuming Music Together: Social and Collaborative Aspects of Music Consumption Technologies." 2006.  
[https://www.researchgate.net/publication/321596161\\_Consuming\\_Music\\_Together\\_Social\\_and\\_Collaborative\\_Aspects\\_of\\_Music\\_Consumption\\_Technologies](https://www.researchgate.net/publication/321596161_Consuming_Music_Together_Social_and_Collaborative_Aspects_of_Music_Consumption_Technologies). Accessed October 14, 2024.
- Pellitero, Miguel. "Streaming Playlists - What is Their Relationship with Music Culture in the UK." ResearchGate, 2018, p. 14.  
[https://www.researchgate.net/publication/327629221\\_Streaming\\_Playlists\\_-\\_What\\_is\\_their](https://www.researchgate.net/publication/327629221_Streaming_Playlists_-_What_is_their)

- [\\_relationship\\_with\\_Music\\_Culture\\_in\\_the\\_UK](#). Accessed October 14, 2024.
- Ragnedda, Massimo, Maria Laura Ruiu, and Felice Addeo. "Measuring Digital Capital: An Empirical Investigation." Northumbria University, UK; Università degli Studi di Salerno, Italy, 2019, p. 2.
- Statista Research Department. "Global Music Streaming Revenue 2023." Statista, <https://www.statista.com/statistics/587216/music-streaming-revenue/>. Accessed October 14, 2024.
- Tencent Cloud. "Background: QQ Music's Three-Dimensional Entertainment Ecosystem." Tencent Cloud, <https://www.tencentcloud.com/customers/detail/2899#:~:text=Background.for%20over%20800%20million%20users>. Accessed October 14, 2024.
- Zevolli, Giuseppe. "The Production and Consumption of Music in the Digital Age." Edited by Brian J. Hraes, Michael Seman, and Tarek E. Virani, Routledge, 2017. [https://www.researchgate.net/publication/317664410\\_The\\_Production\\_and\\_Consumption\\_of\\_Music\\_in\\_the\\_Digital\\_Age\\_Edited\\_by\\_Brian\\_J\\_Hraes\\_Michael\\_Seman\\_and\\_Tarek\\_E\\_Virani\\_Abingdon\\_Routledge\\_2016\\_277\\_pp\\_ISBN\\_978-1138851658](https://www.researchgate.net/publication/317664410_The_Production_and_Consumption_of_Music_in_the_Digital_Age_Edited_by_Brian_J_Hraes_Michael_Seman_and_Tarek_E_Virani_Abingdon_Routledge_2016_277_pp_ISBN_978-1138851658). Accessed October 14, 2024.
- Zhang, Yaqing. "100% Happiness: Nothing Will Always be Bad." QQ Music, <https://y.qq.com/n/ryqq/playlist/8317709780>. Accessed October 14, 2024.

## **Which Has a Bigger Effect on History: the Plans of the Powerful or Their Mistakes? Understanding Historical Impacts: Mistakes vs. Planning By Zelai Li**

It is without question that during major historical episodes, certain decisions made by states or individuals significantly affected the outcome of the entire event. Out of those decisions, some were carefully planned, and some were fatal mistakes that caught all parties off guard. While recognizing both categories has had powerful impact on the outcome of history, this essay will argue that the mistakes of the decision makers have undertaken a more decisive role in the course of history through analyzing series of events that led to and happened during WW1 and WW2. This is because mistakes are often overlooked by the initiators, and this rapidly changes the situation. While plans, however significant are expected by some, therefore does not have the capacity to drastically change the environment of an event.

### **A Cascade of Misjudgements: The Unforeseen Start of World War I**

To begin with, this essay will look back on one of the most horrific tragedy in human history, the First World War from 1914 to 1918. The destructions it has caused were unprecedented and unparalleled to anything that came before. Yet, some historians argue that WW1 was started by a series of mistakes and misunderstanding caused by poor communication between European Powers. Most significantly, the initiator of WW1 – Austria Hungary, never wanted a general European war. Their intentions were to keep the fight in the Balkan regions and keep Serbia quiet to preserve the Bosnian province. (Shen 2) Since the aim of the Austrians were to simply keep their empire as it is, they had no reason to start a European War which then eventually cost them sixty percent of their territory. (Taylor and Francis) It was the Austrian's emperor's mistake of consulting Germany before punishing Serbia that led to the mobilization of multiple European powers, when the Germans sent a blank cheque symbolizing unlimited support, its oppositions such as Russia and France saw it as a threat and prepared for war. They didn't have a choice as well, if Germany didn't back Austria, they risked losing their most powerful ally and become more encircled by its enemies, if Russia didn't mobilize to support Serbia the Tsar would seem weak and vulnerable as he failed to protect a fellow Slav state. ("IB Histroy:WWI-Long" 3:48)  
)France was an ally of Russia and were bound to mobilize, not to mention France always seeked revenge to Germany for the Alsace Lorraine region lost in the Franco Prussian war in 1871.(Rowe) Therefore, when Austria made that call to Germany, they had already started WW1 without planning it, ironically Austria had way more advanced military than Serbia and they were more than capable of punishing Serbia by themselves.

### **German strategic miscalculations in WW1 and their consequences**

Moreover, during the practices of WW1, mistakes are also imperative to the standings of the European powers. The German Schlieffen plan was intended to keep the war on one front in the early stages and it would benefit Germany a lot. However, their mistake begins when they

overlook Britain's resolve to honor their alliance with Belgium that was established during the Treaty of London in 1839. When German troops set foot on Belgian territory, Britain mobilized and became a strong addition to the allies. With the Russians mobilizing way faster than the Germans anticipated in their Schlieffen plan, they now face a two front war that they didn't prepare for. (Rogers and Thomas 57)

Furthermore, Germany initiated the Unlimited U-Boat Warfare in the later stages of the war, which was using submarines to hunt down any ships in Allied waters. One of the German U-boats accidentally sunk an American tourist ship, Lusitania. (Rogers and Thomas 68) Germany did not seek to set ties with America, as the Americans sold weapons to them and remained neutral throughout the war. Yet, the mistake of sinking Lusitania was part of the reason why American joined the war on the side of the Allies. The mistaken calculation of their opponents in the Schlieffen plan forced Germany into a awkward position, followed by the mistake of sinking an American tourist ship made Germany lose control of the situation. At the end the Germans faced almost every Western Power, and it was simply too much for them which cost them the war.

### **Prime Minister Chamberlain's Fatal Misjudgment and the Dawn of World War II**

On the other hand, the start of WW2 in 1939 was also arguably caused by Prime Minister Chamberlain's misjudgment on the situation in Europe. His appeasement policy not only gave Hitler the confidence to expand German Militarism, but it had also sacrificed Sovereignty and interests of its Allies. Hitler didn't have an army strong enough to initiate another European war, however the appeasement policy gave Hitler the opportunity to enter Rhineland and remilitarize it. With the resources and the morale they gained from this operation, the German military grew day by day. Without the appeasement policy, France could've easily stopped German entrance to the Rhineland as the French armed forces were way stronger than the Germans during this time. (Roger and Thomas 83) The final and most important impact of the Appeasement policy is the Munich agreement, where Britain and France betrayed their democratic ally Czechoslovakia by granting Hitler the Annexation of the Sudetenland. The Sudetenland is an integral part of Czechoslovakia, not only did it produce almost half of Czechoslovakia's GDP and war equipment, but also a quarter of its population. Without the Sudetenland, Czechoslovakia is defenseless and doesn't stand a chance against a full invasion of Germany. (Roger and Thomas 86) Nevertheless, they hoped that Hitler's ambition would stop there, however, resembling past events, Hitler broke his promise by annexing the whole of Cezcho Slovakia. At this point, Britain and France had no choice but to put an end to Appeasement policy. The mistakes of the appeasement policy were severe and significant, avoiding conflict and arguments with Germany only boosted Hitler's ambitions and gave Germany the chance to expand and rearm. (Taylor and Francis) Without the appeasement policy, it was difficult for Germany to regain military strength of any kind.

### **Japanese errors in WW2 leads to Axis defeat**



Similarly, the mistakes within practices of WW2 had also significantly affected the outcome of the war. In December 1941, Japan launched a surprise air strike to Pearl Harbor, which again dragged the United States in the war on the opposing side of the Axis powers. President Roosevelt did want to stay neutral, yet this bombardment gave them no choice. (Imperial War Museum) Japan's gamble with Pearl Harbor was that they would leave the Americans in shock and force them to negotiate with Japan instead of going into total war. However, this gamble did not pay off, as the Americans declared total war on Japan and its allies. (Imperial War Museum) The American involvement eventually led to the nuclear attacks on Hiroshima and Nagasaki, these were the only recorded occasions when nuclear weapons were used in war and their strength was unparalleled to any weapon Japan and the rest of the world had seen before. These attacks forced Japan to surrender and claim victory for the Allies. It's safe to say that the mistake Japan made when they bombed Pearl Harbor was the major turning point of WW2.

### **Impact of Planning: Successes from the Chinese Civil War and WW1**

With that being said, it's also imperative that we recognize some plans had also worked out in history even though they can't always compete with the element of surprise and the drastic effect mistakes hold. For instance, during the Chinese Civil War, the Communists had way less advanced weaponry and they were significantly outnumbered by the Nationalists. Mao Ze Dong's military strategy, the Guerrilla Warfare was a significant factor that eventually won the war for the Communists. Unlike the traditional Trench Warfare, the Guerrilla tactics needed the soldiers to always be on the move. During the battles of Shanxi Province, Communists divided themselves into small infantry units, with their knowledge of the terrain, they always caught Nationalists soldier by surprise. (Shiu and Cheng) Also, the Communists party was way more disciplined as they treated the local people with respect while the Nationalists behaved tyrannical. (Price and Senes 136) This gave the Communists an edge during the Guerrilla warfare as the civilians would often transport food and resources to the CCP while misleading the Nationalists by giving them false information.

Furthermore, the British Convoy system during WW1 was also significant in defending British Merchant ships against the German Unlimited U-Boat warfare. In the past submarine battles, it is a common courtesy for the submarines to emerge from water to give the sailors on the boat they're about to sink a chance of escape before they fire. (Roger and Thomas 58) However, the Germans were way too desperate to consider this option, and many sailors and cargos were lost. The British then came up with the convoy system that would send battleships from the Royal Navy to accompany merchant ships entering the Allied waters and this significantly reduced U-Boat attacks. It is worthy to note that the British had only 3 weeks of supplies left with all the U-Boat attacks, and this planned worked out successfully for Britain's favor.

### **Mistakes vs. Plans in History: A Comparative Assessment**

In conclusion, despite certain plans had worked out in major historical battles, it doesn't have the dramatic effects that mistakes often create. While recognizing that both had been impactful throughout history, the two biggest international conflict ever were arguably started by mistakes and ended by mistakes as well. Therefore, no matter how successful plans had been in regional conflicts, they can't alter the situation in major historical events like mistakes do.

## Works Cited

- History Channel*. 3 May 2018, Was Germany Doomed in World War I by the Schlieffen Plan? Accessed 16 June 2023.
- "IB Histroy: WWI-Long Term Causes." *Youtube*, June 2021, [www.youtube.com/watch?v=m0J9OrIYao](http://www.youtube.com/watch?v=m0J9OrIYao). Accessed 16 June 2023.
- Imperial War Museum*. June 2022, [www.iwm.org.uk/history/why-did-japan-attack-pearl-harbor](http://www.iwm.org.uk/history/why-did-japan-attack-pearl-harbor). Accessed 16 June 2023.
- Kahn, David. *Vol. 70, No. 5, America and the Pacific, 1941-1991 (Winter, 1991)*. E-book ed., Council of Foreign Affairs.
- Price, Eunice, and Daniela Senes. *World History: Authoritarian States*. E-book ed., Pearson Baccalaureate.
- "The Reason Japan Attacked Pearl Harbor." *Youtube*, June 2022, [www.youtube.com/watch?v=so4v\\_2zq35k](http://www.youtube.com/watch?v=so4v_2zq35k). Accessed 16 June 2023.
- Rogers, Keely, and Jo Thomas. *Pearson Baccalaureate*. E-book ed.
- Rowe, Michael, Dr. "The Franco-Prussian War 150 Years On: A Conflict That Shaped the Modern State." *King's College London*, 17 July 2020, [www.kcl.ac.uk/the-franco-prussian-war-150-years-on](http://www.kcl.ac.uk/the-franco-prussian-war-150-years-on). Accessed 16 June 2023.
- Shen, Elana. "Causes of World War I." pdf.
- Shiu, Victor, and Chiang Cheng. "Imagining China's Madrid in Manchuria: The Communist Military Strategy at the Onset of the Chinese Civil War, 1945-1946." *Sage Journals*, Jan. 2005, [journals.sagepub.com/doi/abs/10.1177/0097700404270549?journalCode=mexa](http://journals.sagepub.com/doi/abs/10.1177/0097700404270549?journalCode=mexa). Accessed 26 June 2023.
- Taylor and Francis Online*. 11 Mar. 2015, [www.tandfonline.com/doi/abs/10.1080/09636412.2015.1001216](http://www.tandfonline.com/doi/abs/10.1080/09636412.2015.1001216). Accessed 26 June 2023.
- Taylor and Francis Online*. 11 Dec. 2020, [www.tandfonline.com/doi/abs/10.1080/09557571.2020.1855111](http://www.tandfonline.com/doi/abs/10.1080/09557571.2020.1855111). Accessed 26 June 2023.

Aryan Agarwal

30.09.24

### **Upper-Echelon Theory - The Impact of CEO Socio-ethical Backgrounds on Profits and Business Success.**

In the modern business market, business leaders are faced with a host of diverse challenges. These challenges are burdened with a substantial amount of motivation, which can in turn muddle decision-making systems. To cope with these challenges, senior management officers must regulate their resources, understand the current events, and respond accordingly in ways that will optimize corporate results. Nonetheless, there is a breach in the research on the manner in which senior management regulates their decision-making processes in critical business decisions. The constantly changing social landscape places an extreme level of pressure and fear in many senior management officers. Company leaders are engaged in a whirlwind of tasks, assess and identify information, and deal with far more incidents than they can sufficiently and effectively process. Almost all decisions are made within strict deadlines, needing quick and often ill-processed decisions, adversely affecting its validity and credibility.

This can be overpowering for top-level corporate decision-makers. In this light, it is important to comprehend the intellectual processes engaged by top-level managers to course through such environments while addressing the various stimuli that they come into contact with each day. Critical leadership addresses the point with captaincy actions at the highest levels of the companies. Critical leadership can be understood as the functions implemented by CEOs and company directors that have immediate consequences on company operations and profitability. Senior management officers have a distended effect on company productivity, and that influence

is steadily growing. Owing to their significant role in daily corporate operations, there is a theory that has been assigned with the task of understanding the manner in which leaders interpret and respond to critical business decisions—the “Upper Echelon Theory.”

The “Upper Echelon Theory” has been identified as one of the most persuasive concepts in “strategic management literature.” The basic premise of the theory is that a leader’s beliefs, intellect, and personalities mold the manner in which he/she will interpret the situations and resolve these issues. The postulate uses two fundamental methodologies, with studies engaging the role of the corporate leader, such as the CEO, and the other focusing on the role of the senior management team as a whole. The study here is significant since it studies the approaches with which is research that posits the point that the theory is a better predictor of business growth. CEOs will interpret a specific business situation and how they will respond to these situations. There is research that posits the point that the theory is a better predictor of characteristics, Withal, much of the literature centers on the contribution of the CEO’s basic characteristics such as their intellectual level, social relationships, and educational attainment.

Even with the prevailing literature, there is a recognized breach surrounding the CEO’s “black box.” A study conducted by Neely and his colleagues (2020) established the existence of this breach and posited that the gap restricts the conceptual and tangible benefits of the theory. The question posited here is: how do executives understand, process, and resolve different business events given their varying levels of intellect and experience? In this light, there is a need to deconstruct the two primary elements of the theory—the ability of the CEO to understand and process the event and the resolution of the issue given these features of the CEO. The ability of the CEO to understand and process an event is a critical element of the theory,

since it defines the process by which the CEO will assign importance and value to these unfolding events and how the CEO will respond to these events given his understanding and interpretation of the events.

In highly fluctuating and dynamic environments, the theory is labeled as an important tool that is distinctively tailored for CEOs to help them resolve a new and even a previously unknown event. Zhang and Soergel (2020) suggest that the theory can be likened to learning since the process follows the system of building upon previous knowledge. The identification and processing procedure traditionally begins when the leader faces an unexpected event or is subjected to an event in flux. The process can be impacted by factors such as intellect, education, technology, and event importance.

Even with the prevailing literature sets, the actual identification and assessment process is still not completely understood. Analysts continue to disagree on the exact parameters of the process, and the more stringent template is inclined to several identification factors: addressing emotions, establishing the elements, asking for external assistance, assessing one's own and other's judgments, expecting the effects of these actions, studying one's personal motivations, and assessing the effects of these actions on others (Hart 18-20).

Corporate captaincy, regardless of industry, company size, or form, is seen as a critical element in the growth and profitability of a company. In today's business environment, marked with increasing competitive pressures, technological shifts, and changes in legal paradigms, the role of a company's leadership structure, particularly its Chief Executive Officer (CEO), in molding and directing the company is a vital element in this discussion. CEOs hold the primary office in the company and, in many instances, are considered the "face of the company." In this

light, regardless of the sector where the company operates, its size, or business paradigm, the leverage power within the company solely rests in the CEO's hands.

Given this crucial observation, academics and researchers have fueled a growing interest in the role of the CEO in the development and competitiveness of their company. Berson (2008) argues that one of the basic approaches in assessing this relationship is that the leadership strategy adopted by the company's CEO will reflect the values he/she holds. CEOs will significantly impact their company's strategic policies and direction by the values and beliefs they hold on to. Also, the selection process that will be adopted by the firm will be impacted by the preferences of the senior managers and directors of the company. The prevailing literature strongly supports the point that the CEO's beliefs and principles, as seen in their mental, geographical, and job-related requisites, will greatly impact the formulation and development of critical policies, such as capital expenditures and corporate budgeting (Plockinger, Aschauer, Hiebl, and Rohatschek).

With the personal traits that are commonly perceived as critical in developing the decisions, such as age, education, and beliefs seen as integral in the study, researchers have also found other traits that can influence a CEO's decision-making process, such as tenure, skills, and international exposure. One of the basic issues in strategic planning is understanding why companies vary in their policy directions and why companies succeed and fail with these decisions. Communities, whether the area where the CEO is located or where the company operates, will greatly influence the characteristics of the CEO. In this context, it is hard to conceive of situations where the person will come to decisions outside of the influences of the community's parameters. In this light, it must be stated that community elements must not be

construed as social forces imposed by environmental norms. It is the assimilation of specific values within that community by the CEO. To research the effect of community on the beliefs of the individual, there is a need to study the policies that companies make and how these decisions impact their respective communities.

It is posited that the company will spend more in their communities rather than in other locations. For example, Saxenian (1994) identifies the power of the community on the policymaking decisions of a business. For example, if the company is located in a highly urbanized area such as Silicon Valley, it will have an impact on how the company's operations will affect the surrounding community. Machinery and heavy equipment manufacturer John Deere, located on the Iowa-Illinois border, invested significant amounts of money, sponsoring numerous community events and building public facilities such as museums. Nonetheless, this act must not be treated as the effect that a company's location, or its surrounding community, has on its decision-making on whether to invest in the community. Rather, it shows the community principles and tradition that in turn become meshed with those of the CEO and other senior management team members. In terms of family-owned companies, the community element is a major factor in these companies. However, this has been rarely studied in the context of publicly listed firms. It can be argued that the community factor posts a negligible or even non-existent effect on the decision-making and outcomes of large companies and has numerous satellites and hubs.

Withal, this holding cannot be treated as a universal truth, as there are companies with several hubs where the management teams are deployed. Studies have shown that many listed companies are family-operated or owned, and the community plays a significant factor in their



operational dynamics. Research has shown that there is a connection between the understanding of family-owned companies and the operations and corporate profitability and the manner in which community values shape the operating principles of these companies. Lester and Cannella (2006) show how these connections between family-owned companies and communities produce “community-level social capital.” This resource is driven by the link between the company and the community and the latter’s values, norms, principles, beliefs, and holdings being shared with the company (Quttainah 35-37).

### **Upper Echelon Theory**

#### *Significance of the study*

In a similar vein, the “Upper Echelons Theory,” as posited by Hambrick and Mason (1984), suggests that companies reflect the values and beliefs of their senior management, and the output is impacted by the knowledge and background, and the performance is impacted in turn by the beliefs and experiences of the company’s decision-makers. Here, since the CEOs are considered important members of the senior management team, their decisions will impact the policymaking process of the board and will eventually reflect the values of the company that they head. Withal, there is also research that argues that the beliefs and values of the CEO do not significantly affect corporate performance since the output and competitive level of the company are determined by industry-related factors or elements related to company operations (Mukherjee, Sen 1-2). Withal, the overall literature on the subject supports the operation of the “Upper Echelons Theory” in that managerial nuances regarding critical choices and corporate output are impacted by the CEO choice (Plockinger, Aschauer, Hiebl, and Rohatschek).

People should be seen as a critical element in understanding business policy decisions and the outcome of these decisions owing to individual differences. There is a substantial body of research supporting this position. Teece (2007) recognized the element in the concept of “entrepreneurial managerial capitalism” as a key element for a company to deliver maximum output. Penrose (1995) points to CEOs in the center of the study by arguing that leaders are the most significant corporate resource, both human and nonhuman. Felin and Hesterly (2007) support that position by arguing that “managerial human capital” is the most important asset that companies can engage.

However, though it would be incorrect to proffer that analysts have completely abandoned the contributions of the individual members of these companies, efforts to explain corporate outcomes on a “micro-level” basis are the exception rather than the rule of studies in the area. This is possible due to the hesitation with which researchers have approached the issue of bridging the macro- and micro-levels of the corporate structure.

Given this dilemma, there is a breach in the area of how to engage people as a beginning point in understanding corporate dynamics. In this light, it can be posited that the behavior occurs within the company. While the element of community has been used to replace intellect and management experience, the studies using this element have failed to produce replicable results. Nevertheless, though a number of analysts have emphasized the importance of assigning the intellectual element at the center of the discussion, only a few researchers have called for the deportation of community influences for mental elements in understanding critical decision-making and its outcomes. Here, though community influence has not been removed completely

as a research point, it must be emphasized that the factor not be accorded undue significance in the research (Devine).

There are numerous reasons that validate the need for the study. Aside from the growing amount of literature discussing the theory, there are also indications that its parameters are changing. There is substantial evidence that suggests that executives have a greater effect on corporate profitability and direction compared to their effect in the past. In this light, with regards to direction in the study and significance, it is expected that the literature on the “Upper Echelon Theory” will grow, exacerbating efforts to assess concerns that embrace various research efforts. In the same breath, the environments surrounding the corporate world are shifting in ways that impact the part of business leaders and their influence on the business world and society in general. Hambrick argues that these developments create unique possibilities for researchers—to suggest to leaders confronting a changing corporate landscape and to exploit prevailing literature on the theory that will in turn help in resolving larger questions.

Withal, more fundamental issues are also being discussed in the debates, such as accountability, duplicability, and appropriateness of social science studies, including management studies. Though the applicability of concepts of the “Upper Echelon Theory” is growing, so are the challenges that hinder the conduct of a strong, credible study regarding the issue. This underscores the significance of the need to resolve the critiques seen in the field. From its introduction, the theory has been subjected to several reviews and assessments that have supported the operation of the postulate (Neely, Lovelace, Cowen, and Hiller 4-5).

The “Upper Echelons Theory” supports the position that a CEO’s characteristics will significantly affect a company’s strategic decision-making process, including the use of a

company's competitive power within its industry and within the business market. Since the CEO occupies the highest position in the company hierarchy, the CEO has the power to direct the company's critical and daily operations. There are several elements that can be attributed to the connection between the age group of the CEO and the manner in which they approach, understand, and address critical corporate issues. One, it is posited that older managers will not have the same stamina as younger executives. On the upside, older managers will be more experienced compared to their younger counterparts. This group is also seen to have a higher loyalty level and a greater sense of acceptance in the company. Here, rather than lose the company, they would be more tolerant when considering negotiating debt strategies for their companies rather than opting to open their companies to new entities (Setiawan, Harymawan, Adhriani, Pratama, and Santoso).

The theory addresses the breach by discussing the elements of the "Carnegie School of Thought" and the "Bounded Rationality Theory." The latter points to the finding that critical business decisions have complicated information in them; in this context, developing and arriving at a feasible conclusion. Though companies strenuously work to be logical and prudent in their choices by using their internal tools and mechanisms and reflective of market movements, the "bounded rationality theory" recognizes the fact that senior management officers have innate intellectual constraints—limitations in their knowledge and statistical capacity—that will constrict their capacity to secure a "technical rationality" with their decisions.

Strategic instances are only "interpretable" and not "knowable," and critical business decisions are the outcome of dispositional elements rather than an automated process for business profitability. CEOs and senior management officers will instinctively depend on

previous practices, take intellectual "shortcuts," and imprint their personal views and biases on policy development and implementation. In this context, the outcome will depend on the ability and resources available to the CEO in resolving an "actual" event. Here, the decision-making process adopted is extremely subjective since these are based on the CEO's personal prejudices and biases and intellectual resources. Here, the "Upper Echelons Theory" suggests that a business' critical decision making processes and the resulting outcome of these decisions are impacted by the nuances and beliefs of the company's senior management team (Bekos, Chari).

The theory posits that critical decisions such as guiding and setting down the overall direction of the company are determined by the CEO's age, educational background, and even nationality. These statistical characteristics expose instances that develop a knowledge pool of experiences and principles that will help CEOs interpret and analyze events. For instance, CEOs who were raised in urban areas have known many people from diverse ethnic and cultural backgrounds with the attendant skill sets. In this context, these CEOs are more accommodating to decisions to hire individuals from different cultural or ethnic backgrounds.

The "Resource Dependency Theory" connects with the "Upper Echelons Theory" in that it suggests a connection between a CEO's traits and the assets and a company's requirements. The theory, posited by Pfeffer and Salanick (1978), holds that CEOs are selected based on their educational background, skill sets, and potential. These traits are then matched to the assets and requirements of the company; in turn, it is expected that this will help in effectively operating the company and developing sound working relationships among the management and workers (El Abiad, Abraham, El-Chaarani, Skaf, Binsaddig, and Jafar).

Specific personal traits will affect the way that senior management officials will understand, interpret, and act to resolve issues faced by their companies. However, it must be stated here that CEOs and senior management officers will not be able to fully understand and completely interpret all the data being presented to them. Instead, they will interpret and frame the situations according to their limited cognitive abilities. Moreover, they will move to filter the information as they understand it by using an interpretation tool that is biased by their beliefs, personal experiences, and characteristics. These unique traits—age, gender, and ethnicity—affect their intellectual prejudices and are reflected in the way that these formulate the solutions to their problems and ultimately in the manner that these policies are applied.

In the business world, the CEO is considered the primary decision-making party and, as such, the one ultimately responsible for the outcome of such decisions. The CEO is the iconic leader of the company, with his/her decisions supporting or changing the perceptions of the stakeholders regarding the company's present state and future direction. Given this holding, CEOs have a significant influence on the dynamics of the company. The theory supports the principle that managerial decision-making practices are developed from environmental considerations and are affected by societal norms and traditions. Businesses and their leaders are influenced by a network of accepted and anecdotal rules of the corporate world where they conduct their operations.

## **Gender**

Much of the information addressing the purported connection of a CEO's socioeconomic state is samples of their personal values and their impact on the output of their companies. Recent literature has sought to delve into a CEO's attending personality characteristics, such as their ability to sympathize, benevolence, and egotism, and their effect on his decision-making process. Previous corporate literature purports that gender is a critical element affecting the CEO's understanding, experience, and interplay with others. Melis and Nawaz (2023) aver that female CEOs are better at grasping a holistic view of an issue compared to their male counterparts, are more relational, and have a higher capacity to muster support for specific issues from a diverse range of shareholders. In non-profit organizations, the CEO must be able to develop and maintain cordial relationships with various stakeholders and still maintain the group's identity and their range of services. It was also noted that the innate traits of female CEOs make them better financial stewards in philanthropic groups. Nevertheless, it is also this seeming lack of aggressiveness and low self-confidence that can be detrimental to raising support.

Studies on philanthropic groups with female CEOs are seen to have a lower donation rate compared to groups headed by men. Though the capacity of female CEOs to secure donations is not connected to their gender per se, their abilities are unnecessarily tainted by the prevailing patriarchal socio-cultural environment. This social misogynistic environment results in the unneeded development of a gender bias on the ability of women corporate leaders to lead their respective companies. Given the context, investors will rather conduct business with similar traits. In a male-dominated business market, securing donations from affluent male businessmen can work to the detriment of female CEOs (Melis, Nawaz 920-921).

## Education

A noteworthy and interesting concept that can be examined in looking to assess the theory is that a CEO's educational background means that the person belongs to a specific social stratum. Collins (1971) points out that the possibility that managers that can gain an education from a prestigious university is dependent on their parent's socioeconomic class. When the educational background of the company's senior management team differs from the CEO's, the way and understanding that the two parties will examine and resolve issues will differ and can possibly lead to delays in the resolution of issues. Hambrick and Mason (1984) argue that the fact that the CEO has gained a formal education degree is a significant factor in the decision-making process of the company.

Here, the argument points to the research that states that a manager's level of education is positively connected to their broad-mindedness and receptiveness to the use of new technologies in company operations. Here, the use of new technologies can generate discussions and inputs from managers that can in turn lead to a distinct methodology that the company will be known for in the future. Also, it can also be noted that companies that have managers with little or no formal educational background will digress from the standard corporate and sector policies and practices compared to managers who have attained formal education. It is suggested here that managers who have gained their educational levels from a prestigious university will have the needed managerial traits.

However, it must be argued here that the point cannot be universally true given the differing cultural and social dynamics in various countries and regions. Nonetheless, there are still significant similarities among the international community. For instance, Channon (1979)



mentioned that the relationship between corporate success and finishing from a prestigious university is weak; however, there are countries where this position holds true. In this light, it must be resolved that higher education facilities can impart the same skills and traits that these prestigious skills purportedly impart. Hambrick and Mason aver that not only does the CEO's degree dictate the development of their decisions, but also the choice of school. For instance, Harvard Business School graduates will differ in their courses and resources from Chicago Business School graduates.

If these prestigious schools will differ in the courses that they will graduate and bring in to the corporate arena, it is not a far-fetched premise that these graduates will bring in differing skill sets and knowledge sets when they become managers. An engineering graduate will bring a different skill and knowledge set compared to a history or law graduate. The position would posit that different courses will significantly impact their managerial approaches. There is research that seeks to determine whether this factor will affect their decision-making practices.

Regrettably, research has determined that there is no connection between a person's educational level or where they graduated from and their approaches to the consideration of the integration of new technologies and processes when these assume a senior management or CEO position in their respective companies. Moreover, a counterpoint to the position rejected the argument that CEOs that have a degree from a prestigious university will redound to the benefit of these companies; here, people who did not graduate from a prestigious university or have a college degree will not be able to benefit the companies that will hire them. In this position it is posited that these candidates, if they have sufficient knowledge and experience, can contribute to the growth and expansion of these companies as well.

## Social Class

Different companies possess differing social classes; prevailing research has not established or studied the relationship between a person's social class and corporate decision-making. This is the result of the high similarity levels of samples studying the social backgrounds of managers. Channon (1979) found several links between a business leader's social class and the growth approaches of their companies. Channon's research divided these businesses into three categories: family enterprises, corporations, and those operated by a diverse administrative group. Merchant-run businesses were operated by people who came from low-income families or with low educational attainment. On the other hand, companies that are run as corporations or family-run operations are seen to be coming from families or social backgrounds with higher social or economic backgrounds.

Nevertheless, there is a level of difficulty that will result from solely believing that the decision-making and managerial practices will be the result of the person's socioeconomic class. It is also difficult to accept the position that the governance model, or the small size of the company, will have a direct correlation with the type or form of growth strategies that it will employ. The research done by Collins and Moore (1970) points to the statement that corporate leaders coming from lower social classes will adopt aggressive policy options such as acquisition and variegation to generate credibility and enhance his/her companies' stature in the sector and debunk the notion that companies with leaders coming from a higher social class will automatically have a higher profit margin or stronger margin of expansion (Tian 169-171). Several research initiatives have been done on the theory, assessing the connection between a leader's communal characteristics and corporate performance and innovation levels. Withal, the

observable importance of these studies is largely ambiguous, particularly in terms of determining the relationship between a leader's intellectual, community, and mental attributes and a company's competitive output (Ting, Azizan, and Kweh 687).

## **Conclusion**

Bekos and Chari (2023) aver that analysts have studied the applicability of the theory across various national and social contexts, as well as various industries and sectors. Scholars have also supported the applicability of the theory in all corporate sizes, from family-owned and small companies to large multinational firms. Using data information gathered by Geletkanycz and Black (2001), the results posited the element that a CEO's leadership experience and principles have a strong influence on the decision-making processes of their companies. In the latter studies conducted by Lee and Park (2006), the research showed that leaders with diverse traits have a higher likelihood of leading their companies to develop international partnerships, which increases the diversity in their corporate structure.

Nonetheless, the two also post the limitations of the theory. First, the theory has been criticized for a lack of focus on the undergirding systems where senior-level managers and CEOs affect corporate performance, collectively known as the "black box." The utilization of a CEO's social, normative, and innate traits as indicators to establish their intellectual prowess and belief systems is more focused on general tendencies, abandoning the need to study the actual mental processes that influence critical choice-making and its actual effect on corporate performance and profitability. Though empirical leadership traits are helpful to facilitate replicability of observable results, they integrate non-essential findings rather than the needed mental results.

These are also flawed in their assertions and are undependable for determining intellectual or mental factors (Bekos, Chari).

As studies in the effect of top management's personal traits remain a strong avenue for research, the data and the related literature have slowly become more specialized. Here, there was a need to study the critiques that have been leveled against the theory. Regrettably, many of the critiques' findings have been neglected in the process of evaluating the theory. The growth of the research into the theory has exacerbated the difficulty of establishing a universal understanding of these reviews. For instance, the "black box" issue is one that has hounded the research on the theory. Withal, there is no general definition of what constitutes the "box." In this light, researchers and academics must strive to develop the "real-world" application of the theory. Challenges must be addressed to support the theory's validity and utility of the research surrounding the theory (Neely, Lovelace, Cowen, and Hiller 36-37).

### **Works Cited**

Bekos, Georgios, Chari, Simos. "Upper Echelons Theory." 2023

Devine, A. "*Building a micro foundation on the basis of the upper echelon theory.*"

<https://www.diva-portal.org/smash/get/diva2:1095472/FULLTEXT01.pdf> Accessed

September 22, 2024

Hart, Jon (2023). *Unpacking the upper echelon's cognitive black box: a qualitative study of selective attention and decision making in senior executives*. [Unpublished dissertation]

Liberty University

Melis, Andrea, Nawaz, Tasawar. "The impact of CEOs' personal traits on organizational performance: evidence from faith-based charity organizations". *Journal of Business Ethics*, Volume 190, 2024

Mukherjee, Tutun, Sen, Som. "Impact of CEO attributes on corporate reputation, financial performance, and corporate sustainable growth: evidence from India." *Financial Innovation*, Volume 8, issue 40, 2022

Neely, Brett, Lovelace, Jeffrey, Cowen, Amanda, and Hiller, Nathan. "Metacritiques of Upper Echelons Theory: Verdicts and recommendations for future research." *Journal of Management*, Volume 46, issue 6, 2020

Plockinger, Martin, Aschauer, Ewald, Hiebl, Martin, and Rohatschek, Roman. "The influence of individual executives on corporate financial reporting: a review and outlook from the perspective of upper echelons theory." *Journal of Accounting Literature*, Volume 37, 2016

Quttainah, Majdi. "Upper Echelon Theory: role of community and strategy." *International Journal of Innovation and Economic Development*, Volume 1, issue 3, 2015

Setiawan, Doddy, Harymawan, Iman, Adhriani, Desi, Pratama, Fadillah, and Santoso, Arif.

“Does the leverage of a company differ when led by a CEO from a reputable university?”

*Journal of Open Innovation: Technology, Market, and Complexity*, Volume 10, issue 2,  
2024

Tian, Ye. “A literature review on upper echelon theory”. *Advances in Economics, Business, and Management Research*. Volume 219, 169-173

Ting, Irene, Azizan, Noor, and Kweh, Qian. “Upper Echelon Theory revisited: The relationship between CEO personal characteristics and financial leverage decision.” *Procedia-Social and Behavioral Sciences*, Volume 195, 2015, pp. 686-694

El Abiad, Zouhour, Abraham, Rebecca, El-Chaarani, Hani, Skaf, Yahya, Binsaddig, Ruaa, and Jafar, Syed. “The impact of CEO characteristics on the financial performance of family-based listed in the Euronext Exchange.” *Journal of Financial Risk Management*, Volume 17, issue 3, 2024

# 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 Jain<sup>1</sup>

## 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

---

<sup>1</sup> 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.<sup>2</sup> 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.

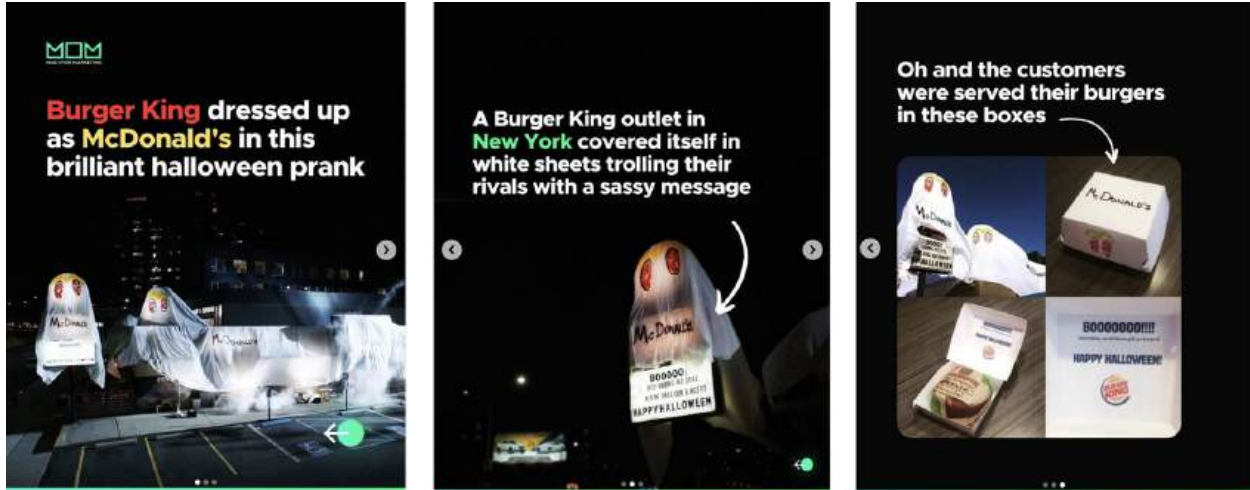


Figure 1. Source: [https://www.instagram.com/p/CzCACJ\\_vlv8/?utm\\_source=ig\\_web\\_copy\\_link&igsh=MzRIODBiNWFIZA==](https://www.instagram.com/p/CzCACJ_vlv8/?utm_source=ig_web_copy_link&igsh=MzRIODBiNWFIZA==)

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

<sup>2</sup> 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.”*<sup>3</sup>

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 <sup>4</sup>. To

---

<sup>3</sup> 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/>

<sup>4</sup> 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.<sup>5</sup>

An example of a tweet with three questions from the actual questionnaire is presented below:

#### *Tweets:*

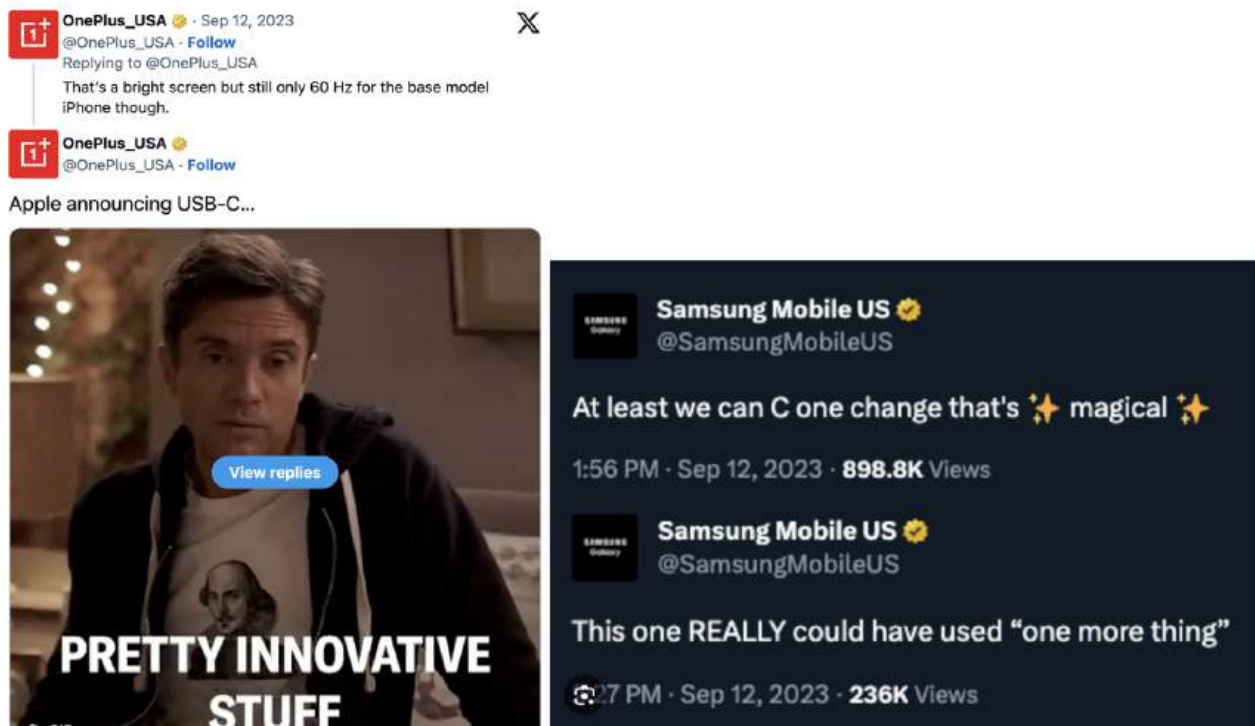


Figure 2. Source: <https://www.marca.com/en/technology/2023/09/13/6501057e22601de0128b45ee.html>

**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)

<sup>5</sup> **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?”

<b>Judgment</b>				
<i>Sample Size: 113</i>				
<b>Advertisement</b>	<b>Mean</b>	<b>Median</b>	<b>Mode</b>	<b>S.D.</b>
<b>Video 1</b>	2.841	3	3	1.424
<b>Video 2</b>	3.177	3	4	1.167
<b>Image 1 + Image 2: C type charger</b>	3.124	3	3	1.158
<b>Image 3: Camera</b>	3.451	4	4	1.275

<b>Image 4 +Image 5: Privacy</b>	3.540	4	4	1.303
<b>Image 7 + Image 8: Flip Feature</b>	3.283	4	4	1.271
<b>Image 9: Photos &amp; Videos Quality</b>	3.266	3	3	1.239
<b>Text 10 &amp; Text 11: Accessories</b>	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?"

<b>Brand Loyalty</b>				
<i>Sample Size: 113</i>				
<b>Advertisement</b>	<b>Mean</b>	<b>Median</b>	<b>Mode</b>	<b>S.D.</b>
<b>Video 1</b>	2.708	3	1	1.443
<b>Video 2</b>	2.805	3	4	1.487
<b>Image 1 + Image 2: C type charger</b>	2.628	3	3	1.219
<b>Image 6: Price</b>	2.584	3	3	1.272
<b>Image 7 + Image 8: Flip Feature</b>	2.593	3	1	1.334
<b>Text 10 &amp; Text 11: Accessories</b>	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.

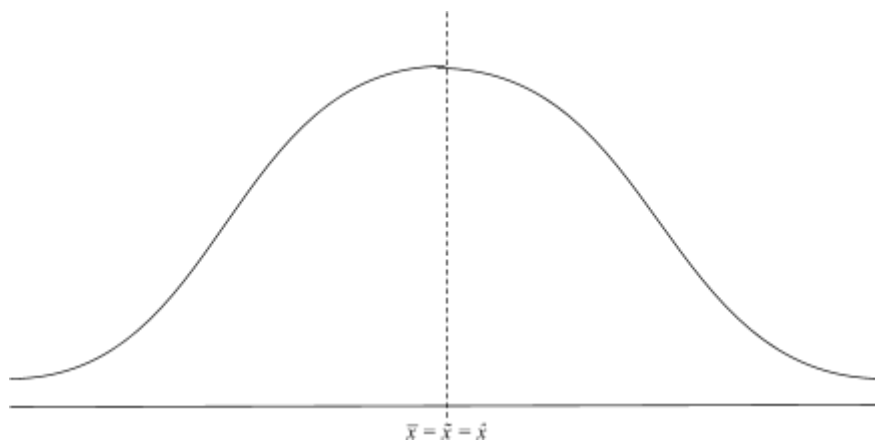


Figure 3. Normal Distribution Curve

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.

<b>Brand Loyalty</b>			
<b>Advertisement</b>	<b>Mean value of those who believed the ad was unfair</b>	<b>Mean value of those who believed the ad was fair</b>	<b>p-value</b>
<b>Video 1</b>	2.050	3.000	0.012
<b>Video 2</b>	2.200	2.960	0.027
<b>Image 1 + Image 2: C type charger</b>	2.545	2.605	0.863
<b>Image 7 + Image 8: Flip Feature</b>	2.605	2.364	0.00004
<b>Text 10 &amp; Text 11: Accessories</b>	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

<b>Image 4 +Image 5: Privacy</b>	2.857	4.042	0.005
<b>Image 7 + Image 8: Flip Feature</b>	1.960	2.983	0.001
<b>Image 9: Photos &amp; Videos Quality</b>	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.
- Tragakēs, 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:



Kevin Hsieh  
Kevin Hsieh, kevinmhsieh@gmail.com

### Acknowledgements

Thank you for the guidance of Simon Alford, mentor from Cornell University in the development of this research paper.



## I. INTRODUCTION

This research paper we experimented with different reinforcement learning algorithms and found different ways to train an artificial intelligence to play the Chrome Dino game. Our goal is to find different ways that we can allow the model to achieve the best performance. We tried training the model in different ways and managed to allow the dinosaur to be able to successfully jump over cacti, duck under birds, and jump over birds. We discovered that an evolutionary modification to Proximal Policy Optimization (PPO) worked best. We aimed to answer the question of what should be used to allow artificial intelligence to learn and play the game. We were trying to find a general algorithm that can be used across different games or projects besides the Chrome Dino game that uses reinforcement learning.

**Related Work.** Artificial Intelligence has been rapidly growing in the past few decades. Researchers discovered different algorithms that can be used for different topics. However, it is hard to create a general algorithm that can work in diverse environments because the AI has to be trained on specific data for the specific goal. Training artificial intelligence also takes a large amount of time. Because of this, researchers try to find ways to improve the generalization and efficiency of models. John Schulman proposed to do multiple gradient updates to every sample. Using new methods which they called proximal policy optimization—also known as PPO, they could outperform other methods by being simpler to implement, more general, and have better sample complexity (Schulman et. al., 2017). However there are many other algorithms. For example, Mastering Diverse Domains through World Models presents DreamerV3 which is a general algorithm which can be used over 150 different tasks. It was used to mine diamonds in minecraft using reinforcement learning. It used PPO as it has become a standard in reinforcement learning (Hafner et. al., 2023). Another paper, Playing Atari With Deep Reinforcement Learning, used reinforcement learning algorithms to play the Atari game. It combined deep learning with reinforcement learning and fed the computer frames of the game using convolutional neural networks. Using their algorithm, their AI could play six of the seven games it was tested on with no adjustment of the algorithms and parameters (Mnih et. al., 2013) . Additionally, an AI using reinforcement learning was trained to be able to play Go. The AI was able to achieve a 99.8% winning rate against other programs and professional human players (Silver et. al., 2016) .

## II. METHODOLOGY

**Plan.** Our first step was to set up the environment for artificial intelligence. The experiment applied deep reinforcement learning algorithms in order for the model to learn. We used a combination of neural networks and reinforcement learning. The neural network used proximal policy optimization and stable baselines3 for learning. Stable baselines3 is a package of deep reinforcement learning algorithms. The game which was played was the Chrome T-Rex Rush which is a game that Chrome allows you to play when you lose internet connection. The Chrome T-Rex Rush was from an open source repository made by Shivam Shekhar (Shekhar. 2023) . In order for the AI to be able to input and gather data from the game. There is a channel used to channel data from the game to the AI. Gymnasium, which is a library for deep reinforcement learning environments, was used to create the environment which served as a gateway between the game and the AI.

**Setup.** After creating a custom environment for the Chrome T-rex game. The game was converted to be able to play a single frame of the game upon a function call. A channel was created to hold variables that were updated with game data which was given to the neural network upon each step. Data that was given was the x location of the dinosaur, the distance of the nearest bird, and the distance of the nearest cactus. The dinosaur's reward system was set up so that for every successful jump over the bird, it would gain a 1 reward upon every successful jump over a cactus or a bird. Upon death, the model would receive a -10 punishment.

**Data Collection.** Data collection was added in order to gather the data of the average lengths and rewards for every rollout of the game. The dinosaur was trained overnight and data was collected. Afterwards, the dinosaur was trained again over 3 nights (approx. 12 hours per night) and data was collected afterwards. This training process was repeated again. Following this, the seed of the game and the AI model was changed to not be randomly generated in hopes to allow more accurate data collection. We can see the data that was collected in the data graphs shown in the results below.

**Experimentation.** The speed of the game was also changed to be constant. The model was trained for a while and the ent-coef was increased to 0.01 when it began underperforming and getting stuck at a local minimum. Data was then collected. Afterwards, optimized parameters for CartPole from RL-baselines3-zoo (RL Baselines3 Zoo: A Training Framework for Stable Baselines3 Reinforcement Learning Agents, 2022) were then used for the dinosaur. The dinosaur was then trained over one night and data was collected. The dinosaur was then trained again over night and data was collected. After this, the dinosaur's observation space was increased from its current observations to include the bird's Y location and the distance between the next two obstacles. The distance between the bird and the dinosaur was also changed to be calculated through just the X value rather than the pythagorean theorem as the bird's Y value changes. The model was trained overnight and its performance data was collected. Next, the dinosaur's environment was changed to be able to support multiple environments running at the same time (Figure 7).

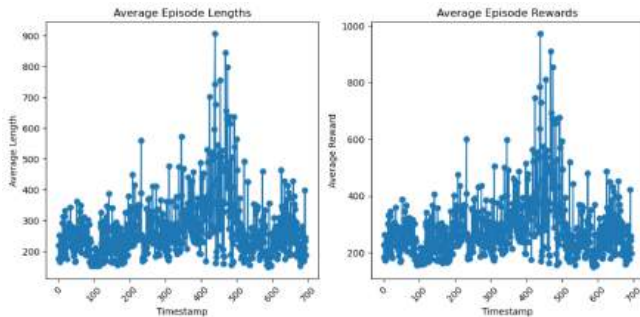
**Evolutionary Algorithm.** A new algorithm was created which hoped to imitate evolution. The best environment's model would be chosen every ~5 minutes and replaced with the best performing model. It would find the model with the most survival time and replace all other models with it and then

continue training. The models would also be backed up to its previous version if its performance dropped below 50% of its survival time length. Additionally, loading and saving models multiple times for the evolutions resulted in the model being no longer able to save and load. In order to combat this issue, it was made so that just the model's policy was saved and loaded during the evolutions. The model was then trained overnight on Macbook with 16 cores training 6 environments simultaneously and data was collected the following morning. The bird spawning was then increased to have a 50% spawn rate every frame with a maximum of one bird per frame. In order to compare the performance of this evolutionary algorithm. It was run multiple times with different amounts of environments. Data was collected for the different amounts of environments.

**Evolutionary Algorithm.** The dino was never able to figure out how to duck under birds. Even when the bird was too high for the dino, it still tried to jump over it and die. In order to get the dino to duck, an incentive was added. The dinosaur was given 1000% its successful jump reward for every frame it was under a bird that was equal to the Y level of the highest bird. After changing the death penalty to -1000 and the dino a reward of 100 for every second it was ducking, it finally learned to duck under birds. Its performance can be seen in Figure 13.

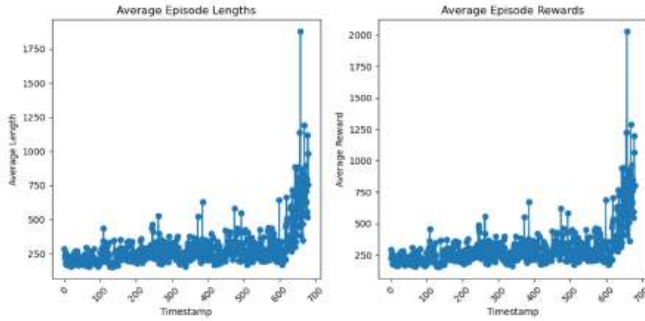
### III. RESULTS

The first training session used the observations of only the dino X location, nearest bird distance, nearest cactus distance. The length of the training was around ~36 hours. The result was that the dinosaur learned to jump over cactuses but eventually lost its ability to do so after continual training.



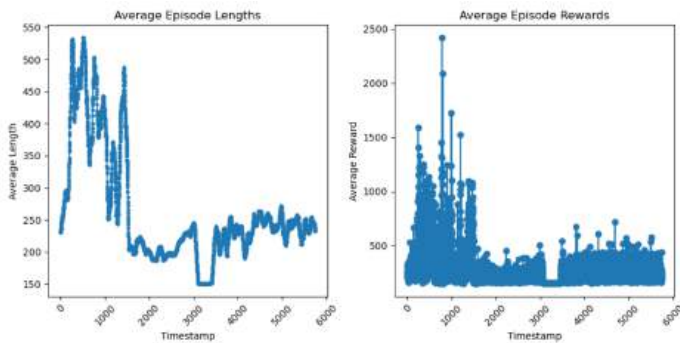
**Figure 1.**

In the second training session. The same observations and time was given as the first training session. The result was a slow training speed and then an eventual learned ability to be able to jump over cactuses. It was similar to the first session but slower.



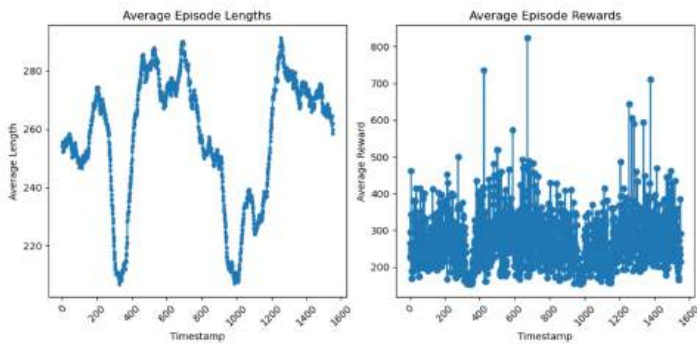
**Figure 2.**

In the third training session, the same parameters were given. It was training for a longer period of time. However, it suddenly lost all its skill as shown in timestamp  $\sim 1500$  and eventually just tried to die as fast as possible by killing itself with the first cactus. This is shown by the second drop to its lowest point at timestep  $\sim 3000$ . In an attempt to salvage the model, the entropy coefficient was increased to 0.01 to encourage it to explore at timestamp  $\sim 3,400$ . This resulted in the model jumping over cactuses but performing poorly continuously.



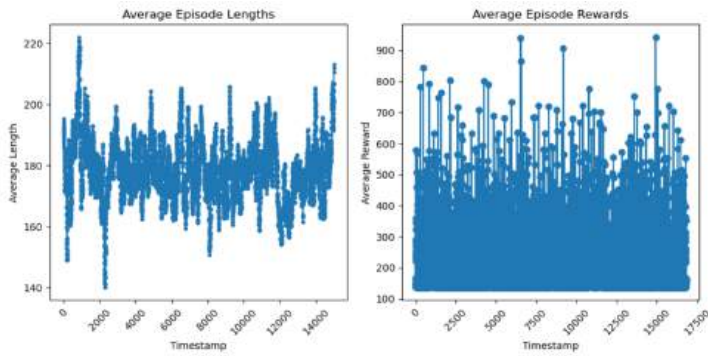
**Figure 3.**

In the fourth training session using the same parameters, the model was trained over 24 hours using the same observations but beginning with a entropy coefficient of 0.01. This resulted in an unstable model that never reached a good performance.



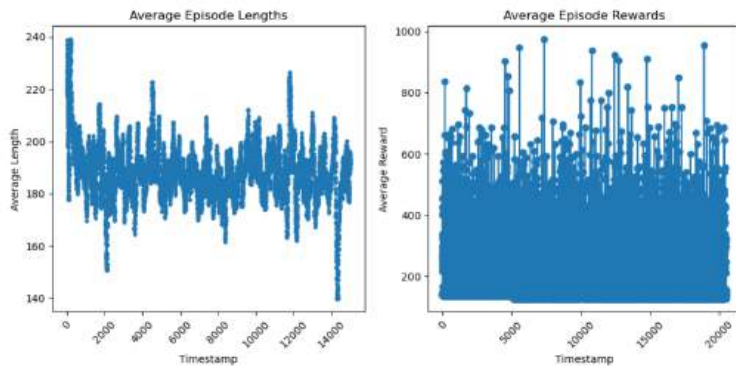
**Figure 4.**

In the fifth training session, the entropy coefficient was changed back to zero. The parameters were set to the optimized cartpole's parameters and then trained for 12 hours. Again, this resulted in an unstable model that performed poorly.



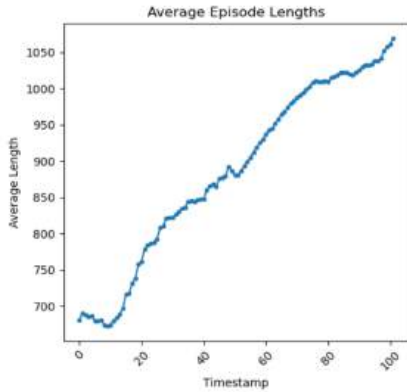
**Figure 5.**

In the sixth training session, the experiment was repeated with the optimized cartpole parameters but with a entropy coefficient of 0.01. The performance increased a bit but still performed fairly poor.



**Figure 6.**

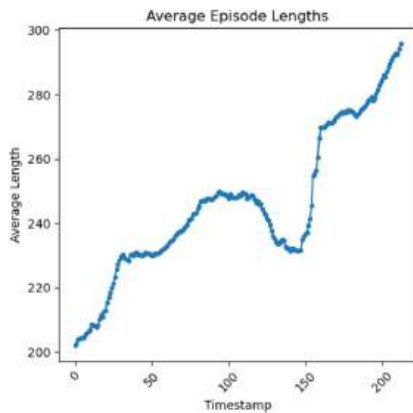
In the seventh training session, the bird's Y location and obstacle distances were added to the model's observation space. This was also when multiprocessing and evolution was introduced. It was trained over 3 hours with 6 environments running simultaneously. The result was a continuous increase in performance.



**Figure 7.**

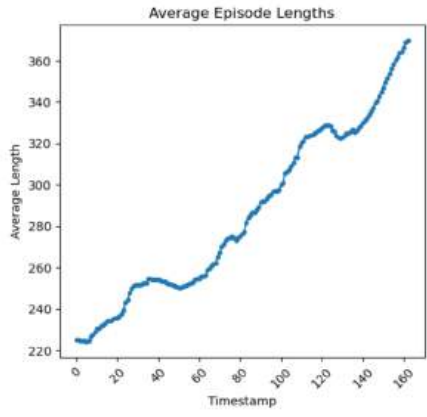
It was discovered that using evolution, multiprocessing, and more observations, we were able to achieve a better performance in a shorter time span. In the following training sessions, we trained the models using different amounts of environments. Birds were also made to spawn with a 50% chance per frame.

The first run was with a single environment with no multiprocessing or evolution which ran for ~13 hours. This allows us to see the effectiveness of just increasing the observation space. The result was that it reached a similar performance to before the observation space was changed. However, its performance still improved as the model had to learn how to jump over birds along with the cacti, while the model did not have to before the observation space was increased.



**Figure 8.**

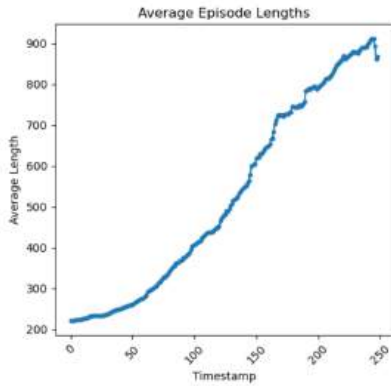
The model was then run with two environments using multiprocessing and evolution for ~11 hours. However, if we just compare the models at time step 150. The model with two environments performed better as the first model reached 230, while the second model reached 350.



**Figure 9.**

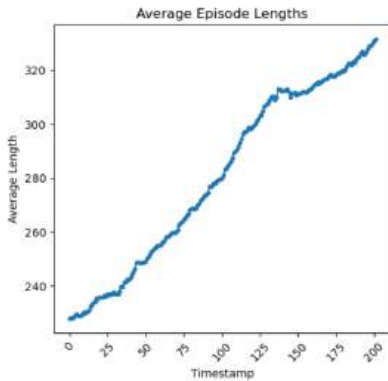


Following this, another model was then run with four environments for ~14 hours. If we compare the models at time step 150. The model with two environments reached 350, while the model with four environments reached a little less than 600.



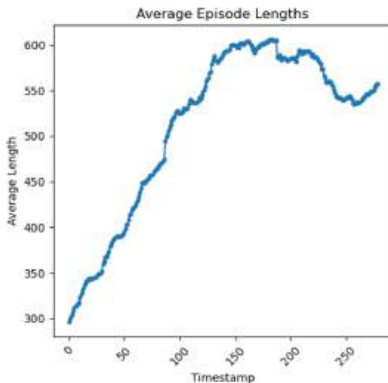
**Figure 10.**

The model was then trained with 6 environments. However, this is when the model's performance begins decreasing. It performs similarly to the model with 2 environments running.



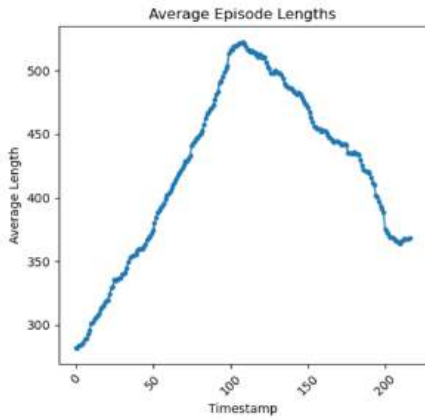
**Figure 11.**

Following the previous model, another model was run with 8 environments. This model reached 600 at time step 150 similarly to the environment with 4 models.



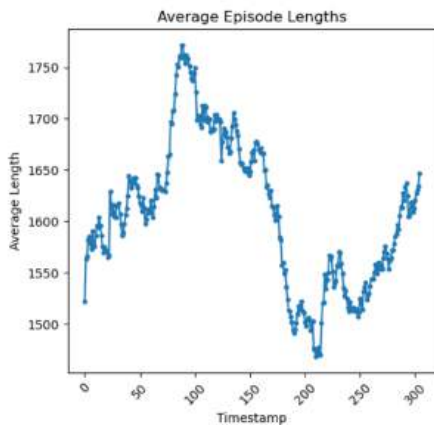
**Figure 12.**

A model was then trained with 16 environments. Its performance speed was much quicker than the previous models, however its performance began dropping after it reached time step 100 to lower than the previous three models.



**Figure 13.**

After giving the dino an incentive to duck. It was finally able to duck which caused it to achieve a much higher performance than it had ever before.



**Figure 14.**

After observing the data collected from this experiment. We can conclude that the model performs better and more efficiently with the evolutionary algorithm compared to without the algorithm. However, the data does not have a clear enough trend to allow us to understand how many environments result in a better performance.

#### IV. DISCUSSION

**General Analysis.** We can observe from Figure 1. And Figure 3. that the model would train for a certain amount of time and lose its capability. I hypothesize that the reason behind this is that the model learns to jump over a cactus and then reaches the section where the bird spawns. This would then result

in the model getting confused as it is introduced to a new obstacle that it has not seen before. It then attempts to experiment which results in the model getting confused with jumping over cacti. In order to combat this problem, the bird's spawning rate was increased and the model was made to take in more observations of the bird. As we can observe from Figure 8, this improves the ability for the model to learn to overcome this obstacle and seems to remove the sudden dip in ability. However, the model would need to be trained for a longer period of time to determine whether this is truly the factor which is causing its performance to drop. We can see that the increase in the entropy coefficient can be also used to prevent the model from finding a local optimum and staying there resulting in a low performance.

**Evolutionary Algorithm Analysis.** Besides this, we can see from the results that the model's performance increases with the evolutionary algorithm. The reason it performs better is probably because it allows the model to experiment and the model that finds does something correct would then replace all other models, while the model which performs poorly would then be removed from the evolution. This would then result in a continual growth in ability unless all models suddenly perform worse. Since the computer it is being trained on only has 16 cores, it should not be able to truly multi-process more than 16 environments. Figures 8-11 attempts to see if the amount of environments is truly affecting the model's performance. We can see a linear growth with environments 1, 2, and 4 as shown in Figures 8-10. However, the performance then decreases with 6 environments and onwards as shown in Figures 11-13. This shows that 4 environments seems to be the optimal, however there does not seem to be any reason that it is so. This would require further research in order to understand the amount of environments' equivalent performance. Despite this, it does show that this algorithm seems to be helpful when trying to increase the performance and speed of the reinforcement learning model.

## V. CONCLUSION

In conclusion, this project shows that evolutionary algorithms should be further researched as it could provide a more efficient way for reinforcement models to learn. Further studies should experiment and discover what is the optimal amount of environments and ways to shape the evolutionary algorithm in order to improve learning. This experiment was limited in time which resulted in more unstable conclusions. I hope that future research could have more time to experiment in more detail and discover ways in which model training can be improved.

**Bayesian Optimization.** I came up with an idea to have a reinforcement learning algorithm, the "teacher," train the training algorithm, the "student," through setting its parameters. The teacher would try to optimize the training parameters such as reward amount, amount of environments, ent\_coef, etc and getting the dino's score as a reward. As a first step towards training a teacher, I looked into hyperparameter tuning of the RL algorithm, aka discovering the best setting of the RL game to train the dino with. To conduct hyperparameter tuning, I used Bayesian Optimization (Bayesian Optimization, 2023). Bayesian Optimization allows us to optimize the parameters for the RL algorithm by modeling a distribution over best parameters and searching parameters most likely to improve the system. However, I unfortunately ran out of time to complete this experiment. I hope to develop this more in the future.

## References:

1. J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal Policy Optimization Algorithms," arXiv.org, 2017. <https://arxiv.org/abs/1707.06347>
2. D. Hafner, J. Pasukonis, J. Ba, and T. Lillicrap, "Mastering Diverse Domains through World Models," arXiv.org, Jan. 10, 2023. <https://arxiv.org/abs/2301.04104>
3. V. Mnih et al., "Playing Atari with Deep Reinforcement Learning," arXiv.org, 2013. <https://arxiv.org/abs/1312.5602>
4. D. Silver et al., "Mastering the game of Go with deep neural networks and tree search," Nature, vol. 529, no. 7587, pp. 484–489, Jan. 2016, doi: <https://doi.org/10.1038/nature16961>.
5. S. Shekhar, "shivamshekhar/Chrome-T-Rex-Rush," GitHub, Sep. 06, 2023. <https://github.com/shivamshekhar/Chrome-T-Rex-Rush>
6. "RL Baselines3 Zoo: A Training Framework for Stable Baselines3 Reinforcement Learning Agents," *GitHub*, May 23, 2022. <https://github.com/DLR-RM/rl-baselines3-zoo>
7. "Bayesian Optimization," GitHub, May 21, 2023. <https://github.com/bayesian-optimization/BayesianOptimization>

# **Labour Laws in India and Canada: A Comparative Study of Legislative Changes and Their Driving Forces**

Reyhaan Aryan

Sanskriti School, New Delhi, India

## **ABSTRACT**

This study examines the historical development of labour laws throughout the world, paying particular attention to important events and reforms that occurred in Canada and India and greatly influenced the current foundations for those laws. We begin by first introducing the history of the development of labour laws from the Industrial Revolution to the present, emphasising important international turning points and the impact of international organisations like the International Labour Organization (ILO). We then focus on the evolution of labour laws in India, looking at important legislative developments like the Industrial Disputes Act and the most recent Labour Codes of 2020, as well as important events like the Bhopal Gas Tragedy and the reforms that followed, the role of the Indian labour movement, and more. Similarly, the study explores the evolution of Canada's labour laws, focusing on the role of events such as the Winnipeg General Strike, the introduction of the Canada Labour Code, and committees formed to address social ills. By comparing these two nations, the paper identifies the fact that India, has more changes in labour due to tragedies such as cases where cities are rampant with child labour. Whereas Canada, has had a significant portion of its laws founded on the basis of protests or general citizen dissatisfaction. This phenomenon offers a nuanced view into how each country may view the opinions of its citizens.

**Keywords:** Labour Laws, India, Canada, labour regulations

## **INTRODUCTION**

Law, in general, is a body of regulation that is drafted and enacted by a government or institution for regulation of behaviour (Robertson, 2006). Over the years, countless scholars have tried to give a concrete definition of law. In the book “The Concept of Law” (1961), H. L. A. Hart states that the first and foremost function of law is to protect law-abiding citizens from wrongdoers which have potential for harm. Moreover, through laws, individuals are enabled to utilise complex economic and social activities. A true society has laws which are free from discrimination on the basis of race, wealth, creed, gender etc. Laws cover vast domains such as: International law, Constitutional and Administrative law, Criminal law, Contract law, Torts,

Property law, Labour law etc. (Horwitz, 1982). This paper will specifically focus on the labour laws of India and Canada and analyse the key moments that led to policy changes in them.

Laws are diverse and differ according to sectors in addition to domains as well, such as, the laws for the regulation of the informal sector would differ from the laws regulating the formal sector (Tokman, 1979).

The informal sector comprises of street vendors, sub contractors etc. whereas the formal sector is made up of government, banking, education jobs, among others. Since there are so many diverse sectors, laws that regulate these sectors cannot be generalised. Problems that plague the informal sector such as depressed wages, carrying out illegal activities, lack of contracts, exploitation etc. (Papola, 1980; Muthusamy and Ibrahim, 2016; Prasad, 2023), are significantly different than problems in the formal sector. This means that laws have to be specifically designed to fit the sector that they are targeting. The process for making these diverse laws depends on the country/institution which is responsible for the legislature. This paper will be focusing on labour laws, which are laws that regulate the workforce and the party employing the workforce. In fact, Labour laws in itself can further be categorised: as related to work conditions, payments, sectors, labour relations, social protection and security, and social and informal institutions. However, before present day labour laws can be analysed, it is important to first understand the history of labour laws.

Despite the plethora of research on labour laws, there exists a noticeable gap in comparative studies, particularly between countries as diverse as India and Canada. This study aims to bridge this knowledge gap by providing a comprehensive analysis of the labour laws in these two nations, considering their unique cultural, economic, and political contexts.

The objective of this study is twofold. Firstly, to scrutinise the structural and practical differences and similarities in the labour laws of India and Canada. Secondly, to assess how these laws impact the labour rights and incidents in each country, considering factors such as workforce demographics, employment rates, and reported labour violations. This comparative analysis aims to offer valuable insights for policymakers, labour organisations, and scholars, contributing to the broader discourse on global labour rights and regulations.

### **Historical Review of Labour Laws:**

England, being the first country to industrialise, was also the first country that formalised labour laws. This resulted in the draftment of the Master and Servant Act (1867) and The Combination Act (1875). Both these acts were formed to protect the employers instead of focusing on the workforce (Arnold and Britain, 1875; Davis, 1875). For instance, the Master and Servant Act was to criminalise breach of contract by workers (Naidu and S. Y. N, (2013). Slowly and steadily, these ideas spread throughout Europe to countries like France, The Netherlands, Switzerland etc.

Over time, labour laws have expanded to encompass a wide range of issues, including but not limited to, minimum wage, maximum working hours, health and safety regulations, and anti-discrimination policies. Framework for labour legislations involve three key stakeholders - the employers federations which represent the interests of the employers, trade unions that represent the interest of the labourers and finally the government which seeks to act as an intermediary between the first two and set regulations for labour. International organisations like the ILO and the EU where applicable also act as a fourth stakeholder that sets frameworks for labour regulations that all countries must follow in order to ensure a decent working condition (ILO, 2024). However, it is important to note that organisations such as the ILO only have lobbying capability and are not enforceable. According to official ILO records, approximately 164 countries have adopted the 48-hour work week, with around 86% of its member states having some form of legal minimum wage. The importance of labour laws cannot be overstated. They are instrumental in safeguarding workers' rights, ensuring fair pay, and providing a safe working environment. Beyond worker protection, these laws play a crucial role in promoting social justice and fostering sustainable economic growth (Saini, 1999). According to the International Labour Organisation and World Bank, well-implemented labour laws can lead to a substantial increase in labour productivity (Kuddo, 2018; Lee et al, 2008). Furthermore, these laws are essential for maintaining healthy employer-employee relationships, which in turn can enhance job satisfaction and productivity. In the broader context, labour laws are integral to achieving several United Nations Sustainable Development Goals (SDGs), particularly those concerning decent work and economic growth. These laws, which include regulations on minimum wages, working hours, health and safety standards, anti-discrimination measures, and the right to collective bargaining, play a fundamental role in ensuring equitable treatment and safeguarding the rights of workers. Some of the most critical ideas in labour law came to be during the Industrial Revolution, a period marked by significant economic and social changes that necessitated new forms of worker protection. For instance, the UK's Factory Act of 1833, one of the earliest labour laws, introduced critical reforms like setting working hours and age limits for child labourers. Such laws laid the groundwork for modern labour legislation, emphasising the need for balance between industrial growth and worker welfare. (Gray, 2002)

## **LABOUR FRAMEWORK IN INDIA**

India's labour law framework is intricate, reflecting the nation's diverse and populous character. With roots in colonial legislation, such as the Factories Act of 1881 (Jhadav, 2019), India's labour law system has evolved into an extensive structure comprising over 200 state laws and about 50 central laws. One thing to note is, in India, labour laws are less pro labour and more intended to ensure that employers have a smoother production process. (Sodhi, 2014)

India, with its massive labour force of over 500 million, faces unique challenges in implementing labour laws. The labour market is predominantly informal, with the Periodic Labour Force Survey (2019-2020) suggesting that about 90% of the workforce is engaged in informal employment (John, 2003). This sector often falls outside the ambit of formal labour law protection. The informal sector's size highlights the disparity in labour law coverage and enforcement in India, underscoring the need for comprehensive policy measures to extend legal protections to these workers. Due to the federal structure of India's governance and in order to properly serve the large workforce, India's labour code is managed by both the centre and the state level governments.

Key legislations like the Factories Act (1948) and the Industrial Disputes Act (1947) were revolutionary at their time of enactment but now face challenges in addressing the needs of a modern, diverse, and largely informal workforce. The Factories Act, which governs worker safety and health, applies to establishments of a certain size, leaving many workers in smaller enterprises unprotected. The Industrial Disputes Act, designed to resolve industrial disputes, is often criticised for its protracted and complex dispute resolution process. (Menon and Majumdar, 2005)

Since India has over 50 separate Central Government labour law legislations and hundreds of State Government legislations, it soon became impossible for organisations and workers to remember and explain each law and its provisions in detail. This section highlights critical provisions of some of the critical labour laws in India.

### **1. Apprentices Act, 1961**

The purpose of this act is to enforce employers to provide adequate technical training to apprentices involved in craftsmanship. It highlights the following details: Obligation of Employer, obligation of apprentice, who is an apprentice, duration of training, reservation for marginalised groups, payment to apprentice, termination of contract, minimum stipend, dispute and penalties.

Under this act, an employer who has employed an apprentice must ensure training is provided to that person, regardless of whether the employer himself is qualified in the trade. Moreover, other responsibilities of the employer are also laid out such as payment of wages at prescribed rates and utilisation of proper instructional staff to train apprentices. (NCIB, 2021a)

To become an apprentice in India, certain requirements must be met. The minimum age at which one can be employed is 14 and the applicant must be medically fit for the trade apprenticeship is sought for. A minimum monthly wage of ₹75 is prescribed and the employer must submit ₹5000



in a government bond to guarantee the performance of the terms and conditions of such contracts. (Gayithri et al, 2019)

## **2. Factories Act, 1948**

The objective of this act was to regulate the labour standards being followed in factories. It focused on providing safety, anti-exploitation, annual leave, welfare measures among others. The notable provisions of this act include provisions on working hours, working conditions, safety regulations, among other things.

As per the regulation laid down by the Adult Working Hours Regulations, workers are prohibited from working more than 48 hours in a week and at least one weekly holiday must be provided. Further, the factory needs to be kept clean (Section 11). There should be systems in place for disposing of waste and effluents (Section 12). Ventilation must be sufficient, and the temperature should be comfortable for employees (Section 13). Dust and fumes must be controlled to stay within acceptable limits (Section 14). Artificial humidification should meet prescribed standards (Section 15). Overcrowding should be avoided (Section 16). Adequate lighting, drinking water, latrines, urinals, and spittoons must be provided (Sections 17-19). Sufficient spittoons should also be available (Section 20).

The act introduced many standards for employers to ensure the safety of workers in the factories. These include but are not limited to proper fencing for machines when they're running to protect workers (Sections 21-27), regular testing of hoists and lifts to ensure that all parts are in order (Sections 28-29), pressure plants to be checked according to regulations (Section 31), ensuring that the floors, stairs, and access points are sturdy and clear of obstacles (Section 32), supply of safety gear for eyes and protection from dangerous substances like dust and fumes must be provided (Sections 35-36), and workers are responsible for using them properly, additional safety measures are required for hazardous substances (Sections 41A-41H) adequate fire fighting equipment should be on hand (Section 38) and a Safety Officer needs to be appointed if there are 1,000 or more workers in the factory (Section 40B).

## **3. Child labour prohibition and regulation Act 1986 & The child labour rules, 1988**

There are a number of Acts which prohibit the employment of children below 14 years and 15 years in certain specified occupations. However, prior to this act, there were no regulations for the employment of minors below the age of 18 and above 14. This act, first and foremost, aimed to, ban the employment of children, i.e., those who have not completed the age of 14, regulate the conditions of work of children in employments where they are not prohibited from working and lay down enhanced penalties for employment of children in violation of the provisions of

this Act, and other Acts which forbid the employment of children (Upadhyay and Kumar, 2017).

Important provisions of this law include provisions that specify that the period of work on each day for the child shall be so fixed that no period shall exceed three hours and that no child shall work for more than three hours before he has had an interval for rest for at least one hour.

No child shall be required or permitted to work overtime. No child shall be required or permitted to work in any establishment on any day on which he has already been working in another establishment. Every child employed in an establishment shall be allowed in each week, a holiday of one whole day, which day shall be specified by the occupier in a notice permanently exhibited in a conspicuous place in the establishment and the day so specified shall not be altered by the occupier more than once in three months.

Every occupier shall maintain, in respect of children employed or permitted to work in any establishment, a register to be available for inspection by an Inspector at all times during working hours or when work is being carried on in any such establishment showing: the name and date of birth of every child so employed or permitted to work; hours and periods of work of any such child and the intervals of rest to which he is entitled; the nature of work of any such child; and (iv) such other particulars as may be prescribed. (NCIB, 2009)

It is imperative to broaden our understanding first by introducing the current labour law framework in India regarding labour, namely, the four labour codes in India which were formed in India by combining all the current bills about legislature into four codes; the Code on Wages, Industrial Relations Code, Social Security Code and Occupational Safety, Health and Working Conditions Code.

### **1. Code on Wages:**

The Code on Wages, a cornerstone of labour reform in India, serves as a beacon for fair compensation practices across industries and regions. It lays down a national minimum wage, a significant stride towards ensuring economic justice and dignity for workers. For example, the minimum wage rate, set at ₹178 per day as mandated by the code, offers a baseline income crucial for meeting basic needs such as food, shelter, and education. Moreover, the code's provision for periodic review, at least once every five years, underscores its commitment to adapting wage rates in alignment with changing economic landscapes, safeguarding workers against inflationary pressures and cost-of-living adjustments.

Furthermore, the Code on Wages doesn't merely establish minimum wage thresholds; it delineates transparent mechanisms for wage determination, factoring in skill levels, geographical variations, and nature of work. This clarity fosters trust and accountability in labour relations,

empowering workers with the knowledge of their rightful earnings and employers with clear guidelines for wage structuring. By instituting such robust frameworks, the code aims to bridge wage disparities and uphold the principle of equal pay for equal work, thereby fostering a more equitable and inclusive workforce.

## **2. Code on Social Security:**

The Code on Social Security represents a paradigm shift in India's approach towards worker welfare, expanding the ambit of social security benefits to encompass a broader spectrum of the workforce. For instance, the inclusion of workers from the informal sector, who constitute a significant portion of India's labour force, marks a pivotal stride towards bridging long standing gaps in social protection. Under the code, schemes for insurance, pensions, maternity benefits, and gratuity are not just statutory mandates but pillars of socio-economic empowerment, offering a safety net against life's uncertainties.

Take, for instance, the provision for maternity benefits. The code mandates a maternity benefit equivalent to the average daily wage for 26 weeks, ensuring that expectant mothers can afford essential healthcare and support during childbirth and postpartum recovery. Additionally, the co-contribution model for schemes such as the Employees' Provident Fund (EPF), wherein both employers and employees contribute 12% of the worker's salary, underscores a shared responsibility towards financial security in retirement. Such provisions not only address economic vulnerabilities but also foster a culture of solidarity and inclusivity, underpinning India's vision of a welfare-oriented society.

## **3. Code on Industrial Relations:**

The Code on Industrial Relations stands as a cornerstone in reshaping the dynamics of labour-management relations in India. By modernising the process of trade union formation and collective bargaining, it seeks to empower workers with a stronger voice in shaping their working conditions and rights. For example, the code specifies clear guidelines for trade union recognition, ensuring that unions represent a significant portion of the workforce, thereby enhancing their legitimacy and bargaining power. Additionally, the introduction of mechanisms for dispute resolution, such as conciliation and arbitration, offers alternative avenues for resolving conflicts, fostering a culture of dialogue and cooperation between employers and employees.

Moreover, the code brings clarity and stability to industrial relations by outlining the rights and responsibilities of both employers and workers. By stipulating provisions such as the maximum working hours (eight hours a day or 48 hours a week) and overtime compensation, it safeguards workers against exploitation and ensures a healthy work-life balance. Furthermore, the code

strikes a balance between the interests of employers and employees by promoting industrial harmony and productivity, thereby laying the foundation for sustainable economic growth and social development.

#### **4. Code on Occupational Safety, Health and Working Conditions:**

The Code on Occupational Safety, Health and Working Conditions heralds a new era in prioritising the well-being and safety of workers in India. With stringent standards and regulations for workplace safety and health, it aims to create a conducive work environment that safeguards workers from occupational hazards and ensures their physical and mental well-being. For instance, the mandate for establishments with more than 500 workers to appoint a safety officer, coupled with regular inspections, underscores a proactive approach towards enforcing safety protocols and preventing accidents.

Additionally, the code's provisions regarding working hours, rest periods, and leave entitlements serve to protect workers' rights and promote their overall welfare. For example, the stipulation that workers cannot be required to work for more than six hours continuously without a break of at least half an hour ensures that they have adequate rest periods to prevent fatigue-related accidents and maintain productivity. Furthermore, the provision for at least one day off every week ensures that workers have time for rest and recuperation, fostering a healthy work-life balance. By prioritising occupational safety and health, the code not only protects workers but also contributes to increased productivity and organisational success.

### **LABOUR LEGISLATION IN CANADA**

In contrast, Canada's labour laws are characterised by a blend of federal and provincial regulations, catering to the diverse needs of its workforce across various sectors and regions (Sack, 2010). Recent amendments to the Canada Labour Code, effective February 2024, demonstrate a progressive shift towards more employee-centric practices. These amendments include more nuanced termination notice entitlements based on the length of continuous service. Such changes indicate an evolving landscape of labour laws in Canada, increasingly focused on employee welfare. (Doorey, 2024)

Canada, with a labour force of about 20 million (Labour Force Survey, 2021), demonstrates a more streamlined approach to labour legislation. Canadian labour laws are renowned for their inclusivity and fairness, covering approximately 96% of the workforce. Laws such as the Canadian Labour Code and the Employment Equity Act are testament to the country's commitment to a balanced and equitable workforce. The Canadian Labour Code sets out comprehensive standards for employment conditions and rights for industries under federal

jurisdiction. Canada, similar to India, operates its labour code from the state and central level due to its federal nature.

Canadian labour laws also feature robust anti-discrimination statutes, prohibiting discrimination on grounds such as race, creed, sex, disability, age, and sexual orientation (Heritage, 2022). This comprehensive approach to anti-discrimination is evident in the broad spectrum of protected grounds across various jurisdictions within Canada. For example, the human rights legislation in Canada explicitly prohibits both direct and indirect discrimination, ensuring that employers cannot enact policies or practices that have an adverse impact on employees with protected characteristics. These protections extend to all aspects of the employment relationship, including hiring, promotions, discharge, and conditions of employment. Moreover, sexual harassment a recognized form of workplace harassment, requires employers to implement specific policies and provide training to address this issue in most jurisdictions.

Statistics and reports further illuminate the context of labour laws in both countries. In Canada, the implementation of new administrative policies by the Canada Revenue Agency in 2024 to determine the province or territory of employment for remote workers for payroll reasons highlights the country's adaptation to emerging work trends. Additionally, amendments to the labour code in 2023 exempted certain sectors from specified hours of work requirements, reflecting the evolving nature of different industries. This includes sectors such as banking, telecommunications, broadcasting, rail, and airlines.

The Canadian Labour Code, a cornerstone of labour legislation in Canada, encompasses a comprehensive framework aimed at promoting fair labour practices, ensuring workplace safety, and fostering harmonious labour-management relations. Central to the code is its commitment to safeguarding the rights and interests of workers while balancing the needs of employers and the broader economy.

One significant aspect of the Canadian Labour Code is its provisions regarding hours of work and overtime compensation. For instance, the code stipulates that employees are entitled to a minimum of one day off with pay per work week or a minimum of four days off with pay per work month. This ensures that workers have adequate time for rest and recuperation, promoting their overall well-being and work-life balance. Moreover, the code mandates that employees cannot be required to work more than eight hours in a day or 40 hours in a work week, with exceptions outlined for certain industries or occupations. These regulations protect workers from exploitation and ensure that they are fairly compensated for any additional hours worked through provisions for overtime pay, typically set at one and a half times the regular rate of pay.

Another key aspect of the Canadian Labour Code is its provisions regarding occupational health and safety. The code mandates that employers provide a safe and healthy work environment for

their employees, free from hazards that may cause injury or illness. For example, employers are required to conduct regular inspections of the workplace to identify and mitigate potential hazards, such as unsafe machinery or hazardous materials. Additionally, the code outlines specific requirements for workplace safety committees or representatives, ensuring that workers have a voice in identifying and addressing safety concerns. The code empowers employees to refuse work that they believe to be unsafe, providing them with legal protections against retaliation from employers.

Furthermore, the Canadian Labour Code includes provisions for collective bargaining and dispute resolution, facilitating the negotiation of collective agreements between employers and unions representing workers. For instance, the code outlines procedures for the certification of unions, ensuring that they have sufficient support from employees to represent their interests effectively. Additionally, the code provides mechanisms for the resolution of labour disputes, such as mediation and arbitration, offering alternatives to strikes or lockouts that may disrupt productivity and economic activity.

### **Critical Instances of Changes in Labour Laws in India**

Many landmark judgements have led to the evolution of labour laws in both India and Canada. In the discussion portion of this paper this paper attempts to highlight 3 instances each where labour laws were directly influenced by the legal system.

#### **1. MC Mehta vs State of Tamil Nadu, 1996**

Sivakasi is a small town in the state of Tamil Nadu. It is one of the worst-hit regions in India in terms of child-labour. A lawyer, MC Mehta, distraught by the conditions faced in this town by children, filed a petition under Article 32. He argued that the employment of children was constitutionally invalid under Article 24 and Article 39(f). The court subsequently issued guidelines and amendments to combat this problem (*Sivakasi Houses World's Largest Concentration of Child Labour in Its Industrial Units*, 2013).

Via this judgement, the court expanded the definition of acts such as the Apprentices Act, and the Beedi and Cigar Workers Conditions of Employment Act. After examining all facts, the court concluded that it was the deteriorating financial condition of the family which made them push child labour on the child. In light of this, the court ordered the formation of the Child Labour Rehabilitation and Welfare fund(henceforth referred to as fund), it was ordered that any person employing children in factories illegally would have to deposit ₹20,000 for the growth and development of the child.

Moreover, the court ordered the government to make sure that an adult of the family is placed in

a proper job instead of the child having to resort to child labour. If the state is unable to provide this facility, then ₹5000 should be submitted to the fund.

To ensure proper facilitation and efficacy of such measures, the court also ordered that:

1. The government shall carry out a survey of child labour within 6 months.
2. A criterion assessing and observing the areas of employment which shall be deemed to be hazardous for employment be formed.
3. Employment should be given to the parents or any adult of the family at the same place where the child was employed.
4. In cases where employment is not provided to the parents or any adult of the family, an amount of ₹85,000 be paid to the parents of the child for the welfare of the child. In case, the parents fail to send the child for education, the payment shall cease.
5. The education of the child in an esteemed institution or facility should be ensured.
6. The inspectors of the concerned shall ensure that education is provided free of cost, as mandated under Article 45 of the Indian Constitution.

## **2. Bhopal Gas Tragedy, 1984**

The Bhopal Gas tragedy of 1984 was a chemical gas leak at a pesticide factory during the night of 2-3rd December 1984. This tragedy is often referred to as the world's worst industrial disaster due to the impact and casualties. The accident occurred at the Union Carbide India Limited (UCIL) pesticide plant in Bhopal, Madhya Pradesh. The government in fact had a 22% stake in the company. The plant was initially only approved for the production of pesticides using component chemicals, however, due the competition, led to the UCIL having to export raw materials and intermediate products to combine all production under one roof to cut costs. This development came with more risks.

At 11 pm on the 2nd of December 1984, a small leak of Methyl Isocyanate (MIC) was noticed by one of the operators. Due to a faulty valve, approximately one ton of water from the coolant which kept the MIC cool was allowed to flow inside the MIC tank. Due to this reaction and the fact that the MIC was no longer being cooled, pressure and heat started to form. The gas flare detection system had been broken for 3 months as well. At around 1 AM, the gas shot up into the Bhopal sky, killing an estimated 3,500 people immediately. Estimates of the total casualties are as high as 20,000 people.

Following this catastrophe, a new chapter was added to the Factories Amendment Act of 1948. Essentially, the amendment focused on ensuring that the risks associated with any industrial activity are made aware to the relevant bodies and communities in the vicinity of the plant. And that workers must have a right to participate in safety management and regulation of the location and expansion of hazardous activities. The chapter also sought to define the “occupier”, that is, one who is responsible for the affairs of the factory and explicitly laid down that the occupier has an obligation to show, in the event of an accident, that due diligence had been exercised to enforce the safety obligations laid down in the Act.

### **3. Vishaka and others v/s State of Rajasthan**

Vishaka and others v/s State of Rajasthan was a landmark judgement for the safety of women in the workplace. It gave birth to the Vishaka Guidelines and laid the groundwork for The Sexual Harassment of Women in the Workplace Act, 2013.

This litigation was the result of a Public Interest Litigation (PIL) filed by several women's organisations led by Naina Kapur and her organisation, Sakshi. Sakshi filed a PIL in the matter of Bhanwari Devi, a government employed social worker in Rajasthan. Bhanwari Devi tried to stop the marriage of Ramkaran Thakur's daughter, who was less than an year old at the time of marriage. As part of her duty in trying to terminate the marriage, she was brutally gang raped.

After this case the court took many actions to ensure the safety of women in the workplace. It defined this safety as a fundamental right for the women and defined the term sexual harassment. Accordingly, “any physical touch or conduct, showing of pornography, any unpleasant taunt or misbehavior, or any sexual desire towards women, sexual favor will come under the ambit of sexual harassment.” The most important step was the formation of the Vishaka Guidelines which would provide protection for women against sexual harassment until the formation of the 2013 act. The features of the guidelines are:

1. Definition of Sexual Harassment
2. Provide a safer working environment
3. Duty of the employer to file a complaint
4. Complaint redressal committee
5. Employer to assist the employee if she is sexually harassed
6. Duty of employer to spread awareness
7. Duty of government to widen the scope of these guidelines

While the section above has looked into incidents that have changed the nature of labour laws in India, the following section will now delve into the noteworthy instances that brought about significant changes in the Canadian labour laws



## **Critical Instances of Changes in Labour Laws in Canada**

### **1. Nine Hour Movement**

The Nine Hour Movement in Canada, also known as the "Nine-Hour Movement" or the "Shorter Hours Movement," was a significant labour movement that emerged in the late 19th century. It aimed to reduce the standard working day from ten hours to nine hours, advocating for improved working conditions, better quality of life for workers, and a more equitable distribution of labour.

The movement gained momentum in Canada during the latter half of the 19th century, paralleling similar movements in other industrialised nations. At the time, many workers, particularly those in manufacturing and industrial sectors, were subjected to long hours of labour in often harsh and unsafe working conditions. The push for shorter working hours was fueled by concerns about worker health and safety, as well as a desire for greater leisure time and improved quality of life.

One of the key features of the Nine Hour Movement was its emphasis on collective action and solidarity among workers. Labour unions, trade associations, and workers' organisations played a central role in mobilising support for the movement and advocating for legislative reforms. Strikes, demonstrations, and public campaigns were organised to raise awareness about the need for shorter working hours and to pressure employers and policymakers to enact change.

The movement achieved significant success in several Canadian provinces. In 1872, the Toronto Typographical Union went on strike demanding a nine-hour workday, marking one of the earliest victories for the movement in Canada. Subsequently, other unions and workers' organisations across the country joined the campaign, leading to the gradual adoption of shorter working hours in various industries and sectors.

### **2. Winnipeg General Strike of 1919**

The Winnipeg General Strike of 1919 was a pivotal moment in Canadian labour history, characterised by widespread labour unrest and collective action. The strike, which began on May 15th and lasted until June 25th, saw approximately 30,000 workers from various industries walk off their jobs in solidarity, demanding better wages, working conditions, and the right to collective bargaining. At its peak, the strike involved nearly half of Winnipeg's workforce. The catalyst for the strike was a culmination of factors including post-war economic uncertainty, rising inflation, and the inequities faced by workers, exacerbated by the harsh conditions of wartime industrial production. Despite its peaceful beginnings, the strike escalated into confrontations between strikers and authorities, leading to arrests, injuries, and even deaths. The federal government decided to step in. It was afraid the strike would spark conflicts in other

cities. Soon after the strike began, two Cabinet ministers met with the Citizens' Committee in Winnipeg. These officials were Senator Gideon Robertson, minister of labour, and Arthur Meighen, minister of the interior and acting minister of justice. Robertson and Meighen refused to meet with the Central Strike Committee.

On the Citizens' Committee's advice, the federal government swiftly supported the employers. It threatened to fire federal workers unless they returned to work immediately. Parliament changed the Immigration Act so that British-born immigrants could be deported. It also broadened the Criminal Code's definition of sedition

### **3. Rand Formula**

The Rand Formula emerged in the context of labour relations in Canada following the Ford Strike of 1945 in Windsor, Ontario. This strike, involving workers at the Ford Motor Company's plant, was a pivotal moment in Canadian labour history and highlighted the power struggle between labour unions and employers over union recognition and collective bargaining rights.

During the strike, the Ontario government appointed Justice Ivan Rand to arbitrate the dispute and find a resolution. Justice Rand's solution, known as the Rand Formula, aimed to address the contentious issue of union dues payment by all workers covered by a collective agreement.

At the time, there were concerns about "free-riding" – situations where workers benefited from the gains negotiated by the union without contributing financially to the costs associated with bargaining and representation. This undermined the financial stability of unions and their ability to effectively represent the interests of all workers in the bargaining unit.

The Rand Formula addressed this issue by mandating that all employees covered by the collective agreement – regardless of their union membership status – pay union dues or fees to the union representing them. This ensured that all workers who benefited from the terms negotiated in the collective agreement contributed their fair share to the union's financial resources.

The implementation of the Rand Formula had significant implications for labour relations in Canada. It provided a stable and predictable source of funding for unions, enabling them to sustain their activities, such as bargaining, representation, and organising efforts. It also promoted fairness and equity in the workplace by ensuring that all workers contributed to the costs associated with union representation.

The Rand Formula has since become a standard practice in Canadian labour relations and has been widely adopted in collective bargaining agreements across various industries. It represents a

compromise between the rights of individual workers and the collective interests of the workforce, promoting solidarity and stability in the labour movement in Canada.

### **Comparing Instances of Changes to Labour Laws in India and Canada**

Our comparative analysis of the factors that led to changes in the legal legislature of India and Canada reveals significant differences in how each country approaches labour laws and welfare. In Canada, changes in labour legislation often result from peaceful protests, organised strikes, and proactive government initiatives. These changes reflect a societal structure that prioritises and swiftly addresses labour welfare through well-organised channels. In contrast, India's legislative changes are frequently reactive, triggered by extraordinary incidents or tragic situations, such as the Bhopal Gas Tragedy or the MC Mehta v/s The State of Tamil Nadu case. This stark difference may point to deeper cultural divergences in how each country perceives and values labour laws.

In Canada, the importance placed on labour welfare is evident in the relatively low threshold of intensity required for legislative changes. Protests and strikes are generally peaceful and organised, allowing for swift governmental response and minimal delays in the enactment of new laws. This contrasts with India, where the intensity of a tragedy is the judge and not the intensity of the ongoing exploitation which caused the tragedy. The Bhopal Gas Tragedy, one of the world's worst industrial disasters, is a prime example of the extreme circumstances required to prompt significant legal reforms.

### **CONCLUSION**

The insights provided in this study could be useful to sociologists and psychologists to analyse the manifestation of the stark cultural differences of both countries within the scope of their legal systems. This study also provides a model to analyse further legislature and find out what formed them and why they were formed. These findings should be interpreted with caution given the multiple limitations of this study, but they encourage future studies to expand upon them.

This study has 4 potential limitations. First, due to the nature of this study, finding relevant past research which was comparative was a struggle. Nonetheless, many studies upon the legal systems of the two countries were utilised. Second, many research papers were inaccessible because of time and financial restrictions. Our study had to heavily rely on books instead of peer reviewed papers due to these restraints. Third, the analysis conducted from the sample size of 3 instances for each country may not encapsulate the overall trend over decades. Lastly, the sources were primarily second hand and third hand and not actual first hand court documents hence they may be vulnerable to biases of previous authors who have been referenced.

## REFERENCES

1. Arnold, T. J., & Britain, G. (1876). The Conspiracy and Protection of Property Act, 1875 (38 & 39 Vict. C. 86), and the Employers and Workmen Act, 1875 (38 & 39 Vict. C. 90). [http://books.google.ie/books?id=CKgXAAAAYAAJ&printsec=frontcover&dq=The+Conspiracy+and+Protection+of+Property+Act,+1875+\(38+%26+39+Vict.+C.+86&hl=&cd=1&source=gbs\\_api](http://books.google.ie/books?id=CKgXAAAAYAAJ&printsec=frontcover&dq=The+Conspiracy+and+Protection+of+Property+Act,+1875+(38+%26+39+Vict.+C.+86&hl=&cd=1&source=gbs_api)
2. Davis, J.E. (1875). The Labour Laws. [http://books.google.ie/books?id=jrYDAAAQAQAAJ&printsec=frontcover&dq=The+Labour+Laws+davis&hl=&cd=1&source=gbs\\_api](http://books.google.ie/books?id=jrYDAAAQAQAAJ&printsec=frontcover&dq=The+Labour+Laws+davis&hl=&cd=1&source=gbs_api)
3. Doorey, David J., Can't Get There from Here: Thoughts on the Idea of Labour Law Reform in the 21st Century (January 26, 2024). Available at SSRN: <https://ssrn.com/abstract=4707810> or <http://dx.doi.org/10.2139/ssrn.4707810>
4. Gray, R. (2002). The Factory Question and Industrial England, 1830-1860. Cambridge University Press. [http://books.google.ie/books?id=72y1SL\\_2E1oC&printsec=frontcover&dq=The+Factory+Question+and+Industrial+England,+1830%E2%80%931860&hl=&cd=1&source=gbs\\_api](http://books.google.ie/books?id=72y1SL_2E1oC&printsec=frontcover&dq=The+Factory+Question+and+Industrial+England,+1830%E2%80%931860&hl=&cd=1&source=gbs_api)
5. GV, R. P. (n.d.). A Study On The Socio-Economic Problems Faced By The Women Workers In The Unorganized Manufacturing Sector: Evidence From Ernakulam District. Manpower Journal, LVII.
6. Heritage, C. (2022, September 15). Rights in the workplace. Canada.ca. <https://www.canada.ca/en/canadian-heritage/services/rights-workplace.html>
7. *How International Labour Standards are used* | International Labour Organization. (2024, January 28). <https://www.ilo.org/international-labour-standards/how-international-labour-standards-are-used>
8. Jadhav, A. D. (2019). The Role of British Legislations and the Working Class Movement in Bombay: A Historical Study of the Factory Acts of 1881 and 1891 in India. SOCIAL REVIEW. International Social Sciences Review Revista Internacional De Ciencias Sociales, 1, 1–10. <https://doi.org/10.37467/gka-socialrev.v1.1965>
9. John, J. (2003). Labour scene in India. International Union Rights, 10(4), 9–11. <http://www.jstor.org/stable/41937317>
10. K Gayithri & Malini L Tantri & D Rajasekhar, 2019. "A critical review of apprenticeship policy of India," Working Papers 440, Institute for Social and Economic Change, Bangalore.
11. Kuddo, Arvo. Labor Regulations throughout the World : An Overview (English). Washington, D.C. : World Bank Group. <http://documents.worldbank.org/curated/en/221471546254761057/Labor-Regulations-throughout-the-World-An-Overview>

12. Lee, S., McCann, D. and Torm, N. (2008), The World Bank's "Employing Workers" index: Findings and critiques – A review of recent evidence. *International Labour Review*, 147: 416-432. <https://doi.org/10.1111/j.1564-913X.2008.00043.x>
13. Menon, S., & Majumdar, I. (2005). India: labour law focus. *International Union Rights*, 12(3), 12–13. <http://www.jstor.org/stable/41937379>
14. Morton J. Horwitz History of the Public/Private Distinction, 130 U. Pa. L. Rev. 1423 (1982). Available at: [https://scholarship.law.upenn.edu/penn\\_law\\_review/vol130/iss6/6](https://scholarship.law.upenn.edu/penn_law_review/vol130/iss6/6)
15. Muthusamy, A., & Ibrahim, D. (2016). PROBLEMS FACED BY INFORMAL WORKERS IN DIFFERENT SECTORS IN INDIA. *Indian Journal of Applied Research*, 25–28. [https://www.worldwidejournals.com/indian-journal-of-applied-research-\(IJAR\)/special\\_articles.php?si\\_val=MjY0](https://www.worldwidejournals.com/indian-journal-of-applied-research-(IJAR)/special_articles.php?si_val=MjY0)
16. NCIB. (2021). Labour Laws In India. In Index (pp. 3–196). [https://ncib.in/pdf/ncib\\_pdf/Labour%20Act.pdf](https://ncib.in/pdf/ncib_pdf/Labour%20Act.pdf)
17. Noonan, J. T. (1962). THE CONCEPT OF LAW. By H. L. A. Hart. Oxford: Oxford University Press, 1961. Pp. viii, 263. 21s. *The American Journal of Jurisprudence*, 7(1), 169–177. <https://doi.org/10.1093/ajj/7.1.169>
18. Papola, T. S. (1980). Informal Sector: Concept and Policy. *Economic and Political Weekly*, 15(18), 817–824. <http://www.jstor.org/stable/4368628>
19. Robertson, G. (2024, July 11). Crimes Against Humanity. <https://www.penguin.co.uk/books/54397/crimes-against-humanity-by-geoffrey-robertson/9781802063240>
20. Sack, J. (2010). U.S. and Canadian Labour Law: Significant Distinctions. *ABA Journal of Labor & Employment Law*, 25(2), 241–258. <http://www.jstor.org/stable/40862947>
21. Saini, D. S. (1999). Labour Legislation and Social Justice: Rhetoric and Reality. *Economic and Political Weekly*, 34(39), L32–L40. <http://www.jstor.org/stable/4408454>
22. Sodhi, J. S. (2014). Labour Law Reform in India. *Indian Journal of Industrial Relations*, 50(1), 100–117. <http://www.jstor.org/stable/24547025>
23. Tokman, V.E. (1979). An Exploration into the Nature of Informal—Formal Sector Relationships. In *Elsevier eBooks* (pp. 1065–1075). Available at: <https://doi.org/10.1016/b978-0-08-024270-5.50007-7>
24. *Sivakasi houses world's largest concentration of child labour in its industrial units.* (2013, July 30). India Today. <https://www.indiatoday.in/magazine/special-report/story/19830115-sivakasi-houses-world-s-largest-concentration-of-child-labour-in-its-industrial-units-770353-2013-07-30>
25. Upadhyaya, S., & Kumar, P. (2017). Amendments in Labour Laws and Other Labour Reform Initiatives Undertaken by State Governments of Rajasthan, Andhra Pradesh, Haryana and UP: An Analytical Impact Assessment. VV Giri National Labour Institute.