

*The*

---

**PIONEER RESEARCH**

---

*Journal*

An International Collection of  
Undergraduate-Level Research

**Volume 7**

2020

**Pioneer**<sup>®</sup>  
academics

Copyright 2021, Pioneer Academics, PBC, all rights reserved

**Publisher**

Hong Kong Digital Publishing Ltd.  
1501, Grand Millennium Plaza, 181 Queen's Road, Hong Kong

ISBN 978-988-77065-8-8

Pioneer Academics, PBC

Email: [info@pioneeracademics.com](mailto:info@pioneeracademics.com)

Website: [www.pioneeracademics.com](http://www.pioneeracademics.com)

*The Pioneer Research Journal* is published annually by Pioneer Academics, PBC. The 2020 issue of *The Pioneer Research Journal* is Volume 7.

Copyright ©2021, by Pioneer Academics, PBC, 30 S. 15th Street, 15th Floor, Philadelphia, PA 19102, USA. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

# Contents

---

<i>Foreword</i> .....	<i>ix</i>
<i>Selection Process</i> .....	<i>xi</i>
<b>Lebanon: An IMF Case Study (International Relations)</b> .....	<b>1</b>
<b>Author:</b> Abhishek Don Hemlani	
<b>School:</b> BASIS International School Guangzhou – Guangzhou, China	
<b>Pioneer Research Concentration:</b> Political Risk and Prediction	
<b>Diversity vs. Inclusion in STEM: Factors that Accurately Measure a Tech Company’s Commitment to Inclusion (STS)</b> .....	<b>14</b>
<b>Author:</b> Adrienne Lai	
<b>School:</b> St. Stephen’s and St. Agnes School – Alexandria, Virginia, United States	
<b>Pioneer Research Concentration:</b> Digital Mapping	
<b>Not so Optimal: The Evolution of Mutual Information in Potassium Voltage-Gated Channels (Biophysics)<sup>1</sup></b>	
<b>Author:</b> Alejandra Duran Urriago	
<b>School:</b> Colegio Santa Francisca Romana – Bogotá, Colombia	
<b>Pioneer Research Concentration:</b> Biological Sensors How and How Well do They Work?	
<b>Can Corporations Balance Economic Profit with Corporate Social Responsibility? (Economics)</b> .....	<b>42</b>
<b>Author:</b> Annapurna Ananya Maddali	
<b>School:</b> The Quarry Lane School – Dublin, California, United States	
<b>Pioneer Research Concentration:</b> The Politics of the Gig Economy: Entrepreneurship, Technological Innovation, and Politics	

---

<sup>1</sup> Alejandra Duran Urriago’s paper is currently under consideration for another journal, and she has therefore decided to forego publication of her paper in *The Pioneer Research Journal*.

<b>Bioethics: To What Extent Should Humanity Perceive Eugenics as Applicable? (Philosophy)</b> .....	58
<b>Author:</b> Bangjun Yue	
<b>School:</b> Institut Le Rosey – Rolle, Switzerland	
<b>Pioneer Research Concentration:</b> Bioethics	
<b>What Is Wolf: Narratives about Human-Wildlife Conflict In Tibetan China (Anthropology)</b> .....	74
<b>Author:</b> Chuhan Zhang	
<b>School:</b> Nanjing Jinling High School International Department – Nanjing, China	
<b>Pioneer Research Concentration:</b> Communication and Culture	
<b>Exploring Neurotrophin Therapy as a Treatment for Rehabilitating Victims of Childhood Abuse (Neuroscience/Psychology)</b> .....	93
<b>Author:</b> Dojin Choi	
<b>School:</b> Coppell High School – Coppell, Texas, United States	
<b>Pioneer Research Concentration:</b> Understanding the Sense of Touch	
<b>“¿Que Dijo?” Adolescent Feelings Towards Language Brokering and Its Effects on Family Conflict (Psychology)</b> .....	126
<b>Author:</b> Ethan Estrada	
<b>School:</b> Liberal Arts and Science Academy High School – Austin, Texas, United States	
<b>Pioneer Research Concentration:</b> Psychology of Immigration	
<b>Coastal Plastic Management as a Response to Rafting Invasive Species (Environmental Studies)</b> .....	140
<b>Author:</b> Jiaying Liu	
<b>School:</b> Indian Springs School – Birmingham, Alabama, United States	
<b>Pioneer Research Concentration:</b> Our Planet’s Plastic Plight	
<b>Semi-Supervised Classification with Generative Adversarial Networks on Medical Datasets with Limited Size and Label (Computer Science/Biology)</b> .....	165
<b>Author:</b> Lutfi Eren Erdogan	
<b>School:</b> Uskudar American Academy – Istanbul, Turkey	
<b>Pioneer Research Concentration:</b> Biology in Computing: Why Are Networks Neural?	



<b>Search Asymmetry Manipulation with a Common Feature</b> (Neuroscience/Psychology).....	181
<b>Author:</b> Maanasi Nair	
<b>School:</b> The International School Bangalore – Bangalore, India	
<b>Pioneer Research Concentration:</b> Vision Impossible: An Introduction to the Psychology and Neuroscience of Seeing	
<b>Complicated Implications of Simple Images: Schizophrenic Patients’ and Their Families’ Views of MRI and fMRI in the United States from 2000-2020</b> (History/STS).....	190
<b>Author:</b> Mengyao (Angel) Yang	
<b>School:</b> San Domenico High School – San Anselmo, California, United States	
<b>Pioneer Research Concentration:</b> History of Medical Technologies	
<b>Analysis of Testing Accuracy Threshold for COVID-19 Through a Computational SIR Model</b> (Computer Science).....	205
<b>Author:</b> Nabo Yu	
<b>School:</b> The Webb Schools – Claremont, California, United States	
<b>Pioneer Research Concentration:</b> Networks: From Searching the World Wide Web to Slowing the Spread of Infectious Diseases	
<b>Cube-Plate Tiling Numbers and Their Identities</b> (Mathematics).....	227
<b>Author:</b> Qianyu Guo	
<b>School:</b> YK Pao School – Shanghai, China	
<b>Pioneer Research Concentration:</b> Fibonacci Numbers and Visual Proofs	
<b>Uncovering the Invisibility: Protecting LGBT Groups in Armed Conflict and After</b> (Sociology/Political Science).....	238
<b>Author:</b> Qing Wan	
<b>School:</b> Nanjing Jinling High School International Department – Nanjing, China	
<b>Pioneer Research Concentration:</b> Gender and Political Conflict	

- The Relationship Between the Anti-Inflammatory Strength of Butyrate and the Progression of Colorectal Cancer (Biology)**.....253  
**Author:** Rohan Kartha  
**School:** Westwood High School – Austin, Texas, United States  
**Pioneer Research Concentration:** The Gut Microbiome
- Ship of Theseus: Looking at Hagia Sophia’s Cultural Complexity in the 21<sup>st</sup> Century (Art History/Architecture)**.....265  
**Author:** Ruoxuan Fang  
**School:** Miss Porter’s School – Farmington, Connecticut, United States  
**Pioneer Research Concentration:** The Medieval Mediterranean: Confluence of Cultures
- Exit, Voice, and Violence: Contemporary Responses of Non-Hindu Groups to the BJP’s Hindu Nationalism (Political Science)**.....286  
**Author:** Rushil Roy  
**School:** Monte Vista High School – Danville, California, United States  
**Pioneer Research Concentration:** Big Questions in Comparative Politics: Nationalism, Modernity, the Social Question
- “Ya me cansé”: The Role of International NGOs in the 2014 Ayotzinapa Massacre (History)**.....305  
**Author:** Seo Yoon (Yoonie) Yang  
**School:** Signal Mountain Middle High School – Signal Mountain, Tennessee, United States  
**Pioneer Research Concentration:** Political Protest in Latin America
- Intangible Assets: The Cause of Veblen Goods – A Breakdown of the Intangible Factors that Drive Pricing Premiums (Business)** .....316  
**Author:** Vardaan Tekriwal  
**School:** Jakarta Intercultural School – Jakarta, Indonesia  
**Pioneer Research Concentration:** Financial Accounting & Analysis: Capital Markets, Entrepreneurship, and Decision Making

- The *Danmu* Register and the Establishment of its Social Existence in a Youth-Centered Speech Community in China** (Anthropology).....335  
**Author:** Xinlin Yu  
**School:** International Department of the Affiliated High School of South China Normal University – Guangzhou, China  
**Pioneer Research Concentration:** Communication and Culture
- How Images Influence People’s Understanding of Science in the Context of COVID-19** (Neuroscience/Psychology).....361  
**Author:** Xinyi Tian  
**School:** Beijing Academy – Beijing, China  
**Pioneer Research Concentration:** Vision Impossible: An Introduction to the Psychology and Neuroscience of Seeing
- The Influence of Government Intervention on Sports Participation for School-Aged Children in China Between 2011 and 2020** (History/Sociology).....378  
**Author:** Xinyue (Abby) Luo  
**School:** International Department of the Affiliated High School of South China Normal University – Guangzhou, China  
**Pioneer Research Concentration:** Only a Game? Players and Spectators: History, Culture, and Sports
- The Slippery Slope: Naturalistic Observations into Non-Verbal Manifestations of Formed Dominance Hierarchies Among Children at Play** (Anthropology).....395  
**Author:** Xuan Jiang  
**School:** The Experimental High School Attached to Beijing Normal University – Beijing, China  
**Pioneer Research Concentration:** Nonverbal Communication
- Nietzsche’s *Agon* with Plato: Truth, Morality, and Modernity** (Philosophy).....412  
**Author:** Yilong Wu  
**School:** University School of Nashville – Nashville, Tennessee, United States  
**Pioneer Research Concentration:** The Examined Life

- The Roles of Extracellular Vesicles in Cancer and the Investigation of Exosomal PD-L1 in Immunotherapy of Colorectal Carcinoma (Biology).....437**  
**Author:** Yuanmuhuang Long  
**School:** High School Affiliated to Shanghai Jiao Tong University – Shanghai, China  
**Pioneer Research Concentration:** Cell Biology
- Types of Graphs Generated by Pascal’s Simplices and a Problem Proposed by Tegua and Godbole (Mathematics).....455**  
**Author:** Zizheng Fang  
**School:** Hangzhou Foreign Languages School – Hangzhou, China  
**Pioneer Research Concentration:** Combinatorics and Graph Theory

# Foreword

---

In Volume 7 of *The Pioneer Research Journal*, more than two dozen young scholars from around the world engage with topics of material, philosophical, and ethical importance. From diversity and inclusion in STEM to information literacy about COVID-19, from eugenics to environmental protection, the work featured in this volume affirms that scholarly inquiry is key to understanding our world and solving the most vexing problems of modern life. In the pages that follow, some of our brightest young minds make substantial contributions to important ongoing conversations in the academic and public spheres, and I am honored to be able to share their compelling work in this journal.

This multidisciplinary volume contains exceptional scholarship by high school students from schools on four continents. In 2020, 974 young scholars from 43 countries and regions conducted research through Pioneer Academics, selected from an international pool of 3,487 applicants. Those admitted to the program participated in an international faculty-led cohort before working one-on-one with leading U.S. professors to conduct original undergraduate-level research in their area of interest. After a rigorous nomination and double-blind review process, 27 papers were selected for publication in this edition of the *Pioneer Research Journal*. The authors represented herein are from schools in China, Colombia, India, Indonesia, Switzerland, Turkey, and the United States.

By conducting the program entirely online, Pioneer Academics has torn down barriers and made undergraduate-level education available to promising young scholars in virtually every corner of the world. To ensure this opportunity is available to as many students as possible, Pioneer provided US \$1.1 million in need-based scholarships, again meeting 100% of demonstrated financial need and fulfilling their mission to remove obstacles to educational access for the most deserving underserved students.

The papers in this edition reflect the rigor and high academic standards that define Pioneer Academics scholars, demonstrating unequivocally what high school students can achieve when given the right tools and encouraged to embrace their intellectual curiosity. Moreover, this journal is a testament to the resilience and tenacity of these dedicated young scholars, who have persisted in the pursuit of academic excellence during a period of prolonged pandemic uncertainty and adversity.

It is our goal to share this work widely, and so the *Pioneer Research Journal* is available in print and online at [www.pioneeracademics.com](http://www.pioneeracademics.com) and is distributed to select colleges, universities, and libraries worldwide.

I am so pleased to share it with you and hope you find it to be inspiring and enlightening.

**David G. Kamitsuka, Ph.D.**

*Dean of the College of Arts and Sciences  
Oberlin College*

*Editor, The Pioneer Research Journal, Volume 7*

## Selection Process

---

To be nominated for publication in *The Pioneer Research Journal*, Pioneer scholar-authors first had to earn a letter grade of A- or higher from their professor mentor. Each professor mentor was invited to nominate one or two papers (depending on the number of scholars they mentored) that met the A/A- grading threshold and that, in their estimation, represented the highest caliber of scholarship from among their mentees.

Following nomination, every paper was reviewed anonymously by a member of our 47-member committee of reviewers, each of whom is a professor with expertise in the subject area of the papers they reviewed.

Contributing Readers scored the papers on a scale of 0 – 4 for four criteria: Engagement with Scholarship, Evidence and Analysis, Writing and Organization, and Scholarly Contribution. These four scores were then tallied, with a maximum score of 16. Papers with sub-scores of 0 in any category were disqualified from publication.

Once the above thresholds were met, the highest-scoring papers were provisionally selected. Upon notification, authors of provisionally selected papers received instructions from their paper's reviewer specifying the revisions required for meeting the standards of an undergraduate research journal. Every paper then underwent one more rigorous review by Pioneer's Writing Center tutors for editorial concerns and the authors were asked to make final revisions based on that review before their papers were finally published.

When all criteria were applied, only 2.7% of the papers generated in the 2020 Spring-through-Summer and Summer-Only terms were selected for publication in the Seventh Volume of *The Pioneer Research Journal*.

# Lebanon: An IMF Case Study

---

*Abhishek Don Hemlani*

**Author Background:** *Abhishek Don Hemlani grew up in China and currently attends BASIS International School Guangzhou in Guangzhou, China. His Pioneer research concentration was in the field of international relations and titled “Political Risk & Prediction.”*

## Introduction

The August 4<sup>th</sup> explosion of 3,000 tonnes of ammonium nitrate in the port of Beirut killed 200 and caused 300,000 people to become homeless, costing the government an estimated \$10-15 billion (BBC 8). The explosion was caused by the unsafe storage of chemicals, and experts have referred to this as one of the “largest non-nuclear explosions in history” (BBC 15). The Diab Administration resigned promptly, blaming the country’s “entrenched political elite”(BBC 36), but Lebanese citizens remain unsatisfied. This cycle of negligence is not new to Lebanon; political mismanagement occurs frequently. The root of the current crisis is grounded in its fractured political system, giving rise to negligence and corruption. It then manifested in poor economic decisions that led to a hyperinflation crisis. The time frame used for the analysis of this crisis is until a bailout plan for financial assistance is agreed to by Lebanon and the IMF, since negotiations have stalled. Former Lebanese governmental officials mainly agree that an IMF bailout is the only solution (Nakhoul 1), with Lebanon seeking \$10 billion in aid (Reuters 1).

This report will analyze the situation in Lebanon and identify critical areas that the IMF should tackle in its overall economic assistance and explore the feasibility of possible solutions to develop further an action plan used in negotiations with the Lebanese government. It will analyze Lebanon’s political, social, geopolitical, and economic risks, and will determine that while the financial crisis is the most imminent, the cause of the issues is political negligence and mismanagement.

## Economic Risk: High

### Overview

Lebanon’s annual GDP is around \$53 billion, with an average GNI per capita of \$7600 (World Bank 2); Lebanon’s economy is characterized by minimum government intervention in private enterprises and a tax-free system (Khalaf and Kingston 1). A deficit existed with high imports and low exports, but elements such as increased tourism and remittances and Lebanon’s position as a banking power helped balance the budget (Khalaf and Kingston 1). Overall, income began to rise

and Lebanese goods found a marketplace worldwide. As a result of Covid-19, and pre-existing economic issues, economic growth is predicted to decrease -11% in 2020 (World Bank 2). Moreover, conditions worsened when the debt-to-GDP ratio reached 150% (Khalaf and Kingston 4). A 2017 McKinsey's analysis reveals that Lebanon relies on low domestic savings and external resources to account for its deficit (Athi 32). This reliance on external sources, a consequence of its government and banking system, could lead to its downfall.

### **Government Ponzi Scheme**

To perpetuate the illusion of economic stability, the government masked its massive debt through unsustainable and mismanaged fiscal decisions. The Banque du Liban (BdL), the central bank, was a significant player and could control the scheme through capital controls, account freezes, and false declarations (Rickards 28), suggesting corruption at the highest level. The government financed itself by selling bonds in dollars, and sometimes in Lebanese pounds, giving the impression that it was managing its fiscal deficit with long-term options (Sester 9). Instead, it sold bonds to commercial banks, which raised money from outside Lebanon by offering a very high-interest rate for depositing dollars (Sester 11). Those deposit inflows and Eurobond purchases covered any fiscal deficit and led to foreign currency debt (Sester 12). In essence, this system led to Lebanese citizens depositing short term dollars into the central bank, which was financing Lebanon's accrued budget deficits. The scheme failed once the deposits stopped. This system worked because the commercial bank's pay-receive ratio was as high as 11% and made a considerable profit (Sester 19). With this scheme, bank and government officials clearly profited through unsustainable, short-term decisions leading to economic downfall. Currently, Lebanon's debt is short-term, high-cost, and has an increasing interest, making it extremely unstable. Consequently, the Lebanese economy has fallen into huge sovereign debt, which has had a crippling effect. This scheme is an example of the government working against its decisions through poor economic choices. This system is unlikely to change without pressure from citizens and external sources, such as the International Monetary Fund (IMF). Because of the nature of the debt and the banks, the Eurobonds, the central bank, and government needs to be restructured, which poses massive uncertainty, liability, and international contribution. This unstable debt requires urgent IMF intervention and action, and further auditing to assess the level of corruption present. The IMF's involvement would require a complete review and restructuring of the government, central and commercial bank relationships, which will be time and resource constraining. It would require the reshaping of the Lebanese economy - a difficult task. The sheer scale of the reform would require a comprehensive plan with direct political cooperation, which is unlikely.

### **Difficult Reform Plan**

The likelihood of recovery for Lebanon is slim due to massive disagreement on the nature of the bailout. The IMF estimates debt losses to be between \$93 billion and \$100 billion, but the government's estimate is much less (Rickards 15). Furthermore, there is no consensus over distributing losses equitably with depositors, creditors, government, and banks (Rickards 16). Achieving this would likely require strong political cohesion. The foundation for rescue does not exist, as Lebanon has no hard currency and no agreement on transferring assets (Rickards



16). Without sizable concessions by the government and reform, the chance of recovery is slim. Moreover, other potential options, such as solving the crisis through economic diversification and exploitation of its vast oil and gas reserves, have been proposed. However, they are unlikely to help in the long-term, and any kind of grants from other countries will not help balance the budget (Sester 22). The potential solution of diversification makes it impossible to end the crisis through economic growth alone (Atl 8). The answer seems to be external aid, where \$30 billion from the IMF, the Center for Economic Development and Reform through Enterprises (CEDRE), and the World Bank could be the maximum loan. The loan would require austerity, massive monetary reform, legal reform, anti-money laundering, and widespread banking reform to be effective. This reform is extremely unlikely given the position of the Hezbollah and the current political fracturing, as expanded upon in the geopolitical and political sections. Without any legitimate course of action, Lebanon is trapped in this cycle of debt and massive uncertainty. The political uncertainty is an enormous hurdle for the IMF, as it involves developing an unprecedented reform plan with potentially massive loans of \$100 billion. This risk factor fits the IMF's mission statement described as helping countries achieve successful reform. For any IMF advice and recommendation to be successful, it would require them to aid Lebanon in achieving excellent transparency and political cohesion, which might be difficult because the IMF's mission statement does not allow for interference in a country's domestic policies (Rickards 17). Given the pressing nature of Lebanese people's problems, it is also imperative that a plan is agreed upon quickly.

### **Citizen Protests**

The massive hyperinflation of the Lebanese currency, caused by the debt crisis and socioeconomic problems brought by Covid-19, has led to the decimation of the luxurious Lebanese lifestyle and triggered massive protests over government accountability. The World Bank pre-Covid-19 projected 50% of the population to be under the poverty line (2), and the current Social Affairs Minister estimates that 75% of the people need financial assistance because of the dwindling value of the Lira and the dire conditions of Covid-19 (Qiblawi 4), highlighting the urgency of the situation. Additionally, as three different exchange rates are used (official, black market, bank rate), and the fact that withdrawals are limited due to currency controls, paying for essential services has become a challenge for Lebanese citizens. They cannot afford bare essentials (Qiblawi 5). This inability displays the economy's dire straits and the necessity for targeting currency to bring broader stability. In light of this situation, the government's proposed financial and food assistance plans were unsuccessful due to political disagreement. Much of the burden has fallen to local NGOs such as the Lebanese Food Bank, funded by donations and thus cannot compensate for lack of governmental support (Qiblawi 15). This situation calls for IMF intervention to aid in currency stabilization and the creation of greater government accountability, which would increase the dependability of government programs. Thus, in order for broader economic reforms to affect the public, the IMF would have to directly work to increase government transparency, allowing for the correct allocation of funds and programs. If this were not accomplished, it would increase poverty rates and lead to hunger and discontent, which would prove disastrous for the country.

Lebanon's economic risk is exceptionally high because of a lack of clarity on the potential recovery and the population's desperate conditions. The solution is

contingent upon greater government reliability and accountability. Many of the debts can be restructured to allow for delayed payments and further austere measures. The IMF must take the citizen component into account because the debt crisis directly affects them.

## **Political Risk: High**

### **Overview**

A sectarian power-sharing system has governed Lebanon since its independence from France in 1943. According to Freedom House, Lebanon ranks 44/100 and is partly free, and with scores  $\frac{1}{4}$  in electing heads of government fairly, suggesting an evident deficiency (4). The head of state must be a Maronite Christian, the prime minister, a Sunni Muslim, and the speaker of parliament a Shi'ite Muslim (Freedom House 5). The system's critics assert that former warlords and sectarian leaders were motivated by regional agendas and self-interest, wherein the main parties are the Sunni, Shi'ite, Druze, and Christian (Perry 3). As a result, Lebanon's political participation is controlled by religion and patronage systems, giving rise to a government working against its citizens. This exclusion prevents real reform from taking place.

### **Creation of a Political Elite**

Political candidates campaign almost exclusively to citizens from the same sect (Ivanescu 5). These candidates may focus on issues extending to Lebanese people, but they argue for a secure representation of their sect, further highlighting religious divisions. Based on this, an analysis of voter-registration lists in 2016 proves that parliamentary seats for each religion did not match the number of voters from each faith, reinforcing the fact that Lebanon has not conducted a census since 1932 (the Economist 5). Furthermore, voters select candidates belonging to their sect, and this not only limits voters' choices but also creates patronage systems that only serve the interests of politicians (Ivanescu 7). These systems are dominated by powerful families and increase political disagreement, becoming extremely resistant to change, and resulting in the exclusion of many people, including women and LGBTQ activists from a political platform (Ivanescu 5). Due to the massive inequalities in politics, the IMF must address this system since the financial policymaking and advice could otherwise be wasted. The lack of representation for citizens would mean the incumbent system's continuation, which gives rise to negligence and corruption. Moreover, the continuation of regional agendas would mean excluding sects, which could give rise to further economic disparity, something the IMF should acknowledge. This solution could work, but the political elite will likely be unwilling to surrender.

### **Corruption and Negligence**

The patronage systems also promote corruption and promote negligence, meaning that government programs designed to provide aid and social welfare have been withheld (Ivanescu 9), thus ensuring the domination of political elites. Lebanon ranks 137/180 in Transparency International's corruption perception index.

According to a report by Transparency International on North Africa, 87% of Lebanese surveyed believe that the government is not doing enough to tackle corruption (16), which is the highest in the region. Additionally, the government also has the highest overall bribery rate of 41% and a high wasta rate of 54% (Transparency International 16). Wasta, as defined by Transparency International, is the use of personal connections to gain access to public services, that the government would otherwise provide (4). Similarly, vote-buying is also highest in the region at 47% (Transparency International 16). This data proves that the Lebanese government is too corrupt, and that corruption likely contributes to its negligence. Furthermore, the waste generated from the power-sharing costs 9% of GDP a year (Economist 4). Another criticism of the power-sharing system is its vulnerability to human rights violations, highlighting negligence. As a result of politics being associated with ethnic or religious influence, it can easily result in “relapses to mass violence or constant low-level discrimination and disenfranchisement” (Ivanescu 7). As a result, parties in this system tend to work against each other instead of in the common interest. The IMF must acknowledge this to facilitate economic self-sufficiency that has equal benefit. A 2005 World Bank study of the health sector found that Lebanon needed 25% more funds than countries with good practices (Economist 4), and the transport department has a budget of \$8 million a year (5), which is ineffective. The IMF must address the lack of funds directed to vital public services, as political negligence leads to poor spending decisions. The lack of available service funds is critical because corruption-permitted neglect creates further inequality and contributes to Lebanon's declining state. Should this continue, it would exacerbate the social crises present, discussed in social risk, and legitimize the debt crisis. Interference in the system is impossible, but the IMF must negotiate with Lebanese officials based on broader change, as the requirements for any loan, which could ensure the development of transparency and accountability.

### **Citizen Protest**

As a result of massive discontent over the political system, Lebanese citizens have been protesting under a united front, as opposed to protesting for each sect, which was the case until recently (Ivanescu 7). According to Amnesty International, they are protesting under demands such as an end to corruption, the sectarian political system, along with the recovery of funds, upholding fair trial, and financial procedures (1). The protests have been met with unlawful force from security forces that used tear gas and metal pellets (Page and Majzoub 4). Thus, it is apparent that the Lebanese people are united against the crisis and are desperate for reform to a failing system. The IMF should acknowledge the role of the citizens. If a strong government was to be instituted, citizens must consent to it to develop Lebanon's self-sufficiency and enact broader economic and social reforms. However, since there is no healthy democracy in Lebanon, it might be challenging to ensure the IMF's neutrality when interfering.

### **Potential Elections (Shock Event)**

Part of the protestors' demands includes holding elections to elect better leadership, which might be unfeasible according to an interview with Professor of Political Science at Lebanese American University, Dr. Bassel Salloukh. Furthermore,

holding new elections under caretaker Prime Minister Hassan Diab will not lead to a radical transformation of the political system (Neumann 3). This fact is due to sectarian political parties' unwillingness to give up power or accommodate any change in the electoral law. Moreover, he suggests that the economic crisis, progressive electoral law, and increased devastation will not affect sectarian parties' power in parliament (Neumann 5). Even in the event of a progressive election as an answer to the economic crisis, the election will be uneventful and will not cause drastic change. The IMF should consider this possibility and push for a government that is willing to enact economic reforms, without compromising its neutrality.

In conclusion, the inherent political problems that have led to Lebanon's economic and social repercussions must be addressed in order to develop Lebanon's future economic self-sufficiency.

### **Social Risk: High**

Lebanon's most pressing social problems are due to its political and economic fallout. Ever since the onset of the Syrian Civil War in 2011, Lebanon has had to bear the brunt of refugees fleeing the country. The government estimates that 1.5 million Syrian refugees are currently in Lebanon, which is the highest per capita ratio in the world, and about a quarter of the Lebanese populace (Fordham 1). Moreover, Lebanon has been disproportionately affected by Covid-19, with many of its hospitals being unable to cope. Additionally, the sectarian division has worsened this condition through inequality and deficiency of services, further adding to social problems.

### **Poor Conditions for Refugees**

Syrian refugees have been exposed to terrible living and working conditions and are greatly affected by the broader economic and social problems facing Lebanon. According to a 2020 Al Jazeera report on the state of refugees, more than 75% of Syrians live below the poverty line (Kranz 7). Additionally, because of inflation and massive bank restrictions, they can no longer access aid from the United Nations, leaving them with almost no source of income (Kranz 6). Similarly, government assistance has not reached them due to negligence (Kranz 8). These conditions are the result of prior neglect, as refugees have very little protection under the Lebanese government, which is not a signatory of the 1951 Refugee Convention that grants the right to work, security, justice, and provision of travel documents (Freedom Fund 4). As a result, except for a few people with a valid work permit, the large majority of Syrian refugees are unable to work or access life-saving healthcare, and hundreds of thousands of Syrian refugee children are denied access to education (Freedom Fund 4). These refugees are also extremely vulnerable to trafficking or exploitation, and harassment from security forces, which the government's 2015-2016 plan to support refugees failed to consider (Freedom Fund 5). Since this plan supports the longer-term strategic framework for 2017-2020, it is likely that the plan's contents will not drastically improve. Pre-existing negligence, coupled with recent discrimination, has only made the situation worse. The IMF must act to preserve the support for this group, such as through international aid, without interference. The IMF has to account for this vulnerable group, ensuring that it receives the required assistance as part of economic reform, and that further economic reforms do not

disproportionately affect them, as part of developing a larger economic model for Lebanon. Nonetheless, the IMF might not change official attitudes towards the refugees due to the sectarian system, which is against refugees.

### **Sectarian Division**

The divisions caused by the sectarian political system create a means for political leaders to engage in clientelism and nepotism, which has become widespread. According to a report by Foreign Policy, “it’s not really about religious identity” (Collard 15). However, the government does not provide essential services such as education and healthcare, and is instead controlled by different sects. Systems like this have led to massive inequality, where the wealthiest 0.1% of the population, around 3,000 people, earns 10% of the total national income, which is what the bottom 50% makes (Collard 18). A notable example of sects controlling the public sector is the electricity crisis, where blackouts have become frequent. The energy shortage is due to under-investment in power plants, arising from negligence, and occurring as a result of the political system (Whewell 14). However, many people resort to running illegal private generators during blackouts, and many owners are former sect militia leaders who defended their sects in clashes (Whewell 22). In return for support during the election, the government will provide these militia leaders with electricity. Thus, the government is playing a major role in perpetuating economic division. The IMF must tackle this, as further ignorance of public infrastructure could severely cripple Lebanon economically, and strengthening infrastructure would help increase government accountability. The IMF should account for any underlying government agendas when it helps create investment plans to revitalize the economy and spur growth, which the political system would hamper.

### **Covid-19 Effect**

The Lebanese government’s inability to adequately respond to the pandemic has contributed to its crisis. The spokeswoman for Lebanon’s medical supply importers claimed that, due to import restrictions, there were no masks or gloves, and importers only brought \$10 million of the \$120 million in goods they ordered (Page and Majzoub 32). Medical supply importers cited a shortage of dollars and government negligence that restricted money transfers (Page and Majzoub 33). This loss is due to government negligence and is evident with the \$1.3 billion owed to private hospitals (Page and Majzoub 30). The virus has negatively impacted Lebanon and will continue to do so as a result of government mismanagement. The effects of the virus also translate into discrimination, as a Deutsche Welle report notes that refugees have to face increased lockdown measures because Lebanese people assume Syrians are likelier to have symptoms because of a “presumed lack of hygiene and the cramped living quarters of the camps” (DW 12). Lebanese President Michel Aoun cited the Syrian refugee crisis as one of three reasons for Lebanon’s state of affairs (DW 13). Despite this, there have been no infections among Syrian refugees (DW 14), which supports unnecessary restrictions. Therefore, the government is unnecessarily discriminating against an already marginalized group of people. Also, in the event of Covid-19 symptoms, Syrians will likely avoid going to medical centers out of fear of deportation because of an unrecognized legal status (DW 20). The government can do more to help this marginalized group by

combating the virus. Thus, the economic mismanagement of funds and resources, along with political agendas of discrimination, has led to an ineffective response. This mismanagement, coupled with lax public sector services, and Lebanon will not withstand increasing infections. The IMF must recognize this and support effective fund allocation, which involves a comprehensive plan for citizens at risk. This strategy could address the discrimination and strengthen the government's response. As previously mentioned, changing attitudes might be impossible.

### **Breakout of Civil War (Shock Event)**

Recently, due to pressure from the French President Macron, the Lebanese parliament has elected German Ambassador Mustapha Adib as Prime Minister (Middle East Eye 2). Many opposition groups rejected the nomination based on the removal of the elsite, and others have compared Adib's promises of a technocratic government to his predecessor and fear he won't deliver (Middle East Eye 5). Lebanese activists have criticized Adib's previous Ambassador appointment due to the sectarian system and believe him to be corrupt (Middle East Eye 7). However, Adib has yet to announce his selections for cabinet members, and these choices could determine the public's reaction (Middle East Eye 22). Given the anger and discontent felt by many, if the new prime minister's choices cause controversy, it could spark a more direct and violent conflict. While the threat of civil war is relatively low, the new leader's subsequent decisions must be taken into account as the IMF designs economic policies and strategies to benefit Lebanon and handle negotiations.

Therefore, Lebanon's social problems are associated with the inherent political agendas and structures, and need to be targeted as part of a domestic investment plan to alleviate Covid-19 infections and create stale public services.

## **Geopolitical Risk: Moderate to High**

### **Overview**

The fragmented Lebanese political system has influenced foreign policy by allowing a small Shi'ite guerilla group known as Hezbollah to achieve an unprecedented level of political and military force in Lebanon, mainly dominating its foreign policy in the region (CEP 1). There is evidence of tremendous influence in the Lebanese government, as the Hezbollah-led parliamentary coalition has held 57 of the 128 seats in parliament following the 2009 elections, and helped elect Michel Aoun to the presidency (CEP 2). These relationships are at the forefront of Lebanon's international troubles. The group uses Lebanon to further its agenda, which often bodes terribly for Lebanon and its future. While Lebanon's position as a banking powerhouse is influential, since they are attracting investment as part of economic recovery, their political relations are paramount.

### **Hezbollah and Israeli Tensions**

Hezbollah's direct activities in Syria and Israel threaten to involve Lebanon and dictate its foreign policy through conflict and political pressure. An example is the 34-day war in 2006 between Israel and Hezbollah, which destroyed 130,000 homes

in Lebanon, and killed more than 1,100 Lebanese (CEP 17). UN Humanitarian relief coordinator Jan Egeland cites Hezbollah blending in with civilians as a direct cause of massive casualties (CEP 18), proving Hezbollah's willingness to involve Lebanon. Another major example is with Hezbollah leaders threatening to blow up Israeli nuclear reactors, and the Israeli government looking to hold the Lebanese government accountable and threatening to strike Lebanese targets in the event of a conflict (Iddon 5). Since Hezbollah exerts massive influence over Lebanese government, Lebanon will likely suffer in the next conflict. The IMF should be wary of Lebanon's position, as the Hezbollah's unwillingness to budge could lead to its becoming the target of international sanctions and scrutiny, which would be detrimental to Lebanon's economic growth. The Hezbollah's political influence makes it difficult to alter, and could hinder the effectiveness of IMF aid.

### **Association with Iran**

Additionally, Hezbollah members continue to pledge allegiance to Iran and receive funding and training (CEP 22). This stance continues to strain Lebanon's relationships with the rest of the Arab world and other countries that disapprove of Hezbollah. Notable examples of this include a canceled state visit by Saudi King Salman after President Aoun made pro-Hezbollah statements, and condemnation of Lebanon by the Bahraini Foreign Minister after Hezbollah criticized a Bahraini counter-terrorism raid (CEP 27). A recent New York Times report by Farnaz Fassihi looks at the effects of the Israel-UAE Deal, wherein Israel is recognized by the UAE in return for suspending its territorial gains (Fassihi 9). Iran's foreign ministry views this as the solidification of a regional partnership that could threaten them, and has increased support for its allies, such as Hezbollah, which has also denounced this deal (Fassihi 15). This event could further push Lebanon into a regional partnership with Iran, and increase its dependence on it, which is terrible for Lebanon's economic self-sufficiency, and could isolate them from vital players in the middle east, such as Saudi Arabia. The IMF should consider this when looking at Lebanon's international growth prospects and understand that this is a developing situation.

### **International Conflict over Aid**

Another source of potential conflict is with the international aid coming from countries such as France and Canada, which have tied aid to deep structural reform. France's foreign minister has called for political and economic reform, such as the formation of a technocracy, as grounds for international assistance to provide temporary economic and social relief (News Wires 4). However, given the immense influence exerted by Hezbollah, fundamental reform is unlikely. Because of such factors opposing Lebanese reform, the US will likely refuse to aid Lebanon, and indirectly help Hezbollah. Moreover, because of stricter sanctions on Syria due to the Caesar Syria Civilian Protection Act and Hezbollah's unwillingness to cut ties, the US will further oppose an IMF bailout (Rickards 30). Other countries like France, Iran, and Russia, all facing economic woes and Covid-19 related issues, may not provide significant support (Rickards 32). Saudi Arabia has also directly refused to provide aid that may benefit Hezbollah (Rickards 33). The apparent lack of significant bilateral support could impact the IMF's economic recovery plans, as it could require a larger initial loan. In response, the IMF will have to take a greater

risk with Lebanon, and put increased pressure on the political system to act, which might be unlikely.

### **Future Conflict with Israel (Shock Event)**

Recently, tensions between Hezbollah and Israel have flared up following the unveiling of Hezbollah's new fleet of precision-guided missiles. It boasted that it could attack cities with its 100,000 missiles that it had previously been unable to do so in 2006 (Iddon 5). The construction of these missiles occurred primarily in "Lebanese and Syrian sites" (Iddon 6), and in the event of a full scale conflict, Israel would place massive responsibility in the hands of the Lebanese government and considerable casualties will occur. However, given the recent Covid-19 outbreak and economic crisis in Lebanon, such a war is unlikely at this time. While this outcome is unlikely, a war could destroy Lebanon's already fragile economic and social systems. If the Israeli government involves its allies and targets Lebanon, it could ruin any IMF plans and lead to massive losses.

Thus, it is clear that the Hezbollah's unwillingness and the influence of Iran directly affect the Lebanese political system, thereby impacting its international growth. Additionally, the IMF faces greater risk by putting more pressure on the political system to enact domestic reform, which might backfire. This area might be the most difficult for the IMF since it is unpredictable, and the IMF has no say.

### **Recommendation**

Lebanon is in dire need of reform, and delayed negotiations will contribute to its deterioration. The problem is inherently political. All other risks stem from it, as the corrupt and negligent political elite made poor fiscal decisions and inadequately responded to social issues, bringing Lebanon to its knees. IMF assistance is desperately needed, but successful reform and recovery cannot occur without reforming the sectarian political system and changing attitudes. Since that might be out of the IMF's domain, the best possible alternative is to encourage economic growth by targeting the effects of negligence and corruption, such as strengthening public services and attracting international aid. Doing so could create greater government accountability and effectiveness in its action plans, but the entrenched political system could remain. The result of the current negotiations remains to be seen and could dictate the terms of recovery and Lebanon's future success.

### **Works Cited**

- Atli, Altay, et al. "It Is Time to Save Lebanon." *Middle East Institute*, 2 Sept. 2020, [www.mei.edu/publications/it-time-save-lebanon](http://www.mei.edu/publications/it-time-save-lebanon).
- "Beirut Explosion: What We Know so Far." *BBC News*, BBC, 11 Aug. 2020, [www.bbc.com/news/world-middle-east-53668493](http://www.bbc.com/news/world-middle-east-53668493).
- Collard, Rebecca. "How Sectarianism Helped Destroy Lebanon's Economy." *Foreign Policy*, 13 Dec. 2019, [foreignpolicy.com/2019/12/13/sectarianism-helped-destroy-lebanon-economy/](http://foreignpolicy.com/2019/12/13/sectarianism-helped-destroy-lebanon-economy/).



- Deutsche Welle. "Syrian Refugees in Lebanon More Scared of Starvation than COVID-19: DW: 06.05.2020." *DW.COM*, 2020, [www.dw.com/en/syrian-refugees-in-lebanon-more-scared-of-starvation-than-covid-19/a-53355378](http://www.dw.com/en/syrian-refugees-in-lebanon-more-scared-of-starvation-than-covid-19/a-53355378).
- Fassihi, Farnaz, and David. "The Israel-U.A.E. Deal and the Beirut Blast Both Box in Iran." *The New York Times*, The New York Times, 16 Aug. 2020, [www.nytimes.com/2020/08/16/world/middleeast/the-israel-uae-deal-and-the-beirut-blast-both-box-in-iran.html](http://www.nytimes.com/2020/08/16/world/middleeast/the-israel-uae-deal-and-the-beirut-blast-both-box-in-iran.html).
- Fordham, Alice. "Syrian Refugees In Lebanon Fear Deportation For Seeking Coronavirus Test Or Care." *NPR*, 6 Apr. 2020, [www.npr.org/2020/04/06/825158835/syrian-refugees-fear-deportation-if-they-seek-coronavirus-testing-and-treatment](http://www.npr.org/2020/04/06/825158835/syrian-refugees-fear-deportation-if-they-seek-coronavirus-testing-and-treatment).
- Iddon, Paul. "Tensions Between Israel And Hezbollah Are High, But Neither Side Wants War Now." *Forbes*, Forbes Magazine, 28 July 2020, [www.forbes.com/sites/pauliddon/2020/07/28/tensions-between-israel-and-hezbollah-are-high-but-neither-side-wants-war-now/](http://www.forbes.com/sites/pauliddon/2020/07/28/tensions-between-israel-and-hezbollah-are-high-but-neither-side-wants-war-now/).
- Iddon, Paul. "In A Third Lebanon War, Israel Will Have To Contend With Increasingly Lethal Hezbollah Missiles." *Forbes*, Forbes Magazine, 27 June 2020, [www.forbes.com/sites/pauliddon/2020/06/27/in-a-third-lebanon-war-israel-will-have-to-contend-with-increasingly-lethal-hezbollah-missiles/](http://www.forbes.com/sites/pauliddon/2020/06/27/in-a-third-lebanon-war-israel-will-have-to-contend-with-increasingly-lethal-hezbollah-missiles/).
- Ivanescu, Isabel. "Lebanon: Saved Then Ruined by Sectarian Power-Sharing (February)." *The Kenan Institute for Ethics at Duke University*, 12 Mar. 2020, [kenan.ethics.duke.edu/lebanon-saved-then-ruined-by-sectarian-power-sharing-february/](http://kenan.ethics.duke.edu/lebanon-saved-then-ruined-by-sectarian-power-sharing-february/).
- Khalaf, Samir G., and Paul Kingston. "Economy." *Encyclopedia Britannica*, Encyclopedia Britannica, Inc., 6 Aug. 2020, [www.britannica.com/place/Lebanon/Economy](http://www.britannica.com/place/Lebanon/Economy).
- Kranz, Michal. "Syrian Refugees in Lebanon Suffer amid Economic Uncertainty." *News Al Jazeera*, 23 Feb. 2020, [www.aljazeera.com/ajimpact/syrian-refugees-lebanon-suffer-economic-uncertainty-200222120627419.html](http://www.aljazeera.com/ajimpact/syrian-refugees-lebanon-suffer-economic-uncertainty-200222120627419.html).
- "Lebanon Launches IMF Talks to Rescue Economy; More Talks in Coming Days." *Reuters*, Thomson Reuters, 13 May 2020, [www.reuters.com/article/us-lebanon-crisis-imf/lebanon-launches-imf-talks-to-rescue-economy-more-talks-in-coming-days-idUSKBN22P339](http://www.reuters.com/article/us-lebanon-crisis-imf/lebanon-launches-imf-talks-to-rescue-economy-more-talks-in-coming-days-idUSKBN22P339).
- "Lebanon." *Freedom House*, 2020, [freedomhouse.org/country/lebanon/freedom-world/2020](http://freedomhouse.org/country/lebanon/freedom-world/2020).
- "Lebanon ." *Data*, 2020, [data.worldbank.org/country/lebanon](http://data.worldbank.org/country/lebanon).
- "Lebanon's Economic Update - April 2020." *World Bank*, 16 Apr. 2020, [www.worldbank.org/en/country/lebanon/publication/economic-update-april-2020](http://www.worldbank.org/en/country/lebanon/publication/economic-update-april-2020).
- Majzoub, AyaMaj. "Lebanon's Protests Are Far from Over." *Human Rights Watch*, 7 Sept. 2020, [www.hrw.org/news/2020/04/27/lebanons-protests-are-far-over](http://www.hrw.org/news/2020/04/27/lebanons-protests-are-far-over).
- Nakhoul, Samia. "IMF Deal Seen as Only Way out for Lebanon." *Reuters*, Thomson Reuters, 4 Mar. 2020, [www.reuters.com/article/us-lebanon-crisis-imf-analysis/imf-deal-seen-as-only-way-out-for-lebanon-idUSKBN20R0SL](http://www.reuters.com/article/us-lebanon-crisis-imf-analysis/imf-deal-seen-as-only-way-out-for-lebanon-idUSKBN20R0SL).

- Neumann, Julia. "Interview with Lebanese Political Scientist Bassel Salloukh : Will a New Lebanon Arise from the Ashes of Beirut?" *Qantara.de*, 11 Aug. 2020, en.qantara.de/content/interview-with-lebanese-political-scientist-bassel-salloukh-will-a-new-lebanon-arise-from.
- Page, Michael, and Aya Majzoub. "Lebanon Protests." *Human Rights Watch*, 2 Sept. 2020, www.hrw.org/blog-feed/lebanon-protests.
- Perry, Tom. "Lebanon Seeks Foreign Government-Backed Financing for Power Plants after Default." *Reuters*, Thomson Reuters, 20 May 2020, fr.reuters.com/article/us-lebanon-crisis-electricity-idUSKBN22W2TS.
- Press Release - September 1, 2020, et al. "Hezbollah's Influence in Lebanon." *Counter Extremism Project*, 31 Oct. 2018, www.counterextremism.com/hezbollah-in-lebanon.
- Qiblawi, Tamara. "75% Of Lebanon Needs Aid after Coronavirus, and Hungry Protesters Are Back on the Streets." *CNN*, Cable News Network, 29 Apr. 2020, www.cnn.com/2020/04/28/middleeast/lebanon-hunger-aid-coronavirus-intl/index.html.
- "Results - 2019 - CPI." *Transparency.org*, 2020, www.transparency.org/en/cpi/2019/results/table.
- Rickards, James. "Crisis in Lebanon." *FDD*, 5 Aug. 2020, www.fdd.org/analysis/2020/08/04/crisis-in-lebanon/.
- Sester, Brad W. "Lebanon's Imminent Financial Crisis." *Council on Foreign Relations*, Council on Foreign Relations, 18 Feb. 2020, www.cfr.org/blog/lebanons-imminent-financial-crisis.
- "Struggling to Survive: Slavery and Exploitation of Syrian Refugees in Lebanon." *The Freedom Fund*, The Freedom Fund, 5 Apr. 2016, freedomfund.org/wp-content/uploads/Lebanon-report-FINAL-5April2016.pdf.
- "Targeting Poor Households in Lebanon." *World Bank*, 21 Apr. 2020, www.worldbank.org/en/news/factsheet/2020/04/21/targeting-poor-households-in-lebanon.
- "The Unprecedented Mass Protests in Lebanon Explained." *Amnesty International*, 16 July 2020, www.amnesty.org/en/latest/news/2019/11/lebanon-protests-explained/.
- "Time for Iraq and Lebanon to Ditch State-Sponsored Sectarianism." *The Economist*, The Economist Newspaper, 7 Dec. 2019, www.economist.com/leaders/2019/12/07/time-for-iraq-and-lebanon-to-ditch-state-sponsored-sectarianism.
- "What People Think: Corruption in the Middle East & North Africa - News." *Transparency.org*, 11 Dec. 2019, www.transparency.org/en/news/what-people-think-corruption-in-the-middle-east-north-africa.
- Whewell, Tim. "Lebanon Electricity Crisis: Stealing Power to Survive." *BBC News*, 13 Dec. 2019, www.bbc.com/news/world-middle-east-50760043.
- "Who Is Mustapha Adib, Lebanon's New Prime Minister?" *Middle East Eye*, 31 Aug. 2020, www.middleeasteye.net/news/lebanon-prime-minister-who-is-mustapha-adib.

Wires, News. "Lebanon at 'Risk of Disappearing', French Minister Says, Pushing Govt to Adopt Reforms." *France 24*, France 24, 27 Aug. 2020, [www.france24.com/en/20200827-lebanon-at-risk-of-disappearing-french-minister-says-pushing-govt-to-adopt-reforms](http://www.france24.com/en/20200827-lebanon-at-risk-of-disappearing-french-minister-says-pushing-govt-to-adopt-reforms).



# Diversity vs. Inclusion in STEM: Factors that Accurately Measure a Tech Company's Commitment to Inclusion

---

*Adrienne Lai*

**Author Background:** *Adrienne Lai grew up in the United States and currently attends St. Stephen's and St. Agnes School in Alexandria, Virginia in the United States. Her Pioneer research concentration was in the fields of science, technology, and society and computer science and titled "Digital Mapping."*

## Abstract

STEM has traditionally been dominated by white and Asian men, and while the industry has grown tremendously, demographic diversity has not changed. It is crucial for tech companies to improve their diversity, as underrepresented minorities and women provide unique perspectives that enhance any workforce. When addressing diversity, tech companies publish a variety of information (diversity reports, workplace awards, etc.) on their websites. The public typically demands that corporations release their demographic data: breakdowns of company jobs based on race, ethnicity, sexual orientation, or gender. However, representation is not synonymous with inclusion. Diversity is the representation of people with different backgrounds, and inclusion is the incorporation of people, so having a diverse workforce does not necessarily mean that everyone feels welcome. This leads to the question: how can a tech company's inclusivity be accurately measured? An examination of secondary literature and data analysis found that representation data does not correspond with a tech company's commitment to inclusion. Reporting demographic data, workplace awards, leadership's commitment to inclusion, mentoring programs, and Employee Resource Groups most accurately reflect a company's inclusivity. In contrast, no measure examined correlated with employee diversity.

## 1. Introduction

Company growth does not necessarily relate to company improvement. While the tech industry expands, STEM workforces are still predominantly white and Asian men. The lack of representation of underrepresented minorities and women hurts tech companies because they are capable employees that provide unique perspectives on problems. In addition, as African American and Hispanic populations are the fastest growing racial groups in the United States, they will represent a significant portion of workers in the future (Alfred et al. 116). Thus, for the tech industry to continue expanding, it is important that these groups are included in STEM. To please the public's demands for companies to improve their minority representation, companies have different plans for inclusion, commitments to diversity, and ways to share their policies with the public. A company may release a variety of the following items to show that it values diversity: demographic employee data, a diversity report, awards, a "Diversity & Inclusion" webpage, a diversity statement and/or plan, an executive position or committee devoted to workplace inclusion, Employee Resource Groups (ERGs) for workers to get support, and partnerships with universities and/or organizations to encourage women and minorities to get involved in STEM and join their company. While each of these metrics gives different information about a company—e.g., the gender, racial, or ethnic diversity of its employees, support available for workers, or its perceived reputation—each one touts a company's diversity.

It is becoming increasingly clear, however, that diversity is not synonymous with inclusion. While diversity is the representation of people with different backgrounds, inclusion is the incorporation of people. Inclusion is key to welcoming minorities and women to STEM as encouraging, supportive environments allow them to fully realize that they have a place in an industry where there are very few colleagues similar to them. Having a diverse workforce is a great start to strengthening a tech company's workforce, but if the atmosphere is hostile, or if minority and female workers are isolated, they will be more likely to switch careers (Shevinsky 2015). Therefore, commitment to inclusivity creates workplace cultures that celebrate differences, which provides the foundation for company diversity to blossom.

A close look at the demographic data published by tech companies shows that while they may be good measures of diversity, they do not adequately capture a company's commitment to inclusion. As employee data is the metric demanded by the public, this leads to the question: how can a tech company's inclusivity be accurately measured? To answer this question, I will first review the secondary scholarship on the exclusive and male-dominated culture of STEM, the role of location on company growth and diversity, the benefits of workplace diversity, the role of education in the representation of minorities in STEM, and the effectiveness of university and company inclusion methods. Next, I will describe a new data set I prepared containing EEO-1 reports, diversity statements, and company inclusion plans found on company websites for 90 tech companies in Silicon Valley, Seattle, and Boston. A statistical analysis of this data suggests that reporting demographic data as a percentage and/or EEO-1, reporting awards, the number of awards, leadership's commitment to inclusion, mentoring programs, reporting ERGs, and the number of ERGs accurately measure inclusion.

Furthermore, none of the metrics studied accurately measure employee diversity, and a comparison of the representation of Hispanic, Latino, and African American women in tech companies to those of the regional STEM private industry and the national private industry indicates that companies that publish EEO-1s do not necessarily have model employee diversity.

## 2. Literature Review

Scholars have studied diversity and inclusion in the tech industry from a number of angles. First, the importance of diversity in global American tech competitiveness will be discussed. It will be followed by an examination of the male-dominated and exclusive nature of STEM, the role of location on company growth and diversity, and the benefits of workplace diversity. Then, the role of education in underrepresented minorities and female students STEM involvement will be highlighted. Finally, the effectiveness of university and company inclusion measures will be discussed.

Mary V. Alfred, Sarah M. Ray, and Michele A. Johnson argue that increasing diversity in the tech industry is necessary for it to remain competitive in the global tech economy (116). The tech industry is a rapidly growing sector that revolutionizes how people communicate, interact, and live their lives. The number of engineering and computer science jobs increases at twice the rate of the national average, and these jobs tend to have higher pay, making STEM a beneficial occupation (U.S. Equal 1). Also, the industry has self-sufficient growth, as new technologies lead to spinoff inventions that expand the sector. However, while tech companies are at the forefront of innovation, the industry lacks diversity. Compared to other private industries, the tech sector has greater proportions of white and Asian men and smaller proportions of African Americans, Hispanics, Latinos, and women (U.S. Equal 2). Furthermore, there is less minority representation at higher level positions in companies, which indicates a stigma against promoting them; company leaders are promoted from lower level employees, so the leadership team should reflect the diversity of the company (U.S. Equal 20). Alfred and her colleagues argue that this is a flaw in STEM because “the fastest growing population groups are racial minorities, thus making long-term economic growth of the United States dependent on its racial minority population” (Alfred et al. 116). In fact, it is predicted that by 2050, racial minorities will be around half of the U.S. population, so it is key for them to be welcomed in tech so that it can grow (Palmer et al. 492).

While some companies promote the pipeline problem, the idea that the lack of diversity is due to minorities’ lack of interest in STEM, Sarah Thébaud and Maria Charles, Brittany Bloodhart et al., Katy Shevinsky, and Maria Ong et al. claim that the main reason for the representation gap is due to the exclusive, competitive, male-dominated, and individualist tech culture (Shevinsky 203). While there are more female college graduates than male graduates, men make up a majority of the STEM majors (Bloodhart et al. 1). Even if women graduate with a STEM degree, some do not pursue a STEM career, and most will stop working in STEM due to feelings of isolation, lack of advancement, and a negative work culture (U.S. Equal 7, 9). Additionally, while STEM’s reputation is changing to include women and minorities, the atmosphere remains the same. Ong et al. argue that many universities’ and companies’ tutoring and self-confidence

building inclusion plans promote the idea that these groups need extra support or to be “fixed,” which does not change the tech industry’s exclusive reputation (208). Also, Katy Levison reports that while employees report discrimination, not much is done to correct it, and some companies discourage reporting (51). Based on her experience in Silicon Valley, Levison writes that she is hesitant to speak up about company problems because she does not want to be labeled as a liability like other outspoken women (51).

Furthermore, Ong et al. and Thébaud and Charles argue that since STEM values characteristics like competitiveness, individuality, and raw talent that are traditionally associated with men, women feel isolated in STEM (207; 4). Thébaud and Charles cite the “geek” and “programmer” images as promotions of the idea that social skills traditionally associated with women (caring, emotional, etc.) are not valued in STEM, which causes women to be less interested in it (4). According to Bloodhart et al., the ingrained gender bias creates the stereotype that women are less capable than men in STEM fields; when women outperform or outnumber men in STEM college courses, both male and female students underestimate the performance of women (7). This false perception translates to the workplace as employers will pay more for white male workers they view as more productive, and faculties are less likely to hire women and male students with non-white sounding names as they are viewed as incompetent (Thébaud and Charles 3-4). Thus, the white male dominance in the tech industry is reflected in the discrimination of underrepresented minorities and women in STEM.

Gijsbert Stoet and David C. Geary, Enrico Moretti, Dante Di Gregorio and Scott Shane, Massimo Colombo et al., and Bryan Stephens et al. examine the role of location in tech companies’ diversity and growth. While one might think a community with more gender equality would negate the negative perceptions that discourage female involvement in STEM, Gijsbert Stoet and David C. Geary argue that the opposite is true: countries with greater gender equality have larger gaps in STEM female representation (581). They attribute this phenomenon to countries with greater gender equality tending to be welfare states where the monetary benefits associated with tech jobs are less appealing, so individuals are more likely to pursue careers in subjects that interest them (Stoet and Geary 582). When individuals follow their passion, they argue that women tend to pursue non-STEM careers, even if their science skills are similar to those of men pursuing STEM, because as reading tends to be their best skill, women think they will enjoy a career utilizing it more than a STEM career (Stoet and Geary 582). Thus, the study provides evidence for the pipeline problem as it argues that women are naturally less STEM inclined than men because when selecting which career interests them, more women choose non-STEM.

While diverse workplaces foster innovation, gender balance is not a factor tech companies consider when placing their headquarters. Historically, Moretti states that companies are located near natural resources and their consumers, so they have easy access to materials and their clientele (121). Moretti argues that now, tech companies are attracted to areas with thick labor markets, specialized services (venture capital and part manufacturing), and knowledge spillovers. He claims that the market size is important to attract workers with specialized skills (engineers and scientists) because more job opportunities lead to more employee talent in the area (128). With a large talent pool available, more companies want to be located there as well, so special services that support startups like

consulting, advertising, and manufacturing move there to gain clients (Moretti 133). Furthermore, Moretti argues that a key service for the tech industry is venture capital, since venture capitalists provide crucial support to startups in funding, mentoring, and organization. Di Gregorio and Shane argue that it is important for them to be located near their client company to provide the best aid, so venture capitalists tend to invest in companies in close proximity, which attracts companies to the area (211).

While one might expect tech company clusters to create a competitive environment, it spurs innovation. According to Stephens et al., the high level of inventive thinking fosters knowledge spillovers, as collaboration is natural and inventions lead to spinoffs that lead to new companies (269). Similarly, Colombo et al. claim that proximity to research universities increases knowledge spillovers, as research is passed from faculty to professionals through seminars, internships, conferences, and graduates (144). However, they argue that if a university's research is commercially funded, there are fewer knowledge spillovers, as there are more intellectual property rights (Colombo et al. 188). Additionally, Di Gregorio and Shane's study shows that highly acclaimed universities generate more startups as the institution's reputation helps gain funding (212).

While the rich talent pool and proximity to venture capitalists, research universities, and other tech companies allows tech hubs to prosper, the growth of tech hubs can be increased by their location being diverse. Stephens et al., Samila and Sorenson, and Cheruvelil et al. examine the benefits of a diverse workforce in innovation and teamwork. Stephens et al. argue that immigrants play a big role in tech as they bring knowledge from around the world and provide unique perspectives (271). They point out that largest U.S. tech hubs benefit from immigration: Silicon Valley has a large, diverse population due to strong immigration, and cultural diversity is highlighted in New York City (Stephens et al. 269, 271). Also, Samila and Sorenson argue that racially and ethnically diverse communities benefit creativity as people of different backgrounds are constantly interacting, leading to more innovative designs (772, 774). Cities with smaller subareas are less segregated because people must branch out into more areas, which fosters inclusion (Samila and Sorenson 782). Thus, Samila and Sorenson argue that diverse companies have faster rates of economic growth and foster more innovative thinking (790).

In addition, Cheruvelil et al. argue that diversity is important in STEM because it creates efficient teams. Productive and collaborative teams are valuable because the finished design is better than what each individual would have created alone (Cheruvelil et al. 31). According to Wilton et al., diversity improves the workplace by introducing people to new perspectives and encouraging them to share ideas (60). One study found that "teams randomly assigned an ethnically diverse set of members learned more from each other, and therefore correctly solved the problems presented to them at higher rates" (Samila and Sorenson describing Philips, Northcraft, and Neale (2006) 775). Thus, having multiple perspectives in a team fosters problem solving and creative thinking.

Palmer et al., Alfred et al., and Thébaud and Charles discuss the role of education in encouraging underrepresented minorities and women to pursue STEM careers. Education is crucial for future employee demographics in STEM because school is where people determine what they are interested in. Palmer et al. state that high school influences what degrees students pursue and whether



minority students continue with STEM in college because it is when students first consider possible career paths; teachers introduce students to their future or help eliminate potential professions (458). Also, as Alfred et al. claim, biased disciplinary actions against students of color negatively affect their education as they distrust their teachers and become less invested in their education (122). Some might think that to get students more engaged in school, they should select their classes, so they can explore their interests. However, Thébaud and Charles's study shows that schools with more STEM requirements have a greater gender balance among students who want to pursue STEM in college (4). Increasing high school STEM requirements might be necessary for tech to continue to grow as the ratio of students who receive STEM undergraduate and graduate degrees has decreased in the U.S. (Palmer et al. 491).

Similar to high school, a person's college experience can have profound effects on their future; people figure out what career interests them, gain professional skills, build strong faculty relationships, and expand their social network. This is not different for women or minority students studying STEM; their undergraduate experience can give them the support to continue in STEM or decide that another path is best for them. In a study by Robert T. Palmer, Dina C. Maramba, and T. Elon Dancy II, an undergraduate student reported that at the "end of his first year he observed that more than one-half [of students of color] dropped out or changed to non-STEM majors" (496). Thus, graduate school is especially critical for female and minority students, as the gaps in representation are even more pronounced (Alfred et al. 124). Universities have different ways to try to improve the retention rates among female and minority students in STEM. For example, prioritizing diversity and offering financial aid help increase diversity in STEM at an institution (Ong et al. 228; Palmer et al. 493).

The lack of belonging in STEM is increased for first-generation college students. As students are less engaged and confident in intimidating environments, isolation is a huge contributor to students dropping out as it is associated with decreased performance in classes (Palmer et al. 501; Duran et al. 136). For first-generation students of color in STEM fields, this isolation is further pronounced, so they are more vulnerable to changing career paths. As the majority of first-generation students are underrepresented minorities, the increased lack of belonging contributes to their lack of representation in STEM. Research done by Antonio Duran and his colleagues shows that campus housing, peer networks, learning centers, positive faculty relationships, and involvement in campus co-curriculars increase feelings of belonging at college (Duran et al. 136).

Similar to the resources to help first generation students feel comfortable, universities provide resources to support female students and students of color as they pursue STEM degrees; Duran et al., Maria Ong et al., and Palmer et al. discuss the effectiveness of these methods. Peer support groups connect peers with similar experiences and goals, so everyone is comfortable with each other (Palmer et al. 496). Peer support does not have to be with people of the same race or gender; the most important element is that the group members share an understanding of each other (Ong et al. 223). In the study conducted by Maria Ong and her colleagues, a female Latino student said that, as an undergraduate, she reached out to a female Latino postdoctoral candidate in her department and when they met, she felt instantly comfortable to share her struggle, even though

she did not know her (223). This shows the power of a relationship with someone with similar experiences and interests, so peer support groups are beneficial to supporting female students and students of color in STEM.

According to Ong et al., student racial, gender, and sexual orientation diversity conferences are great support resources as well because students build connections with students with similar backgrounds and interests and meet role models that show them that they can be successful in STEM (223). While Palmer et al. argue that involvement in non-STEM activities negatively impacts the retention of students of color in STEM (as they are not fully engaged in STEM disciplines), Ong et al. argue that they are beneficial because students gain friends who will support them (493; 224).

Like peer support, faculty support is key during a student's college experience. Underrepresented minority students greatly benefit from having a mentor who strongly believes in them and/or with a similar background as them. According to Palmer et al., a student's grade might not change due to faculty guidance, but their confidence will strengthen (492). In Maria Ong, Janet M. Smith, and Lily T. Ko's study, an African American woman spoke about the impact a visiting professor (who was also African American) made when they insisted that she take their class even though she thought she would fail as her white faculty advisor told her that it was too hard for her (224). She excelled in the course and gained more confidence in herself due to a professor believing in her, which shows the power of faculty support (Ong et al. 224). The effects of discouraging mentors are also shown as she did not believe she could pass due to her advisor's discouragement. Thus, since faculty support is crucial in building the confidence of female students and students of color, it is critical for them to have supportive faculty relationships.

For this reason, Palmer et al. argue that research opportunities are important for improving retention rates because in research positions, students connect with faculty (493). They gain mentors to support them through their college career, which makes them feel more included. Furthermore, they argue that research allows students to enhance their knowledge and understanding of STEM and gain new skills, which improves their confidence and helps them determine their career goals (Palmer et al. 493). Realizing that they are capable of conducting high level research helps students believe they have a future in STEM, so they are more likely to continue in the industry.

Similar to colleges, companies with hostile workplace cultures have issues retaining underrepresented minorities; Shevinsky writes, "Women who work in tech agree that the problem is harassment, discrimination and a generally hostile work environment. Many cite difficulties being hired, promoted, or getting funding" (205). As diversity creates high performing collaborative teams, it is important for companies to foster an inclusive environment (Cheruvilil et al. 32). Companies have different methods for creating a more inclusive community, such as Employee Resource Groups (ERGs), where employees can connect with professionals with similar backgrounds as them and creating diversity and inclusion plans. Another method is diversity interventions. In "Addressing Unintended Consequences of Gender Diversity Interventions on Women's Sense of Belonging in STEM," the authors argue that while Video Interventions for Diversity in STEM (VIDS) increase workplace awareness of gender bias and recognition of sexism, they inadvertently cause women to feel isolated, act as a

social identity threat, and increase female's negative feelings for STEM (Pietri et al. 528-529, 532). They argue that additional methods like articles about overcoming adversity and UNITE modules (in which participants learn to combat sexism and that biases can change) can mitigate these negative consequences (Pietri et al. 543). Thus, not all intervention methods result in progress towards fostering a more inclusive culture, similar to how not all diversity reporting metrics truly reflect the inclusiveness within a company.

### 3. Data Collection

While the secondary literature provides insight into many aspects of diversity and inclusion in the tech industry, it does not discuss how to measure commitment to inclusion: how does the public know if a company is actively working to foster diversity in the workplace? In order to determine the best inclusion metric, employee demographic data and company inclusion plans were compiled for 90 tech companies in Silicon Valley, Boston, and Seattle. The data comes from a variety of sources: EEO-1 reports, company diversity reports, and company webpages.

I used companies' Employer Information Reports (EEO-1s) to get the demographic data of employees. The EEO-1 is a document that private companies with more than one hundred employees file with the U.S. Equal Employment Opportunity Commission, a federal agency responsible for enforcing laws against workplace discrimination, each year ("EEO-1 Survey"; "Fact Sheet"). Companies break down their staff by job, gender, and race (one box is "White, Male, Professionals" for example), and employees are asked to self-identify their race, but if they do not answer, their employer chooses ("Fact Sheet"). While some tech companies publish their EEO-1 report, most do not, so a limitation is that only companies that publish the data will be included. The companies that chose to publish are proud of their diversity numbers or think they will boost their reputation, so the percentages from the data are probably above average. Furthermore, within the job categories, technical and non-technical jobs are not separated. For example, the "Professional" category includes computer programmers and lawyers. The most recent EEO-1 report published is from 2018, but some companies do not have this most recent report available—instead they have it from 2015-2017—so the dates of filing are not consistent among the companies. The date range can make it difficult to determine which company truly has the best diversity, as the data is from different periods in the companies' histories.

Some companies publish a diversity report that includes details of their efforts to increase diversity and foster an inclusive environment (plans, affinity groups, partnerships, progress, awards). The reports typically release demographic data as percentages and do not break down the employees by gender and race (there is no white women category for example). Thus, I used the report data as an indication of a company's commitment to inclusion instead of a supplement to the EEO-1 data. I also used the report and plans outlined to measure a company's commitment to inclusion. Diversity reports are created by a company, so they are positively biased. Each company's diversity report has a different format, so each one has different elements, which can make the commitment to inclusion metric unstandardized as some companies do not

publish certain pieces of their diversity efforts. For example, Apple received three workplace and diversity awards in 2020, but they are not listed in its report or on its website, so in this study, Apple is viewed as having no awards ("Apple Awards"). The reports were mostly released in 2018-2020, but some were published in 2016 and 2017. Similar to the EEO-1 reports, the different date ranges make it difficult to compare companies as the reports describe plans from different timepoints in company development.

A company's EEO-1 and diversity report are typically found on a company's "Diversity and Inclusion" webpage. If a company does not have a formal diversity report, the same information is typically found there. I used the webpage and its information (plans, awards, statements) as a measure of commitment to inclusion. Furthermore, I looked at the company "Leadership" webpage to see if there is a leadership position devoted to diversity and inclusion, which counts towards company diversity commitment. Similar to the diversity reports, each company's page includes different elements, so the commitment to inclusion element is reliant on what they choose to report, not necessarily what they are doing. As a company's website is created by the company, its goal is to make the company look good, so there is a positive bias.

I created the sample of 90 tech companies by selecting the tech companies that release their EEO-1 reports and adding them to a random sample of tech companies to bring the total number of companies to 30 in each region. I found the EEO-1 releasing companies by searching for which tech companies release their EEO-1s in each region. There are 27 Silicon Valley companies, 2 Seattle companies, and 2 Boston companies included that have released EEO-1 reports. To create the random sample for Silicon Valley, I transcribed the tech companies on the Employbl website onto a spreadsheet (1100+ companies), imported it to Wolfram Mathematica, and then randomly selected tech companies until the total number of Silicon Valley companies was 30. To create the random sample for Boston and Seattle, I transcribed the tech companies on the first twenty pages of the location specific company list on the BuiltIn website onto a spreadsheet (~350 companies), imported it to Mathematica, and then randomly selected companies until each region had a total of 30 companies. I replaced companies if they did not have a website or were not headquartered in Silicon Valley, Seattle, or Boston.

At the start of data collection, I emailed companies in the three regions about potentially providing their EEO-1 report or a similar breakdown of their workforce demographic data for this project. However, I did not email all the companies in the sample as I realized this method of data collection was time-consuming and did not typically result in data. I felt that examining more metrics used to measure diversity and inclusion was a better use of time. However, I did get one response from a Boston company. I decided to use this data, even though not all the companies got a chance to respond to a request, because there is a lack of Boston companies that report their data (only 1 not including this company). They asked that their company name be kept private, so they are referred to as BostonA.

Also, while the most recent official EEO-1 report released is from the year 2018, NetApp only provided its 2020 EEO-1 report draft—the EEO-1 report they submitted using payroll data from 2019. It is important to keep in mind that these numbers were not approved by the U.S. Equal Employer Commission officially

when the data was analyzed. However, as it uses data from a period that has already occurred, we may assume that the data is accurate.

#### 4. Data Methodology

In order to evaluate a company's success at fostering inclusion, I used these sources to assemble a new composite dataset recording 16 diversity and inclusion metrics in a sample of 90 tech companies distributed evenly between Silicon Valley, Seattle, and Boston. These areas are significant for my study since Silicon Valley has a multitude of large tech companies and startups, Seattle has the headquarters of some large tech companies like Microsoft, and Boston has many top educational institutions including MIT.

In order to make sense of this varied data, I propose a new “commitment to inclusion” metric, which I calculate for each company by counting how many of the following elements it has on their website: EEO-1, diversity report, inclusion plan, diversity and inclusion statement, leadership position or committee dedicated towards diversity and inclusion, mentoring program, workplace awards, “Diversity & Inclusion” webpage, Employee Resource Groups (ERGs), demographic percentages, pay equity, and partnerships with universities or companies to get women and/or underrepresented minorities involved in STEM. The number of awards or ERGs was not factored in. I counted diversity statements if the company addressed their commitment to or the importance of fostering a diverse workplace. I counted company inclusion plans based on the details the company provided on how it creates an inclusive workplace. Companies received half points if they solely mentioned they had a diverse workplace for the statement, or if they alluded to a plan but did not provide specific details. In the mentoring category, companies could earn two points: mentoring employees was one point, and mentoring students was another point. The pay equity metric included whether the company mentioned it had pay equity or measures the pay gap annually.

**Table 1.** *Company Total Commitment to Inclusion Scores by Region.*

Silicon Valley Companies	Total Commitment Score	Boston Companies	Total Commitment Score	Seattle Companies	Total Commitment Score
Apple Inc	10.0	Boston Scientific Corporation	12.0	Amazon.com Inc	9.0
Cisco Systems Inc	14.0	BostonA	3.5	Microsoft Corporation	14.0
Alphabet Inc	11.0	Catalant Technology	1.0	Convoy	0.0
Adobe Systems Incorporated	11.0	SessionM	1.0	SAP Concur	3.0
Intel Corporation	8.0	PlacePass	0.5	Joe	0.0

Slack Technologies Inc	8.0	Vesper	0.0	Quadrant Resource	1.0
Yahoo Inc	3.0	Own Up	0.0	Rad Power Bikes	3.0
Palo Alto Networks Inc	9.0	Payfactors	0.0	PayScale	4.0
NVIDIA Corporation	10.0	PatientPing	0.0	AppSheet	0.0
View Inc	1.0	Humana Studio_h	1.5	Formidable	1.0
Twitter	6.0	Alyce	0.0	Evergreen Home Loans	1.0
PayPal Inc	9.0	Dimagi	0.5	Project Archer	0.0
MobileIron Inc	1.0	Honeycomb	0.0	Ookla	2.0
Square	8.0	InsightsSquared	0.5	Callison	7.0
LinkedIn Corporation	9.0	Trux	0.0	Chef Software	1.0
NetApp, Inc	10.0	BrandBot	0.0	Avalara	4.0
Intuit Inc	9.0	Cybereason	1.0	Hare-brained Schemes	2.0
Hewlett Packard Enterprise	9.0	Acquia	3.0	WinDocks	0.0
Lyft	11.0	Inphantry	0.0	JetClosing	0.0
VMware Inc	11.0	Zoovu	0.0	Twistle	0.5
Pandora Media Inc	8.0	Hakeema	0.0	Fuse Networks	0.0
Salesforce.com	12.0	ASICS Digital	0.0	AeroTEC	2.0
Mozilla Corporation	7.0	Robin	0.0	LeafTail Labs	1.0
Facebook Inc	12.0	Pepper Gang	0.5	Ekata	1.5
Pinterest	5.0	Flipside Crypto	0.5	Cartogram	0.0
eBay	11.0	Ten Percent Happier	0.0	Autho	2.5
Airbnb	8.0	Zome Energy Network	0.0	Apptentive	1.0
Wind River	2.0	Demiurge Studios	0.0	Neal Analytics	0.0
Sandisk	1.0	Akamai	4.0	Vcare Software Development	0.0
Helpshift	0.0	AdmitHub	0.0	RealSelf	2.0

I calculated a company's workplace diversity by determining what percentage of its "Technical" and "Professional" EEO-1 category workers are Hispanic, Latino, or African American women. I chose the "Technical" and "Professional" categories because they represent the most STEM related positions. I selected Hispanic, Latino, and African American women to indicate a company's diversity because these workers have the most discrimination against them, as they are part of two underrepresented groups. Thus, these workers'

representation can be used to indicate diversity in STEM positions in the entire tech company, as they face the greatest barriers to entry.

**Table 2:** *Percentage of STEM jobs occupied by Hispanic and African American women by tech company.*

Tech Company	Year	Percentage of STEM Jobs Occupied by Hispanic/Latino/African American Women (%)
Apple Inc	2018	3.01614
Cisco Systems Inc	2018	2.99853
Alphabet Inc	2018	2.24734
Adobe Systems Incorporated	2018	3.31823
Intel Corporation	2018	3.25097
Slack Technologies Inc	2018	5.02874
Yahoo Inc	2015	1.89565
Palo Alto Networks Inc	2018	2.66739
NVIDIA Corporation	2018	0.85488
View Inc	2016	0.641026
Twitter	2016	2.27713
PayPal Inc	2017	3.11948
MobileIron Inc	2016	2.50784
Square	2016	2.01699
LinkedIn Corporation	2016	2.99043
NetApp, Inc	2020	2.71809
Intuit Inc	2016	2.97974
Hewlett Packard Enterprise	2018	3.49593
Lyft	2017	3.79859
VMware Inc	2018	1.94659
Pandora Media Inc	2017	7.08447
Salesforce.com	2018	2.55459
Mozilla Corporation	2017	2.15311
Facebook Inc	2018	3.85237
Pinterest	2018	3.9267
eBay	2018	1.70261
AirBnb	2018	4.60941
BostonA	2020	3.7037
Boston Scientific Corporation	2018	3.34019
Microsoft Corporation	2018	2.95944
Amazon.com Inc	2016	1.87395

## 5. Data Analysis

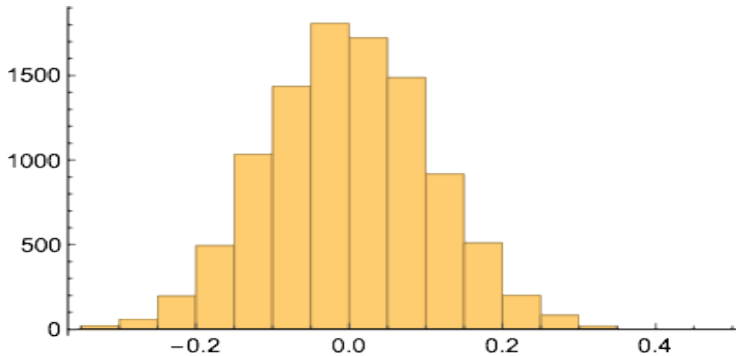
Using the representation data from the EEO-1 reports and the company commitment to inclusion information from websites, I examined how accurately different metrics reflect a company's commitment to inclusion and diversity. First, I compared employee diversity to the company's commitment to inclusivity. Then, I examined the relationship between reporting demographic data, commitment to inclusion, and company diversity. Next, I compared EEO-1 reporting companies' employee diversity to regional and national private industry employee diversity. Then, I compared reported awards and retention to diversity and inclusion. Finally, I calculated the correlation coefficients of leadership, mentoring, and ERGs to commitment to inclusion and diversity.

For each comparison, I calculated the correlation coefficient between the variables using the Wolfram Mathematica software. All 90 companies were included in each calculation unless otherwise specified. When comparing metrics to company commitment to inclusion, I calculated the correlation coefficients between the metric and the commitment score excluding the metric; for example, to measure the correlation between reporting awards to inclusivity, I calculated the correlation coefficient between reporting awards and the commitment score without award reporting factored in. When I calculated the correlation coefficient between the number of awards or ERGs and commitment to inclusion, I used the commitment score that did not include whether awards and ERGs were reported. Also, when metrics were compared to employee diversity, I only used companies that report EEO-1s because these are the only companies where the representation of Hispanic, Latino, and African American women in STEM jobs is known.

First, I calculated the correlation coefficient between the diversity of the company and the commitment to inclusion to be 0.1315. This low and weak correlation coefficient shows that there is a very small correlation between workforce diversity and commitment to inclusion, so employee diversity is not a good measure of company inclusivity.

While employee demographic data does not have a relationship with company inclusivity, companies report their demographic breakdowns or company diversity percentages to show that they are a welcoming corporation. A study by Leigh S. Wilton and her colleagues found that white men associated gender diversity reporting with the corporation being creative, fair, and prestigious, and believed a company was more open minded when it mentioned that it was “gender diverse” (Wilton et al. 59-60). To determine if reporting demographic data correlates with a company’s commitment score, I selected companies that have published diversity percentages and/or an EEO-1 report from 2016-2018 as a reporting data company because the recent years show a commitment to keep the public informed. The Boston company that provided EEO-1 data solely for this study is not counted as reporting an EEO-1 report, as they have not publicly shared the information. I found that the correlation coefficient between data reporting and commitment score is very strong and positive (0.856619), which supports the study’s findings. On the bootstrapping statistics method histogram (Figure 1), the correlation coefficient is positioned far on the right tail, so it is a significant result. The very strong correlation shows that data reporting is a good measure of a company’s commitment to inclusion. This phenomenon can be attributed to the positive bias of company data reporting: companies report to boost their reputation, so it makes sense it would only be published if the company’s numbers and commitment were good.



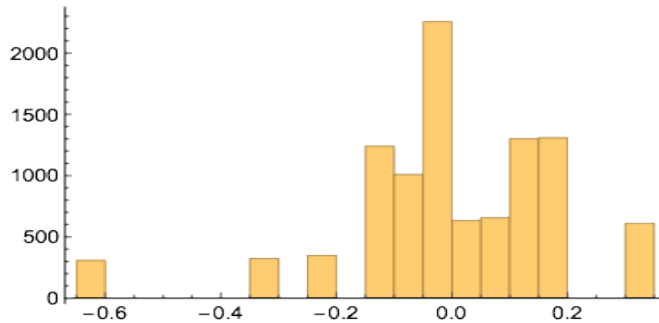


**Figure 1:** Histogram of Random Correlation Coefficients between Reporting and Commitment Score.

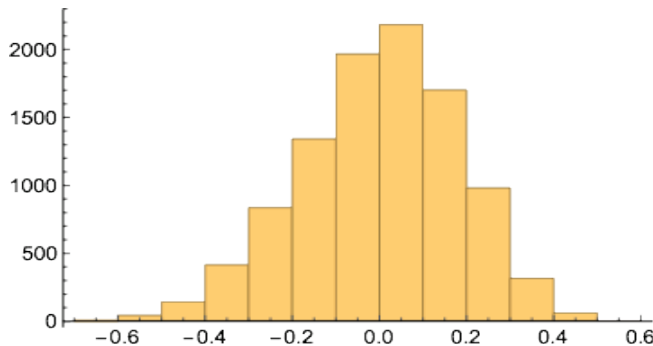
However, Wilton et al. also found that when companies released gender diversity information that included a description of race (the proportion of employees who are white women for example), the boost to the company's reputation in the opinion of white men was less than the boost without the mention of race (Wilton et al. 59). This could explain why few companies release their EEO-1 reports: the combination of gender and racial demographics would not boost their reputation by as much as it could without it. To test the validity of race demographics increasing positive company perception less than gender demographics, I calculated the correlation coefficients between solely EEO-1 reporting and commitment score, and between solely percentage reporting and commitment score. EEO-1 reporting represents gender and racial breakdowns in STEM jobs, while the percentage reporting represents the gender breakdown. The correlation coefficient between EEO-1 reporting and commitment score is 0.825827 and the correlation coefficient between percentage and commitment score is 0.862346. The correlation coefficient to commitment score with percentage was higher than the correlation coefficient with EEO-1 reporting, which suggests that the reputation boosts companies gain by solely reporting gender demographic breakdowns are greater than the reputation boosts companies gain by reporting gender and racial demographic breakdowns. Both correlations are significant by the bootstrapping method, so they are significant. As both the EEO-1 only reporting and percentage only reporting correlations are strong and positive, they are good measures of a company's inclusivity.

While employee demographic reporting corresponds with a higher commitment to inclusion, does it correlate with a demographically diverse workplace? To find out, I calculated the correlation coefficients between both types of data reporting (EEO-1 and percentages) and the representation of Hispanic, Latino, and African American women in the workforce (the percentage of STEM jobs they make up). The EEO-1 correlation coefficient is 0.159404 and the percentage correlation coefficient is 0.223937. Both correlation coefficients are not located far from the center on the bootstrapping histograms (Figures 2 and 3), so the correlations might have happened by chance. The correlation coefficients show that reporting percentage data corresponds with companies with diverse workplaces more than EEO-1 data. This suggests that companies reporting gender-diversity have more diverse workforces than companies that

issue EEO-1s. However, the results can be due to the fact that companies that reported percentages had to have reported EEO-1 data as well to be included, because companies with EEO-1 data are the only ones that can be used to examine the representation of Hispanic, Latino, and African American women in STEM. Thus, it is unclear whether the percentage reporting alone correlates with greater demographic diversity. Both correlation coefficients are small, which suggests that reporting either EEO-1 or percentage data does not correlate with the representation of underrepresented minority women in tech companies.



**Figure 2:** Histogram of Random Correlation Coefficients between EEO-1 Reporting and the Representation of Underrepresented Minority Women in STEM jobs. The calculated correlation coefficient is 0.159404.



**Figure 3:** Histogram of Random Correlation Coefficients between Percentage Reporting and the Representation of Underrepresented Minority Women in STEM jobs. The calculated correlation coefficient is 0.223937.

Percentage reporting seems to boost a company's reputation more than gender and race demographic data, which begs the question: why do companies publish the EEO-1? Few companies release their EEO-1 reports; I found that 27 Silicon Valley companies, 2 Seattle companies, and 1 Boston company published their EEO-1 report out of the approximately 9000 companies in these three regions combined ("US Tech"). They might report their data because their workforce has greater representation than what is normal in the tech industry. To test this hypothesis, I compared the percentage of all workers in tech companies who are Hispanic, Latino, or African American women to the percentage of STEM private industry workers who are Hispanic, Latino, or African American women

in California, Washington, and Massachusetts. I took the regional data from the U.S. Equal Employment Opportunity Commission's 2018 Job Pattern EEO-1 data.

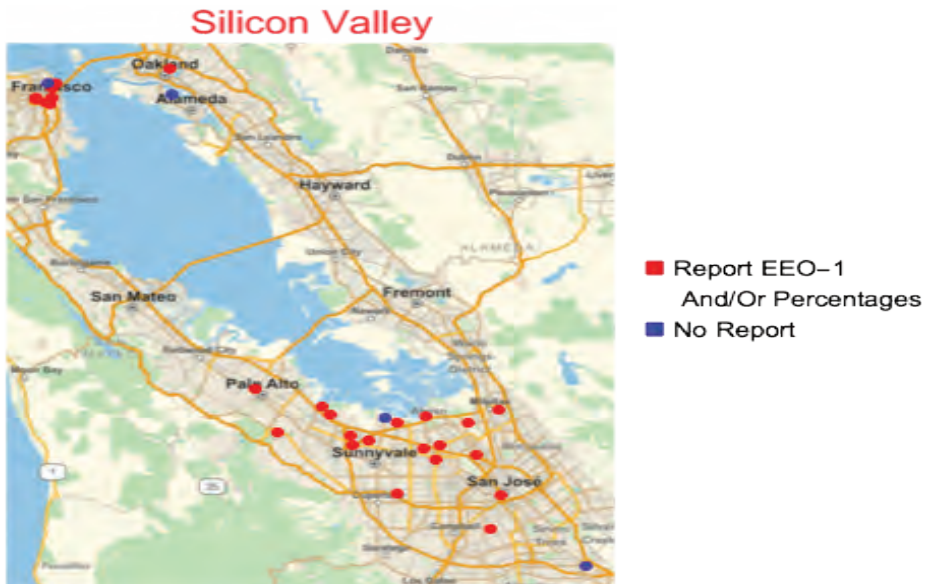
**Table 3:** *The Percentage of all Company Workers who are Hispanic, Latino, or African American Women.*

Tech Company	Year	Region	Percentage of All Workers Who Are Hispanic/Latino/African American Women (%)
Apple Inc	2018	California	1.28997
Cisco Systems Inc	2018	California	2.08283
Alphabet Inc	2018	California	1.56138
Adobe Systems Incorporated	2018	California	1.93396
Intel Corporation	2018	California	2.63788
Slack Technologies Inc	2018	California	3.26493
Yahoo Inc	2015	California	1.36579
Palo Alto Networks Inc	2018	California	1.34505
NVIDIA Corporation	2018	California	0.646445
View Inc	2016	California	0.217391
Twitter	2016	California	1.59214
PayPal Inc	2017	California	1.53519
MobileIron Inc	2016	California	1.58103
Square	2016	California	1.11046
LinkedIn Corporation	2016	California	1.87829
NetApp, Inc	2020	California	1.78299
Intuit Inc	2016	California	1.69176
Hewlett Packard Enterprise	2018	California	2.62933
Lyft	2017	California	1.83918
VMware Inc	2018	California	1.12611
Pandora Media Inc	2017	California	4.02477
Salesforce.com	2018	California	1.39966
Mozilla Corporation	2017	California	1.62162
Facebook Inc	2018	California	2.92366
Pinterest	2018	California	2.58324
eBay	2018	California	0.85324
AirBnb	2018	California	3.24675
BostonA	2020	Massachusetts	2.5
Boston Scientific Corporation	2018	Massachusetts	1.37271
Microsoft Corporation	2018	Washington	2.24859
Amazon.com Inc	2016	Washington	0.34566

**Table 4:** *The Percentage of all Private STEM Industry workers (Private Industry workers for the US) who are Hispanic, Latino, or African American Women in 2018.*

Region	Percent of Company that is Hispanic and African American Women
California	2.4287
Massachusetts	1.6574
Washington	1.46019
US	2.81539

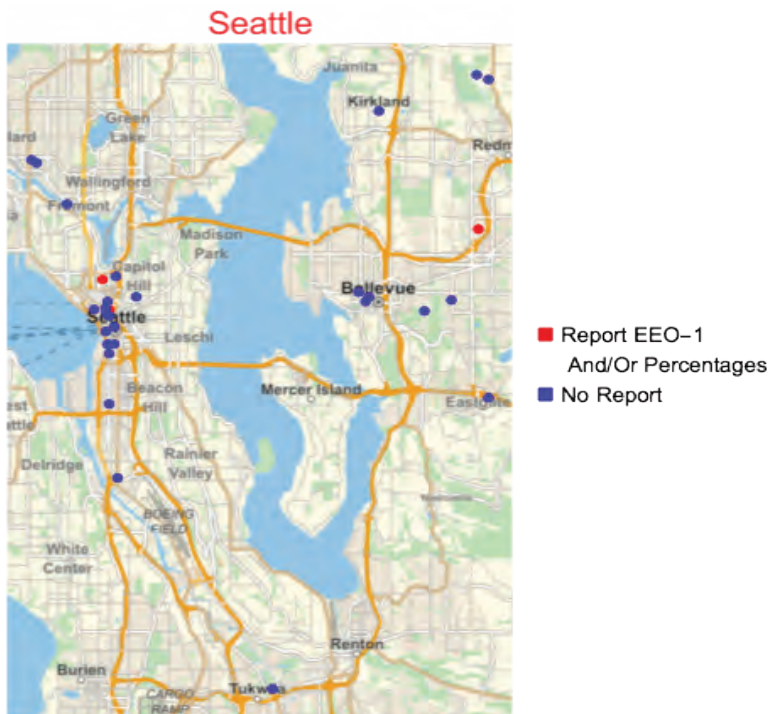
In all the studied regions, the percentage of private industry STEM workers who are Hispanic, Latino, or African American women is less than their representation in private industry in the United States of America. The maps of Silicon Valley, Boston, and Seattle (Figures 4-6) show which companies in the sample reported demographic data and which did not. California had the highest percentage of underrepresented minority representation, which corresponds to California tech companies scoring higher on commitment to inclusion than the other regions (Table 1). However, only 7 out of the 27 (25.93%) California tech companies' entire workplace demographic data had greater representation of Hispanic, Latino, and African American women than the regional average, which indicates that companies do not solely release EEO-1 data because the company's data is greater than the regional average. This might be because the California region percentage is the proportion of the total number of Hispanic, Latino, and African American women STEM private industry workers to the total number of STEM private industry workers in California; it is not the average of tech companies' Hispanic, Latino, and African American women representation. In Boston, only the BostonA company reports a representation percentage higher than the regional percentage, but as it does not publicly share its demographic data, it cannot be included in the analysis of why companies publish EEO-1 reports. In Seattle, both companies' percentages were below the regional percentage. Compared to the national representation of Hispanic, Latino, and African American women, only three companies, all headquartered in Silicon Valley, report workplace diversity that is greater than the national average. The majority of the companies' representation being below the regional and national diversity levels indicates that EEO-1 reports do not highlight model diversity in a corporation.



**Figure 4:** Silicon Valley Company Reporting Map (2016-2018).



**Figure 5:** Boston Company Reporting Map (2016-2018).



**Figure 6:** Seattle Company Reporting Map (2016-2018).

Another reason companies publish EEO-1 reports is to show that their company is transparent, which can help them win awards. The number of awards a tech company has won was calculated by adding together awards relating to the workplace, company inclusion, and diversity. Only awards that were listed on the company's website were included. The correlation coefficient between the number of awards and total demographic data reporting (EEO-1 and percentage) is 0.381441, and the correlation coefficient between reporting awards and total demographic data reporting is 0.332678. Both correlations are significant values

by the bootstrapping method, and both are moderately weak, which indicates that there is a slight relationship between award reporting and quantity and total demographic data reporting. The correlation of number awards being higher indicates that releasing demographic data might help companies win awards.

It was found that the correlation coefficient between reporting awards and publishing only EEO-1s is 0.308607 and the correlation coefficient between number of awards and EEO-1 reporting is 0.356626. These significant, moderately weak correlations suggest that publishing EEO-1s has a slight relationship with the number of awards a company receives and whether a company chooses to publish their awards.

However, are workplace awards a good metric for the commitment to inclusion in tech companies? The correlation coefficient between reporting awards and the commitment to inclusion of a company is 0.439392, while the correlation coefficient between the number of awards and the commitment to inclusion is 0.504118. Both correlation coefficients are significant values (determined by the bootstrapping method) and are moderately strong, so there is a relationship between the reporting and number of awards and commitment to inclusion. This suggests that awards are a decent way to measure commitment to inclusion in STEM companies.

While awards indicate a company's commitment to inclusion, do they measure a company's demographic diversity? The correlation coefficient between reporting awards and the representation of Hispanic, Latino, and African American women and the correlation coefficient between the number of awards reported and female underrepresented minority representation are -0.189721 and -0.0230979, respectively. These negative correlation coefficients indicate that there is an inverse relationship between awards and diversity—more awards and reporting of awards correlated with a less diverse company. This might explain why tech companies report their awards: they want their company to seem prestigious for its diversity when, in reality, its diversity is worse than that of other companies. However, these correlations are not significant values by the bootstrapping method, so it is possible that the correlations were found by chance. Also, only the EEO-1 reporting companies were included in these correlation coefficient calculations, since EEO-1 reports are used to calculate representation, so the finding cannot be generalized to all tech companies. Companies that publish awards but do not release their diversity data might be more diverse, but that question cannot be answered in this study. The negative correlation could also be due to the fact that not all companies publish the number of workplace awards they have won on their website. The negative correlations devalue awards as measures of company diversity because they suggest that companies with less minority representation report winning more awards than companies with greater workforce diversity. Therefore, awards are not an accurate measure of company diversity.

Another way to measure inclusion is through the retention of minority employees. If the workplace is intimidating, there are weak inclusion plans, and if a company does not have the right support for employees, they will stop working there. To observe the retention rates, only companies that report EEO-1s are included because EEO-1s include information on the yearly changes between the number of employees of a certain race and gender; white women is a category that includes the total number for the year the report is published and the previous

year. The overall number of employees per company is included too, so I calculated the percent change of the total company workers that are Hispanic, Latino, or African American women for each company. On some EEO-1 reports, the data for the previous year was not provided, so I recorded their change as 0. Additionally, some company's EEO-1 reports from past years were accessible, so I averaged the percent changes in employee representation in these companies. Also, I calculated the average company growth using the EEO-1 report data to see if the change in representation of underrepresented minority women matched the overall change in a company's workforce.

**Table 5: Percent Change in Representation of Underrepresented Minority Women and Company Growth.**

Tech Company	% Change in % Representation of Underrepresented Minority Women in STEM jobs	Company Growth
Apple Inc	5.10666	6.3432
Cisco Systems Inc	4.80277	0.83099
Alphabet Inc	12.9423	21.0672
Adobe Systems Incorporated	6.75462	21.7402
Intel Corporation	8.87382	2.63522
Slack Technologies Inc	19.2283	56.3511
Yahoo Inc	-18.1058	2.4522
Palo Alto Networks Inc	-8.11694	19.7174
NVIDIA Corporation	0.966656	6.47285
View Inc	-23.4783	30.6818
Twitter	52.4238	-11.4311
PayPal Inc	1.44097	4.56976
MobileIron Inc	65.1186	-9.15619
Square	0	0
LinkedIn Corporation	13.7443	5.98821
NetApp, Inc	0	0
Intuit Inc	34.6775	-7.11816
Hewlett Packard Enterprise	-14.4898	-36.2992
Lyft	38.7932	115.502
VMware Inc	8.7489	8.08772
Pandora Media Inc	17.8173	-15.7757
Salesforce.com	8.63685	19.9502
Mozilla Corporation	23.6421	1.03141
Facebook Inc	23.8103	57.9245
Pinterest	-4.35676	32.0697
eBay	25.7285	3.55021
AirBnb	0	0
BostonA	0	0
Boston Scientific Corporation	11.9343	9.33208
Microsoft Corporation	10.4997	13.0058
Amazon.com Inc	23.2969	49.3463

In 38.71% of the EEO-1 reporting companies, the percent change in the representation of underrepresented minority women is greater than the overall

company growth, which shows their commitment to increasing representation. These companies gain a reputation for open-mindedness by increasing representation, so releasing the EEO-1 report might be to inform the public of its increasing demographic changes annually. The percent change in employee diversity in 61.29% of EEO-1 reporting companies was less than the percent growth of the overall company, which indicates these companies hired fewer underrepresented minority women compared to other employee demographic groups as they expanded their workforce. Additionally, the majority of the companies (83.17%) had positive percent changes in the representation of racial minority females, which shows that these companies are interested in increasing representation among their workforces. However, 16.13% of tech companies experienced a decrease in the representation of Hispanic, Latino, and African American women; moreover, in 12.9% of the companies, the retention rate decreased while the company growth increased, suggesting that the company has a negative work environment. Thus, companies with lower retention rates may release their EEO-1 reports to gain a better reputation. In Table 6, Boston Scientific Corporation has steady company growth, around 9.3 % from 2016 to 2017 and 2017 to 2018, and the growth of representation of underrepresented minority women decreased from a 16.3% change to a 7.6% change; the percent change in the representation of Hispanic, Latino, and African American women decreased, while the company growth remained constant. This phenomenon was seen in other companies in the study, so the data shows that as companies grow at constant rates, their change in increasing diversity decreases. The pattern might mean that as companies expand in smaller amounts, fewer new minority female employees are employed than in the previous year due to the smaller increase in the company overall. Also, it could mean that the company had strong retention rates as few new hires were added, or that the workplace is hostile as it is having issues keeping minority workers, so people quit and no new hires want to join.

The correlation between the average yearly change in representation of women of underrepresented minorities and company growth was calculated to determine which of the previous explanations for the pattern is correct. The correlation coefficient is weak and positive (0.151522), which suggests that an increase in company growth is only slightly related to an increase in the percent change of representation of underrepresented minorities. Thus, for more growth in a company, the change in representation of female minority employees yearly also increases, which supports the explanation that the representation percent change is due to how many employees the company is adding as a whole. The percent change in representation and company diversity have a weak positive correlation of 0.173075, which suggests that there is a slight correlation between the yearly percent change in the representation of Hispanic, Latino, and African American women and the diversity of a tech company. These results suggest that an increase in the representation percent change results in an increase in diversity as the corporation increases its diversity while increasing its size. Combined, the two correlations show that with an increase in company growth, the rate of change for the representation of minority women increases, which leads to more diversity in the company. This shows how tech expansion is reliant on diversity, as increased diversity has a slight positive correlation with expansion.

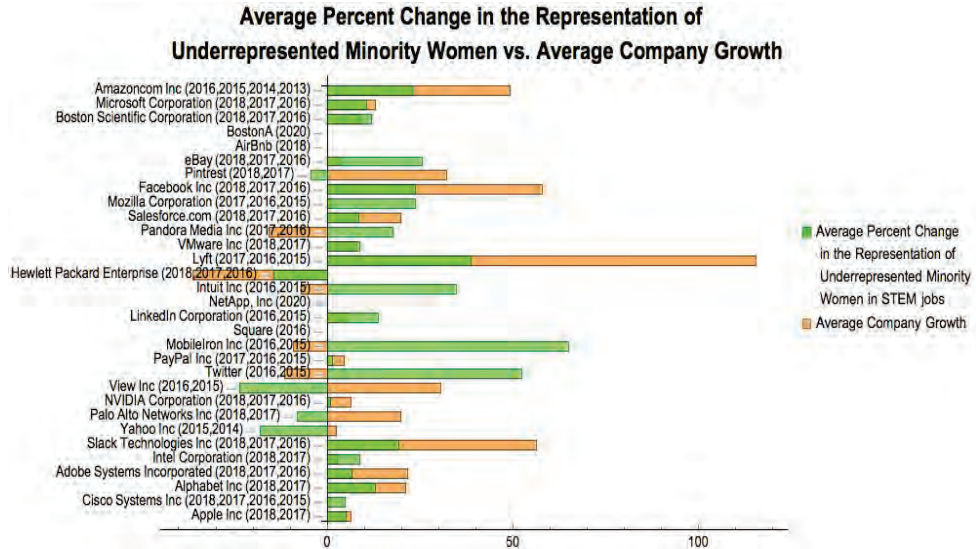


**Table 6:** *Yearly Retention among Tech Companies. (Not applicable for companies where EEO-1 data is only available for one year).*

Tech Company	EEO-1 Years	Retention of Hispanic and African American Women in STEM Jobs (%)	Company Growth
Apple Inc	2017, 2018	5.10666	6.3432
Cisco Systems Inc	2015, 2016, 2017, 2018	2.24602, -0.339903, 12.5022	3.33756, -0.863401, 0.0188162
Alphabet Inc	2017, 2018	12.9423	21.0672
Adobe Systems Incorporated	2016, 2017, 2018	7.79523, 5.71401	17.2298, 26.2506
Intel Corporation	2017, 2018	8.87382	2.63522
Slack Technologies Inc	2016, 2017, 2018	37.596, 0.860541	61.5034, 51.1989
Yahoo Inc	2014, 2015	-18.1058	2.4522
Palo Alto Networks Inc	2017, 2018	-8.11694	19.7174
NVIDIA Corporation	2016, 2017, 2018	0.938782, 0.99453	13.9304, -0.984737
View Inc	2015, 2016	-23.4783	30.6818
Twitter	2015, 2016	52.4238	-11.4311
PayPal Inc	2015, 2016, 2017	2.36751, 0.514431	7.86837, 1.27115
MobileIron Inc	2015, 2016	65.1186	-9.15619
Square	2016	N/A	N/A
LinkedIn Corporation	2015, 2016	13.7443	5.98821
NetApp, Inc	2020	N/A	N/A
Intuit Inc	2015, 2016	34.6775	-7.11816
Hewlett Packard Enterprise	2016, 2017, 2018	-34.9404, 5.96136	-66.5006, -6.09784
Lyft	2015, 2016, 2017	68.0041, 9.58213	167.85, 63,1542
VMware Inc	2017, 2018	8.7489	8.08772
Pandora Media Inc	2016, 2017	17.8173	-15.7757
Salesforce.com	2016, 2017, 2018	6.1216, 11.1521	24.1098, 15.7906
Mozilla Corporation	2015, 2016, 2017	27.1204, 20.1638	-3.85321, 5.91603
Facebook Inc	2016, 2017, 2018	32.6873, 14.9333	75.0645, 40.7846
Pinterest	2017, 2018	-4.35676	32.0697
eBay	2016, 2017, 2018	43.0748, 8.38231	10.6943, -3.59388
AirBnb	2018	N/A	N/A
BostonA	2020	N/A	N/A
Boston Scientific Corporation	2016, 2017, 2018	16.2798, 7.58876	9.42531, 9.23886
Microsoft Corporation	2016, 2017, 2018	131277, 7.87173	15.8271, 10.1845
Amazon.com Inc	2013, 2014, 2015, 2016	15.6045, 36.0465, 18.2396	47.3078, 47.7915, 52.9396

Furthermore, the correlation between percent change of the representation of underrepresented minority female employees and company commitment is negative and weak (its correlation coefficient is -0.0530153),

which suggests that there is no correlation between the percent change of representation and diversity commitment. This indicates that demographic data is a poor measurement of inclusion as there is no relationship between increasing company diversity and its commitment to inclusion.



**Figure 7:** Representation of Underrepresented Women in STEM Jobs and Total Company Employees.

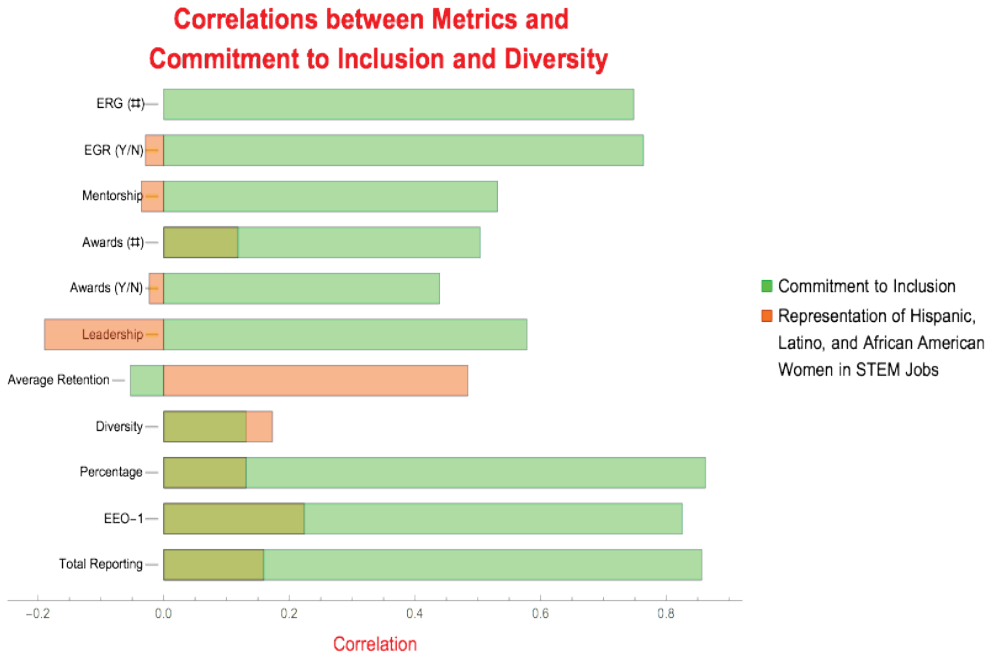
Pre-college programs and other organizations with a commitment to getting women and people of color involved in STEM have been proven by existing literature to provide support for individuals as they explore STEM. Do tech company partnerships with universities or organizations that support minorities in STEM correlate with the percent change in the representation of underrepresented minority women in STEM? To find out, the correlation coefficient between the number of partnerships a tech company has and their percent change in representation was calculated to be 0.0127577. This low correlation coefficient indicates that these partnerships do not have a relationship with the changes in the representation of Hispanic, Latino, and African American women in STEM.

The secondary literature shows that the commitment university leaders have to diversity and inclusiveness is important in developing a welcoming community (Ong et al. 228; Wilton et al. 60). In this study, leadership in companies was measured by whether there is a member of the company's executive board and/or if there is a leadership committee devoted to diversity, inclusion, and equity. Companies received two points if they had both a leader and a committee. The correlation coefficient between leadership and company commitment to inclusion (without factoring in leadership) was found to be 0.578454. The moderately strong correlation suggests that there is a relationship between the commitment to inclusion of leaders in a tech company and the company's overall commitment. The correlation coefficient between leadership and the representation of underrepresented minority women in STEM jobs is

0.484394, which suggests a moderately strong relationship. Both correlation coefficients are significant values by the bootstrapping method as they are unlikely produced due to chance. Thus, the commitment of company leaders to inclusion can be used as a metric to measure company inclusion. The correlation coefficient between leadership and diversity being less than the correlation coefficient between leadership and commitment suggests that a leader's commitment to inclusion has a greater effect on companywide inclusivity than diversity. This shows the large influence leaders have in a company, as what they are passionate about becomes what the company values. Also, as tech leadership is mostly made up of white men (more so than any other department) leadership team diversity especially needs to increase.

Additionally, secondary literature indicates that faculty and peer mentors are vital sources of support for female and minority students as they pursue STEM degrees and careers (Duran et al. 136; Palmer et al. 492; Ong et al. 237). Company mentorship has the potential to have the same effect, as company mentors help new employees navigate the company. I recorded company's mentorship opportunities; they included programs open to women or students of color in the community and programs within the company itself. Companies earned two points for having both options. The correlation coefficient between mentorship and commitment to inclusion is 0.531476, which is a significant number based on the bootstrapping method. This moderately high, positive correlation indicates that as mentorship increases, so does the company's commitment to inclusion. In contrast, the correlation coefficient between mentorship and demographic diversity is 0.118443. The slight relationship is interesting since after employees are hired, mentors would provide support, yet the low correlation suggests they have very little, if any, effect.

Finally, the secondary research identified counterpaces (environments where students can connect with students of the same background and similar interests as them) as important sources of support for female students and students of color as they pursue STEM careers (Ong et al. 207). Here, students feel comfortable and build confidence as everyone has similar experiences and understands each other. In tech companies, these groups are called Employee Resource Groups (ERGs) and share a similar purpose: employees can connect with workers like themselves. The correlation coefficients between ERG reporting and the number of ERGs and commitment to inclusion are 0.763665 and 0.748282, respectively. The correlation coefficients are significant by the bootstrapping method and indicate a very strong, positive relationship, so ERGs are a good metric of a company's inclusion. Contrarily, the correlation coefficients between reporting ERGs and the number of ERGs and the representation of underrepresented minority women working in STEM jobs is -0.0357484 and -0.0291007, respectively. The negative relationship suggests that the more ERGs or more reporting of ERGs corresponds with less demographic diversity. However, the correlation coefficients are very close to 0, which suggests that there is no correlation between ERGs (reporting and number) and the diversity of workers in STEM companies. Thus, ERGs cannot be used as a metric for the diversity of a company.



**Figure 8:** Correlation Coefficients between Metrics and Commitment to Inclusion (not including the factor tested against in the commitment score) and Diversity.

## 6. Limitations

A limitation of the results is that they are skewed towards companies that report EEO-1 data. I selected companies that publish their EEO-1 report due to the breakdown of employees by job type, race and ethnicity, and gender, but most companies do not publish this metric. Out of the over 9,000 tech companies in Silicon Valley, Boston, and Seattle, 30 published the data (BostonA does not publish their data but it was available for this study), and I included them in the total sample of 90 companies. Thus, the proportion of companies reporting EEO-1 data is much greater in this study than in the United States currently, so the results are shifted to more accurately represent what these companies practice. Furthermore, it is possible that the results on the diversity breakdown metrics are skewed to more accurately represent Silicon Valley than Boston and Seattle because 27 of the 30 EEO-1 publishing companies are located in Silicon Valley, so the diversity demographic data represents Silicon Valley more than other two cities. Silicon Valley is the tech hub of the United States and is known for being more diverse than other tech hubs, so the results are probably skewed to be higher than they are nationally. Also, for the diversity metric, I used the representation of Hispanic, Latino, and African American women in STEM jobs as women of color are the least represented group in STEM. Their diversity representation is the lowest out of any minority group, which might have shifted the results to be lower than they are nationally. Lastly, with a study of this nature, a potential reason that only correlations between the factors and commitment to inclusivity

were found is because the factors were used to calculate a company's commitment score. However, this was prevented because I compared each metric to the commitment score calculated without using the factor being compared; when I compared mentoring to inclusivity, I used the commitment score calculated without mentoring points.

## 7. Conclusion

While the public demands that tech companies release their demographic breakdowns of employees, this study found that representation data does not correlate with a company's commitment to inclusion. Inclusion is key when observing the diversity of a workplace because a welcoming environment keeps underrepresented minority employees involved in a STEM company. Thus, it is important to find which measure best reflects inclusivity. While demographic data is not an accurate representation of a company's inclusivity, reporting data (as a percentage and/or EEO-1) has a strong, positive correlation with a company's commitment to it, so it is a good metric. Furthermore, inclusion and diversity awards (reporting and amount), leadership's commitment to inclusion, mentoring programs, and ERGs (reporting and the amount) correlate with a company's commitment to inclusion, so they accurately measure inclusion. Interestingly, these metrics do not correspond with a company's diversity representation—in fact, none of the metrics studied accurately reflected the representation of underrepresented minority women in STEM jobs; retention had the highest correlation coefficient of 0.484394, which suggests a moderately low relationship. Furthermore, it was found that almost none of the companies that release their EEO-1 demographic data had representation of Hispanic, Latino, and African American women that was above the regional STEM private industry or the national private industry number, which indicates that companies that publish EEO-1s do not necessarily have model employee diversity. Thus, society must reevaluate how it determines if a company is diverse and inclusive, and what it demands from companies to be transparent, as different metrics correlate with different aspects of a company's commitment to inclusion and diversity.

## Works Cited

- Alfred, Mary V., et al. "Advancing Women of Color in STEM: An Imperative for U.S. Global Competitiveness." *Advances in Developing Human Resources*, vol. 21, no. 1, 1 Dec. 2018, pp. 114-32.
- "Apple Awards 2020." *Comparably*, [www.comparably.com/companies/apple/awards/2020](http://www.comparably.com/companies/apple/awards/2020). Accessed 13 Sept. 2020.
- Bloodhart, Brittany, et al. "Outperforming yet undervalued: Undergraduate women in STEM." *PloS One*, vol. 15, no. 6, 25 June 2020, pp. 1-13.
- Cheruvellil, Kendra S., et al. "Creating and maintaining high-performing collaborative research teams: the importance of diversity and interpersonal skills." *Frontiers in Ecology and the Environment*, vol. 12, no. 1, Feb. 2014, pp. 31-38.

- Colombo, Massimo G., et al. "The contribution of university research to the growth of academic start-ups: an empirical analysis." *The Journal of Technology Transfer*, vol. 35, 19 Mar. 2009, pp. 113-40.
- Di Gregorio, Dante, and Scott Shane. "Why do some universities generate more start-ups than others?" *Research Policy*, vol. 23, no. 2, Feb. 2003, pp. 209-27.
- Duran, Antonio, et al. "A Critical Quantitative Analysis of Students' Sense of Belonging: Perspectives on Race, Generation Status, and Collegiate Environments." *Journal of College Student Development*, vol. 61, no. 2, Mar.-Apr. 2020, pp. 133+.
- "EEO-1 Survey." *U.S. Equal Employment Opportunity Commission*, www.eeoc.gov/employers/eo-1-survey. Accessed 2 Aug. 2020.
- "Fact Sheet for EEO-1 Survey Filers." *U.S. Equal Employment Opportunity Commission*, www.eeoc.gov/employers/eo-1-survey/fact-sheet-eeo-1-survey-filers. Accessed 2 Aug. 2020.
- Lee, Jenni. "What Young Women in Tech Really Need." *Lean Out: The Struggle for Gender Equality in Tech and Startup Culture*, edited by Elissa Shevinsky, OR Books, 2015, pp. 173-78.
- Levison, Katy. "Sexism in Tech." *Lean Out: The Struggle for Gender Equality in Tech and Startup Culture*, edited by Elissa Shevinsky, OR Books, 2015, pp. 47-60.
- Moretti, Enrico. "Forces of Attraction." *The New Geography of Jobs*, Boston, Houghton Mifflin Harcourt, 2012, pp. 121-53.
- Ong, Maria, et al. "Counterspaces for Women of Color in STEM Higher Education: Marginal and Central Spaces for Persistence and Success." *Journal of Research in Science Teaching*, vol. 55, no. 2, Feb. 2018, pp. 206-45.
- Palmer, Robert T., et al. "A Qualitative Investigation of Factors Promoting the Retention and Persistence of Students of Color in STEM." *The Journal of Negro Education*, vol. 80, no. 4, Fall 2011, pp. 491-504.
- Pietri, Evava S., et al. "Addressing Unintended Consequences of Gender Diversity Interventions on Women's Sense of Belonging in STEM." *Sex Roles*, vol. 80, no. 26, 17 Aug. 2018, pp. 527-47.
- Samila, Sampsa, and Olav Sorenson. "Community and Capital in Entrepreneurship and Economic Growth." *American Sociological Review*, vol. 82, no. 4, Aug. 2017, pp. 770-95.
- Shevinsky, Elissa. "The Pipeline Isn't the Problem." *Lean Out: The Struggle for Gender Equality in Tech and Startup Culture*, edited by Shevinsky, OR Books, 2015, pp. 203-16.
- Stephens, Bryan, et al. "Austin, Boston, Silicon Valley, and New York: Case studies in the location choices of entrepreneurs in maintaining the Technopolis." *Technology Forecasting and Social Change*, vol. 146, Sept. 2019, pp. 267-80.
- Stoet, Gijsbert, and David C. Geary. "The Gender-Equality Paradox in Science, Technology, Engineering, and Mathematics Education." *Psychological Science*, vol. 29, no. 4, 14 Feb. 2018, pp. 581-93.
- Thébaud, Sarah, and Maria Charles. "Segregation, Stereotypes, and STEM." *Social Sciences (Basel)*, vol. 7, no. 7, 9 June 2018, pp. 1-18.

- U.S. Equal Employment Opportunity Commission. *Diversity in High Tech*. May 2016, [www.eeoc.gov/sites/default/files/migrated\\_files/eeoc/statistics/reports/hightech/diversity-in-high-tech-report.pdf](http://www.eeoc.gov/sites/default/files/migrated_files/eeoc/statistics/reports/hightech/diversity-in-high-tech-report.pdf). Accessed 14 Aug. 2020.
- "US Tech Hubs." BuiltIn, [builtin.com/tech-hubs](http://builtin.com/tech-hubs). Accessed 15 Sept. 2020.
- Wilton, Leigh S., et al. "In Good Company: When Gender Diversity Boosts a Company's Reputation." *Psychology of women quarterly*, vol. 43, no. 1, Mar. 2019, pp. 59-72.



# Can Corporations Balance Economic Profit with Corporate Social Responsibility?

---

*Annapurna Ananya Maddali*

**Author Background:** *Annapurna Ananya Maddali grew up in the United States and currently attends The Quarry Lane School in Dublin, California in the United States. Her Pioneer research concentration was in the field of economics and titled “The Politics of the Gig Economy: Entrepreneurship, Technological Innovation, and Politics.”*

## Abstract

This paper will examine the idea of social responsibility, and how corporations today have embraced this concept in their approach to social debates and social change. Milton Friedman, in his 1970 *New York Times Magazine* article on this topic, argued that the only social responsibility a business has is to “increase its profits.” Accordingly, Friedman is well known as a critic of social responsibility in business. This paper will reevaluate Friedman’s main argument by asking whether corporations and businesses can create value for society — by advancing social responsibility — that transcends their normal profits-seeking function.

Friedman’s argument that a business’s only social responsibility is to produce profits is directly in conflict with John Mackey’s idea of what the role of a business is. John Mackey, the co-founder of Whole Foods, believes that corporate philanthropy is a good thing for businesses and the long-term benefit of stakeholders and investors. Mackey is an advocate of the idea of social responsibility of firms, meaning he believed that firms can and should create some benefit for society. Friedman rejected this idea of social responsibility, claiming the only responsibility a business has is to maximize profits for shareholders.

By conducting a case study using one of the leading corporations of today, Whole Foods, this paper will strive to answer the question of whether big businesses today can create profits for shareholders and create economic value for themselves, while also creating non-economic value for themselves that is beneficial for causes such as environmental sustainability and poverty. In other words, this paper will investigate whether both ideas can be simultaneously achieved, or if a tradeoff exists.

## Introduction

In today’s economic climate, it is not enough for a corporation to simply generate a product or service that is in demand. In addition to creating a valuable and useful product or service for consumers, many corporations are also giving back to society.



Corporations share their wealth with the parts of society that can benefit from these contributions in two ways: one way is by giving part of their profits to charity; the other way is even more importantly, by pursuing goals — or social responsibilities — that are not profitable and may even reduce their profits. Social responsibility is when businesses, in addition to maximizing profits and shareholder value, act in a manner that benefits society.<sup>1</sup> Social responsibility is becoming increasingly important to both investors and consumers because it reveals that the true values and goals of a corporation are not just to be profitable, but also to contribute to the common good by improving societal welfare. On the surface, social responsibility seems like a very ethical and morally admirable choice for corporations to pursue as a goal.

However, the idea of social responsibility raises a very important question: are companies pursuing social responsibility, whether it be through monetary donations or by committing themselves to causes, because they truly believe that the morality and justice of the causes are just as, or even more, worthwhile than profits? Or are they engaging in socially responsible actions to build a good brand image and public reputation, which generates greater profits? Milton Friedman, in “The Social Responsibility of Business is to Increase its Profits,” argues that a business itself does not have a social responsibility to the public. This is because, according to Friedman, a business is an artificial association assembled by individuals for the specific purpose of making profits. While individuals within a business can have social responsibilities, whether it be to their families, countries, or other entities, Friedman believes that a business is *only* responsible for its shareholders — who, after all, created that artificial association.<sup>2</sup> Friedman argues that if a corporation pursues social responsibility, it is stealing from shareholders, or consumers, or employees. In other words, using money that is not the possession of the corporate executive for a purpose to which shareholders have not consented to fulfill a social responsibility reduces the ability of a business to do what it is supposed to do: generate and maintain steady profits.

Mackey, in opposition to Friedman, believes that individuals within a corporation do have a diversity of values that the company pursues together; they do not just pursue profits. This combination of the values of individuals reflects on the corporation, becoming the corporation’s social responsibility. The social responsibility and contributions that corporations have and make do not only create social value by improving it for the better but can also make the corporation itself very valuable and increase customer loyalty and brand image.

To understand whether a corporation can balance economic profits with social responsibility, this paper will employ a distinction between economic and non-economic value. There is a difference between a corporation creating value for society by being socially responsible and improving societal welfare, and generating profits for shareholders, creating value for consumers, and capturing economic value in the long-term. But can these two goals be combined?

While Friedman’s view is orthodox in economics and business, recently a new trend called “shared-value creation” has emerged. The theory of “shared-value creation” was developed by Harvard professors, and is the practice of creating economic value in a way that also creates value for society by addressing its needs and

---

<sup>1</sup> Ganti, A., 2020. *Social Responsibility Definition*. [online] Investopedia. <https://www.investopedia.com/terms/s/socialresponsibility.asp> [Accessed 17 August 2020]

<sup>2</sup> Friedman, Milton. “A Friedman Doctrine-- The Social Responsibility Of Business Is to Increase Its Profits.” *The New York Times*, *The New York Times*, 13 Sept. 1970, [www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html](http://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html).

challenges.<sup>3</sup> It has become increasingly popular among leading corporations as a framework for redefining their corporate values. Having shared values means that employees, stakeholders, and investors share common work attitudes and sentiments, which builds a shared interest in success.

Huge companies like Johnson & Johnson, Google, and Walmart donate millions of dollars to local and global causes like public health, workforce education, community development, and more.<sup>4</sup> These companies choose these causes and organizations that require funding to continue doing what they do. Contributing and donating money to fund and aid such humanitarian and welfare organizations that are doing something for the betterment of society, especially when a corporation generates as much money per annum as Google does, means that a business and the individuals within the business are not just upholding and fulfilling their social responsibility. The corporations are also improving society and the structure of society overall by adding value through their injections of money and donations into these organizations. The social value that a company creates includes the collective benefit and betterment of a community. By donating money, that is exactly what these corporations are doing: donating money and associating their large company names with organizations working towards notable causes. Doing so is what brings the value that corporations produce for society. These companies, by fulfilling their social responsibility and generating great profits, are proving that corporations can simultaneously create value and capture value. Many companies are skewed in their values, with some only looking to generate profits, and others only looking to create value by bettering society. But these companies, rather than creating profits or improving aspects of society, are instead simultaneously creating profits and improving aspects of society.

The question now becomes: Can larger corporations create economic profits for themselves while also benefiting society and doing good with their social responsibility? Then, is there another way for corporations to practice both at the same time in a way that benefits all aspects of society?

## Theory

Milton Friedman was an incredibly influential economist in the second half of the 20th century, best known for his belief in a free-market system.<sup>5</sup> In his essay called “The Social Responsibility of Business is to Increase its Profits,” Friedman argued that businesses do not really have any public social responsibility, and the only responsibility that they really have is to shareholders. Friedman develops an “agent theory of the corporate executive,” and defines an executive as an employee of the shareholders, who are the owners of the corporation. As a result of this definition, Friedman argues that the corporate executive, as an employee of the board, has an obligation to advance the board’s agenda, which includes within it primarily the priority of creating a profitable company. Friedman states that the corporate

---

<sup>3</sup> “Creating Shared Value.” FSG, July 17, 2018. <https://www.fsg.org/publications/creating-shared-value>.

<sup>4</sup> Greenwood, Chelsea. “10 Of the Companies That Give the Most to Charity in the US.” Business Insider. Business Insider, November 14, 2018. <https://www.businessinsider.com/companies-that-give-the-most-to-charity-in-the-us-2018-9>.

<sup>5</sup> Chappelow, Jim. “Milton Friedman Definition.” Investopedia. Investopedia, February 5, 2020. <https://www.investopedia.com/terms/m/milton-friedman.asp>.

executive “has a direct responsibility to his employers. That responsibility is to conduct the business per their desires, which generally will be to make as much money as possible while conforming to their basic rules of the society.”<sup>6</sup> This means that the money that corporate executives spend is really the money of the shareholders, so spending this money on social causes to fulfill social responsibility is the same as spending someone else’s money for an executive’s purpose. Many critics of Friedman have rejected his definition of social responsibility by saying it is too narrow and controversial. Critics argue that Friedman’s definition does not do any justice to the multiple aspects and goals that can be acknowledged by shareholders of a corporation, like, for example, the goals of being profitable while also being socially responsible.

Based on this definition, Friedman next makes a set of very bold and controversial arguments that render corporate social responsibility financially irresponsible and a violation of the fiduciary duties of the executive. In effect, Friedman goes on to argue that spending for corporate social responsibility involves theft “insofar as [a corporate executive’s] actions in accord with his ‘social responsibility’ reduce returns to stockholders, he is spending their money. Insofar as his actions raise the price to customers, he is spending the customers’ money. Insofar as his actions lower the wages of some employees, he is spending their money.”<sup>7</sup> In other words, spending a corporation’s money to fulfill social responsibility decreases the profits that the shareholders receive because that money is going somewhere else rather than going to the shareholders. If this money is being taken away from shareholders, then they may move to increase prices to compensate for their loss, which would result in consumers paying more for the same products or services. If consumers pay more, they are spending more of their own money to make it possible for the executive to be socially responsible. Similarly, instead of increasing prices, shareholders may move to cut the wages of some employees, which means the employees’ money is also being spent to fund whatever socially responsible cause an executive is choosing to pursue.

Therefore, if a corporate executive is socially responsible and spending money, that money is either coming from the shareholders, the customers, or the workers, so the money an executive is using is not their own. Since the decision about which cause to support is likely to be disproportionately shaped by the executive’s preferences, even if an executive consults the shareholders, he or she is pushing the business to spend on whatever cause they deem worthy — and very likely to the detriment of the company’s profitability. Friedman wants to show us that any attempt to establish a collective consensus on some controversial cause among the shareholders, and between the shareholders and the executive, is likely to be only an illusion. Friedman says that, instead, a business should “use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game, which is to say, engages in open and free competition without deception or fraud.”<sup>8</sup> In other words, because there is no way for a business to collectively be socially responsible, Friedman says to concentrate on increasing profits by allocating most of the resources of a company in the most efficient manner without spending too much time on what the social responsibility of a business should be. Doing so, he

---

<sup>6</sup> Friedman, Milton. “A Friedman Doctrine - The Social Responsibility Of Business Is to Increase Its Profits.” *The New York Times*, 13 Sept. 1970, [www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html](http://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html).

<sup>7</sup> Ibid.

<sup>8</sup> Ibid.

argues, will increase profits as long as a corporation continues to conduct itself in a way that is corporally ethical and non-fraudulent.

In a 2005 debate, John Mackey responded to Friedman's well-known argument by developing an alternative framework for thinking about profits and social responsibility in a corporate setting. Mackey is an American businessman who co-founded the well-known corporation Whole Foods in 1980, but is also an advocate for social responsibility, and what he calls "conscious capitalism". Mackey argues that Friedman's view that a business is an artificial entity that cannot have any social responsibilities is too narrow. This is because Friedman's definition cannot do justice to business operations of Whole Foods along with those of many other companies. Mackey argues that Friedman undersells the humanitarian or philanthropic aspect of capitalism. Though Mackey agrees with Friedman that profits are important for a business, he believes that profits should not be the only focus of a business. Instead, in Mackey's opinion, businesses have multiple objectives that they pursue simultaneously, and which reflect the priorities of multiple different "stake-holders" in the business. Shareholders, for Mackey, are but one of several different kinds of stake-holders in a company. Mackey writes, "From an investor's perspective, the purpose of the business is to maximize profits. But that's not the purpose for other stakeholders—for customers, employees, suppliers, and the community."<sup>9</sup> Even though the goal of the investors is to maximize profits, other stakeholders - not to be confused with shareholders - define the purpose of the business in terms of its own needs and desires. According to Mackey, a healthy business should embrace all these stakeholders and balance their values, instead of limiting itself only to the priorities or interests of investors — who are, after all, only one type of stakeholder. To clarify, shareholders own part of a public company through shares of stock, while stakeholders have an interest in the performance of a company for reasons other than stock performance or appreciation.<sup>10</sup> Shareholders, in other words, are just one of several different sorts of stakeholders existing within a company.

Stakeholders can be consumers, employees, shareholders, and communities. Each group of people has a certain set of values which together make up the business values because these people represent the business. However, as soon as Mackey begins discussing how to balance these potentially divergent values, and how to think of the relationship of corporate philanthropy to profits, a certain confusion begins to appear in his framework. For instance, Mackey argues that a certain amount of corporate philanthropy is good not just for ethical reasons, but also because it improves the social reputation of the business. The latter makes the business profitable because customers are loyal to businesses with good reputations, and thus advances the long-term benefits of investors. Mackey then makes his case by using Whole Foods and its philanthropic responsibility as an example. In addition to donating profits throughout the whole year, Whole Foods has "5% Days" on which 5% of the store's sales are donated to organizations the corporation deems worthwhile. This practice, according to Mackey, attracts hundreds of new customers, many of whom become regular consumers. He concludes that 5% Day not only allows Whole

---

<sup>9</sup> Milton Friedman, John Mackey. "Rethinking the Social Responsibility of Business." Reason.com - Free Minds and Free Markets. Accessed August 28, 2020. <https://reason.com/2005/10/01/rethinking-the-social-responsi-2/?print=>.

<sup>10</sup> Banton, Caroline. "Shareholder vs. Stakeholder: What's the Difference?," March 13, 2020. <https://www.investopedia.com/ask/answers/08/difference-between-a-shareholder-and-a-stakeholder.asp>.

Foods to support worthwhile causes, but it is also a successful marketing tactic that has benefited Whole Foods' investors.

Mackey continues saying that shareholders of a public company, like Whole Foods, own their stock voluntarily. If they do not agree with the philosophies or social responsibilities of a business, they can always sell their investment, or stocks, and exit their "relationship" with the company. He says that many shareholders can come to amicable decisions about philanthropic policies most of the time, but also brings up the question of where to draw the line. Mackey discusses how even though Whole Foods has responsibilities to the community which, for him, includes employees, customers, and even the suppliers, the company does not neglect its responsibilities to investors. Mackey also stresses that some of the philanthropic policies were built into the company's identity early on in its history: the 5% donation policy, which Mackey acknowledges is an arbitrary amount, was something that Mackey and his co-founders agreed on when they were starting the company in the 1980s, and ever since then the company has been transparent to all its investors about his preexisting priority. Mackey claims that in his experience, "most investors understand that it [corporate philanthropy and social responsibility] can be beneficial to both the corporation and to the larger society."<sup>11</sup> This means that investors do want to see corporations maximizing profits so they get their return on investments, but at the same time, understand that a company can generate profits while still having a social responsibility and giving some of those profits back to the community.

Friedman responded to Mackey's essay, claiming that Mackey had proved the exact point Friedman argued in his essay. Friedman begins by stating that Mackey's idea of corporate philanthropy being beneficial is simply wrong. He uses Mackey's example of 5% Days, claiming that these supposedly charitable days hosted on behalf of the Whole Foods corporation is only for the benefit of the stakeholders. He says Whole Foods was within their rights of spending their own money, "using 5 percent of one part of their [Whole Foods'] wealth to establish, thanks to corporate tax provisions, the equivalent of a 501c(3) charitable foundation."<sup>12</sup> Friedman sets up his argument by claiming that a mere 5% of Whole Foods' profits is the equivalent of setting up a charity, but it is a charity without a mission statement or any separate bylaws. He then argues by posing a question: why not just pay stockholders their dividends, or the money that rightfully belongs to them, and let them spend it on whatever causes or purpose they want, rather than distributing the money on behalf of the corporation? Friedman answers his questions by saying "a stockholder can make a larger gift for a given after-tax cost if the corporation makes the gift on his behalf than if he makes the gift directly."<sup>13</sup> He defends his argument by continuing to say that if a corporation donates money to a socially responsible cause on behalf of stockholders, the deductibility of corporate charity is eliminated.

Friedman also argues that there is a clear distinction between social responsibility and responsibility in general, such as financial responsibility. He says the goal of maximizing profits is one with an end from a private property point of view, meaning it can be completed. However, Friedman claims, maximizing profits from a social responsibility perspective is not one that can be fulfilled because there is always a greater profit that can be made if a corporation wants to be socially

---

<sup>11</sup> Milton Friedman, John Mackey. "Rethinking the Social Responsibility of Business." Reason.com - Free Minds and Free Markets. Accessed August 28, 2020. <https://reason.com/2005/10/01/rethinking-the-social-responsibility-2/?print=>.

<sup>12</sup> Ibid.

<sup>13</sup> Ibid.

responsible. Friedman discredits Mackey's idea of a corporation being able to maximize profits while entertaining the notion of being socially responsible. In Friedman's opinion, the social responsibility of a private corporation always leaves corporations wanting more of a profit. On the other hand, the financial or legal responsibility of a private property "is a sophisticated means of enabling people to cooperate in their economic activities without compulsion."<sup>14</sup> This compulsion that Friedman is referring to is the compulsion or desire to use profits and make donations to continue being socially responsible. This system based on responsibility in a free market allows people to focus on maximizing their economic activities and separate their corporate responsibilities of maximizing profits from their personal social responsibilities. This separation, Friedman concludes, "assure[s] that each resource is used for its most valued use and is combined with other resources in the most efficient way."<sup>15</sup> In other words, all resources of a corporation - whether it be labor, capital, or land - are to be used as efficiently as possible for the sole purpose of fulfilling a corporation's primary goal of maximizing profits.

Friedman and Mackey agree about certain aspects of a corporation's responsibility in society but disagree on the shared value creation aspect. Friedman completely discredits the notion of shared value creation, believing that corporations do not create any value for society because any value they create for society by being socially responsible is in an attempt to create greater value for themselves. Instead of being "socially responsible", not out of the goodness of their hearts, but to make themselves more appealing, Friedman says these corporations should just disregard the idea of social responsibility and focus solely on their profits. Alternatively, Mackey believes that social responsibility can be achieved, which firms contribute to and benefit society while also creating an economic profit.

While Friedman urges corporations to focus on maximizing profits and making themselves more valuable, Mackey emphasizes the idea of social responsibility where corporations can create value for society through their social responsibility while simultaneously generating large economic profits.

### **Case Study: Whole Foods**

Whole Foods is one of the many corporations that emphasizes social responsibility. Founded in 1980 by John Mackey, Renee Lawson Hardy, Mark Skiles, and Craig Weller, Whole Foods started as a small store that sold natural and organic products in Austin, Texas, but quickly expanded to become a multinational supermarket chain. The corporation has prided itself on staying true to its mission statement of selling only the highest quality organic and natural products. Another thing Whole Foods highly values is giving back to the community. Each store has a unique way of connecting with their community, whether it be through charity events, donations, or setting up organizations and foundations targeting global issues such as community education. Encouraging its chains to set up these foundations devoted to specific causes in society — such as environmental preservation and public education — is how Whole Foods corporation defines its social responsibility. Additionally, the corporation implemented an annual event called the 5% Day to donate 5% of all after-tax proceeds and profits to a charity of choice. According to the company balance sheets, the actual income tax expense for the fiscal years indicated differs from the

---

<sup>14</sup> Ibid.

<sup>15</sup> Ibid.

amount computed by applying statutory corporate income tax rates to income before income taxes. In 2014, the excess charitable contributions made by Whole Foods was \$1 million. In 2015, this amount increased to \$9 million, and in 2016 to \$10 million.<sup>16</sup> This pattern of increasing their charitable donations and money they allocate towards doing something socially beneficial as their profits increase shows that they truly do value social responsibility. But their social responsibility extends beyond just donating to third parties; the company has also set up three separate foundations to further their community contribution.

The first organization, called Whole Planet Foundation (WPF), was founded in 2005. WPF focuses on promoting self-sufficiency for its partners and suppliers in developing areas of the world. The initiatives WPF pursues with its partners include combating poverty, aiding natural disaster relief, and providing micro-credit loans. WPF identifies and partners with microfinance organizations (MFIs) who provide and administer responsible financial services to their members. Since its establishment, WPF has funded over 4 million microloans and interacted with people across 77 countries while helping entrepreneurs improve their own lives in countries where Whole Foods Market sources products.<sup>17</sup> By helping people who are not in financially stable positions with loans, Whole Foods has allowed thousands of people to pursue their dream of building businesses, all while building value for themselves as well.

Another organization Whole Foods has developed is the Whole Kids Foundation, founded in 2011. It is a nonprofit 501(c)(3), founded by Whole Foods Market and guided by the same values, principles, expertise, and standards for quality ingredients, food production methods and nutrition as Whole Foods. The organization supports community education initiatives to inspire families and schools to improve children's nutrition and wellness. Since it was established, this organization has connected more than 8 million schoolchildren — across the United States, UK, and Canada — with regular access to natural foods with real nutritional benefits. All of the proceeds donated to this cause make a direct impact on the children in schools and communities, because Whole Foods itself covers the cost of the Whole Kids Foundation operating budget.<sup>18</sup>

The last organization founded by Whole Foods is called the Whole Cities Foundation and was established in 2014. Their mission is to focus on promoting healthy eating education and access to nutritious food for under-served communities. Since its start, Whole Cities Foundation has partnered with over 191 nonprofits across 106 cities in the United States and Canada and has contributed more than \$2.1 million to local communities. It offers grants to nonprofits with food access projects like urban farms, mobile markets, and community gardens which provide fresh and affordable produce closer to homes.<sup>19</sup>

With all of these organizations, Whole Foods has done more for society than simply donate a percentage of their profits; they have actually founded organizations associated with their corporation to interact with consumers, and

---

<sup>16</sup> United States Securities and Exchange Commission Form 10-K Whole Foods Market, Inc. (n.d.). Retrieved September 26, 2020, from <https://www.sec.gov/Archives/edgar/data/865436/000086543616000366/wfm10k2016.htm>

<sup>17</sup> Our Impact. (n.d.). Retrieved September 07, 2020, from <https://wholeplanetfoundation.org/our-impact/>

<sup>18</sup> Impact. (n.d.). Retrieved September 07, 2020, from <https://www.wholekidsfoundation.org/impact>

<sup>19</sup> Our Approach. (n.d.). Retrieved September 07, 2020, from <https://www.wholecitiesfoundation.org/what-we-do/our-approach>

create more value rather than just making hefty donations. The social responsibility of Whole Foods is to provide communities of people with healthy, nutritious food and lifestyles. By contributing so much to the common people through its environmental, social, and humanitarian values, Whole Foods has built a brand or reputation of being an all-natural and organic supermarket chain that provides consumers with the best possible experience. But they are also widely regarded as a highly philanthropic corporation. This means that Whole Foods has built significant value for society by making it a better place through contributions, and, as John Mackey would claim, they have been pursuing their social responsibility and improving societal welfare, while simultaneously generating high profits.<sup>20</sup> This practice called shared value creation, or the idea of the competitiveness of a company and the health of the communities around it being dependent, has allowed for Whole Foods to maintain their profits and continue to be one of the top performing supermarket corporations, while also being one of generous and philanthropic through donations and setting up foundations.

Though Mackey would claim that this social responsibility differentiates Whole Foods from competitors, Friedman would argue that this social responsibility does not exist at all because, along with their reputation, their consumer base and consumer loyalty to Whole Foods increase in addition to the profits of Whole Foods. Friedman would argue that this social responsibility is, in fact, a marketing tactic, because Whole Foods' motivation and intentions behind their social responsibility is driven by their ability to make profits from it in the long-run. Therefore, the whole meaning of being socially responsible towards a cause is undermined because their social responsibility is being pursued with intentions other than doing something for the common good.

Though their approach to social responsibility is debatable, it is important to recognize Whole Foods' impact and how they differentiate themselves from competitors to make themselves the consumer favorite. Whole Foods has significant competition with firms such as Sprouts and Trader Joe's, which both have a similar mission of providing consumers with healthy and organic high-quality produce. But these competitors have not created nearly as much social value for society as Whole Foods. Although Trader Joe's and Sprouts do make donations and have partnerships with charitable organizations, they have not embedded the philanthropic work into their core business values in the same way that Whole Foods has. This is also reflected in the fact that neither Sprouts nor Trader Joe's has established the types of philanthropic foundations that we observed with Whole Foods. Whole Foods has created multiple organizations, each with a different goal, to address different problems faced by their consumers and the general public. Mackey and the shareholders of Whole Foods believe that every person who interacts with the corporation is a stakeholder. These stakeholders' best interest is addressed and implemented with the establishment of these organizations founded by Whole Foods, whether it be access to natural and healthy food, or following their dreams to become entrepreneurs.

Whole Foods has arguably created more social value for society, whether it be by providing consumers with healthy and organic food alternatives, or directly donating and contributing to society to make it better. Let us try to understand the

---

<sup>20</sup> Milton Friedman, John Mackey. "Rethinking the Social Responsibility of Business." Reason.com - Free Minds and Free Markets. Accessed August 28, 2020. <https://reason.com/2005/10/01/rethinking-the-social-responsi-2/?print=>.



real impact Whole Foods has created on society. If Whole Foods and all their organizations suddenly disappeared, it would significantly impact society compared to if Trader Joe's or Sprouts disappeared. This is because, even though all three corporations have a similar goal of providing consumers with healthy food substitutes, Whole Foods has made it their mission to not only involve themselves in their consumer food choices, but also directly involve themselves in consumer lives. So, if Whole Foods did disappear, all the benefits and contributions they could make in the future that positively impact consumer lives would also vanish, leaving communities which interact with Whole Foods, and maybe even be dependent on Whole Foods contributions, abandoned. So, the social responsibility of Whole Foods has become an asset to the corporation because their social responsibility has made it so that certain communities of people are dependent on Whole Foods and their contributions, as well as their healthy and organic produce.

This hypothetical situation that evaluates the effect that the disappearance of Whole Foods would have on consumers and the community the corporation impacts disproves Friedman's idea of the only responsibility a corporation has being its profitability. The approach of Whole Foods disagrees with both Friedman and Mackey's stances. Friedman claims that businesses should not focus on having social responsibility because any value they create for society is with artificial intentions. It is important to differentiate between creating artificial value and having artificial intentions. Friedman is claiming that any value they create is artificial with artificial intention because they are not intrinsically motivated to improve society but are instead motivated by the reward they will receive in the form of profits. This argument is supported by Mackey's statement about how Whole Foods uses social responsibility as a marketing tactic, proving that their social responsibility could be pursued with an artificial intent, like Friedman said. However, in contrast to what Friedman claimed, this "social responsibility" does have benefits to society and positively impacts different communities of people, making Whole Foods' social responsibility to communities beneficial to an extent.

## Shared Value Creation

Shared value creation involves creating economic value in a way that also creates value for society by addressing its needs and challenges. The academic concept was introduced by Professors Michael Porter and Mark R. Kramer from the Harvard Business School in 2011 in an article called "Creating Shared Value". In their publication, Porter and Kramer claim that a profit maximizing approach has been one that companies have adopted as their main business strategy. These companies do everything in their power to increase profits for their firms. However, many corporations have begun to adopt a socially responsible approach to share their wealth with the community. Porters and Kramer argue that these corporations are not adopting this idea of being socially responsible very effectively, leading to huge economic losses. Many corporations also fail to consider the long-run profits and are focused on being socially responsible in the short-run. Porter and Kramer then introduce creating shared value instead of focusing on social responsibility because creating shared value accommodates the profit-maximizing interests of firms and their combined interest in doing good for society.<sup>21</sup> Corporations such as Johnson and

---

<sup>21</sup> Porter, Michael E. and Kramer, Mark R. (2015, August 25). Creating Shared Value. Retrieved September 13, 2020, from <https://hbr.org/2011/01/the-big-idea-creating-shared-value>

Johnson, IBM, and Google have all begun to adopt and create shared value, which has benefitted everyone who interacts within and with their companies.

This idea of creating shared value is different from having corporate social responsibility because rather than just focusing on bettering society, creating shared value also focuses on creating economic profits for a firm.<sup>22</sup> While social responsibility is constructed and derived from personal values of individuals or a group of people in a firm, shared value creation focuses on difference techniques and “strengthening the local cluster of supporting suppliers” and other entities, in order<sup>23</sup> to increase the efficiency and sustainability of corporations while also keeping community welfare and benefits in mind.<sup>24</sup>

Many corporations such as Whole Foods claim that they have an immense amount of social responsibility, which is an economic cost to them. Mackey claims that Whole Foods is socially responsible, which comes at a direct economic cost in the short run. However, from being socially responsible, the goal of the corporation is to attract more consumers. As we saw in the previous section, this makes their social responsibility a direct economic cost in the short-run, but one from which Whole Foods makes both a direct and indirect economic profit in the long-run. This relationship between short-run and long-run profits complicates Mackey’s claim that Whole Foods is acting, in part, purely out of social responsibility in a self-sacrificial way. The corporation creates social value for different communities of people by providing healthy meals, and giving out microloans, which does come at an economic cost to them. However, as we saw, part of the motivation behind being socially responsible is making a profit in the long run by using social responsibility as a marketing tactic to draw consumers in. Mackey clearly states that “[donating] not only allows us [Whole Foods] to support worthwhile causes but is an excellent marketing strategy that has benefited Whole Foods investors immensely,”<sup>25</sup> his intention to be socially responsible and contribute to the common good is overlooked by his intention to generate profit, complicating any social responsibility he and Whole Foods claimed to have. Although this completely contradicts the idea of social responsibility, it does show that those who practice this idea of having social responsibility reveal an internal confusion about how they think about the relationship between responsibility and profits, and which one ultimately takes priority in their business philosophy.

Integrating these good intentions with profit-making intentions transforms this from social responsibility to the desire to create shared value. By definition, corporate social responsibility seeks to add value to an organization's activities by ensuring they have a positive impact on society, the environment, and the economy.<sup>26</sup> But corporate social responsibility does not focus on the profitability of the organization; instead it focuses on how much it can improve societal welfare and improve aspects of society based on the values of people within these organizations. This decision between being socially responsible or just maximizing profits places companies in an unpleasant bind from which they have no productive way out. These

---

<sup>22</sup> Ibid

<sup>23</sup> Ibid

<sup>24</sup> Ibid

<sup>25</sup> Friedman, M., Mackey, J., & Rodgers, T. (2020, August 21). Rethinking the Social Responsibility of Business. Retrieved September 13, 2020, from <https://reason.com/2005/10/01/rethinking-the-social-responsibility-2/>

<sup>26</sup> Corporate Responsibility: Factsheets. (n.d.). Retrieved September 13, 2020, from <https://www.cipd.co.uk/knowledge/strategy/corporate-responsibility/factsheet>

companies can either focus only on profits, and potentially drive away customers, or they can engage in a socially responsible activity or causes which will decrease the profitability of their company. However, in his essay, Friedman argues that companies do not even need to decide between social responsibility and profit maximization. He says some things that appear to be socially responsible actions benefit companies, and in that case, this social responsibility is actually intelligent and sensible marketing strategies.<sup>27</sup> Whole Foods, instead of being socially responsible to only benefit society, is doing so because by doing good for society, they can simultaneously generate a profit.

Creating shared value has been incredibly effective and useful for businesses to create economic value and profit from contributing to and solving social challenges. Creating shared value allows businesses to pursue “mutually supportive”<sup>28</sup> activities. Businesses rely on consumers and communities to buy their products. Communities, in return, rely on businesses for job creation and economic development and advancement. This idea of creating shared value encourages corporations to focus on long-term strategies and growth, rather than short-term profit maximization goals. It also encourages businesses to invest in local markets and communities to strengthen the productivity of local suppliers and drive down internal business costs of production.<sup>29</sup> With this information, understanding what Whole Foods has done and how it has interacted with the community makes it evident that Whole Foods has only been benefiting from creating shared value. Not only are they able to outsource production and fund local markets and farmers, but this also decreases their own costs of production, which becomes a win-win situation for Whole Foods, local suppliers, and consumers as well. Creating shared value is an extremely important business value or strategy to have. Rather than completely focusing on being philanthropic with social responsibility, which often comes at the cost of economic loss, creating shared value encourages businesses to do good for society which will end up benefiting the business as well. There is a disadvantage for a profitable company, even when its CEO speaks or thinks of his or her actions as purely socially responsible. This idea of being socially responsible is misleading and perpetuates the illusion that pure social responsibility is possible. On the other hand, there is an advantage to thinking in terms of shared value creation. The more the number of corporation executives thinking in terms of creating shared value increases, the more they may discover profitable and socially beneficial business ideas that they would not otherwise have considered, which combines interests to be socially responsible while still maximizing profits.

Shared value creation allows everyone, including local suppliers, producers, shareholders, workers, and consumers who are involved with corporations such as Whole Foods, to benefit. The shared value created by these firms moves in a circular movement, and everything contributed to society comes back to these firms,

---

<sup>27</sup> Friedman, Milton. “A Friedman Doctrine - The Social Responsibility Of Business Is to Increase Its Profits.” *The New York Times*, The New York Times, 13 Sept. 1970, [www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html](http://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html).

<sup>28</sup> Creating Shared Value: A Strategic Imperative for Companies. (2020, September 17). Retrieved September 13, 2020, from <https://www.csis.org/analysis/creating-shared-value-strategic-imperative-companies>

<sup>29</sup> Mark Kramer Co-Founder and Managing Director, Kramer, M., & Director, C. (2018, July 17). CSR vs. CSV - What's the difference? Retrieved September 16, 2020, from <https://www.fsg.org/blog/csr-vs-csv-what%E2%80%99s-difference>

strengthening a firm's place in society, but also the way a firm interacts with all aspects of society. Social responsibility, on the other hand, is more ineffective than creating shared value because there is never any way to determine whether everyone shares the same values and responsibilities. Being socially responsible and active can also stir controversy in the community. For example, many professional athletes, in light of the Black Lives Matter movement, were very open in conveying their stance on these issues. Because these professional athletes have such a large platform, and their words are received by such a wide audience, the stances these athletes have did not align with some public opinions. Because these athletes conveyed their contentious claims, which were perceived to be controversial, viewership of national league sports that are heavily broadcasted and viewed in America declined by almost 30%.<sup>30</sup> This "responsibility" that these athletes had to society was a way of bettering society, but it had a direct effect on the viewership of their sports, showing a dissonance in the social responsibility of individuals and the way the general public receives their socially responsible actions.

It is also difficult to identify whether every person in a firm, such as the stakeholders and shareholders, are truly happy with the socially responsible cause a business executive has decided to donate to. Lastly, and most importantly, social responsibility should be enforced with the genuine motivation and intent of executives to really help society, whether it benefits an executive's firm or not, rather than being motivated by the desire to increase profits and build a reputation for being charitable.

When an executive claims to have a social responsibility, it is extremely difficult to ascertain whether an executive is truly being socially responsible to be socially responsible; they could be "socially responsible" with the intention of making a profit out of their philanthropic work in the future. Because of the uncertainty for both consumers and producers that comes with being socially responsible, the whole idea itself is flawed because there is always a degree of dishonesty and discontent. Shared value creation, on the other hand, leaves all aspects of society content, because every person involved in this interaction between consumers, producers, suppliers, and third parties is benefitting in some way or another. Each party involved has certainty that they will receive some value, either social or economic, from shared value creation, whether it be in the short-run or long-run.

## **Conclusion**

The question of whether a firm can balance corporate social responsibility while also creating economic profits is a complicated one. The table below summarizes the differences between a traditional business model, which Friedman endorses and compares this to a socially responsible business model, which is what Mackey believed was integral to a firm's success.

---

<sup>30</sup> Saad, L. (2020, September 14). Farming Rises, Sports Tumbles in U.S. Industry Ratings. Retrieved September 16, 2020, from <https://news.gallup.com/poll/319256/farming-rises-sports-tumbles-industry-ratings.aspx>.

**Table 1.**

<b>Traditional Business - Friedman</b>	<b>Socially Responsible Business - Mackey</b>
Focuses on creating only economic profit - separates productivity from personal values	Economic profits are secondary, and benefitting society by fulfilling personal values is more important
Does not believe in giving back to society through donations and charity – by continuing to produce and manufacture goods, a firm is already giving to society	Believes that giving back to society is even more important than generating economic profit - improving societal welfare through philanthropic actions is most valuable
People within the firms have a social responsibility, but the firm itself does not. The firm is constructed and shaped by the employees' values and ideologies	The firm is an entity that is socially responsible. The firm is built on the personal values of the employees; therefore, it is the firm's responsibility to adopt and represent these employees' personal values - this becomes their social responsibility.

The table above briefly compares the two business models described by both thinkers. As Milton Friedman would argue, social responsibility is an artificial idea created by firms to make it seem like they are very philanthropic. A firm itself cannot have any social responsibility, but the people within a firm can, based on their personal values and goals. This “do-good” approach leads to consumers perceiving an organization as more charitable which increases profits. Friedman says that the idea of a firm trying to profit from philanthropy invalidates any social responsibility consumers have. Mackey, on the other hand, criticizes Friedman’s approach as being too narrow of a description of the social responsibility of corporations. Mackey argues that the social responsibility of a company is a compromise between stakeholders that combines the interests and values of all people within a company. Mackey says that though social responsibility comes at an economic loss, the purpose of social responsibility is not to make a profit; it is to help consumers. But Mackey also said that being socially responsible is good for a firm because it brings in more consumers in the long run, meaning more profits. This then raises the question of whether there is a balance between social responsibility and economic profits, or whether the intention to make economic profits is masked with the idea of social responsibility. Because of Mackey’s claim that social responsibility serves as a profit-making marketing tactic in the long run, the integrity and meaning of social responsibility is tarnished. Social responsibility is an extremely murky idea because there is no way to determine true intentions behind being socially responsible; if the intention is to make a profit then it is not social responsibility. Social responsibility is also an extremely risky practice for businesses, because even if a business is truly socially responsible and does not expect to make a profit from its social responsibility, it is still an economic loss for them. Instead of being socially responsible, creating a shared value between producers, suppliers, and the community is far more beneficial to all involved parties. Corporations that create economic value in a way that also creates value for society by addressing its needs and challenges create a shared value. This shared value is far more valuable for corporations because it allows for corporations and societies to combine shared interests. For companies, this shared interest means being as profitable as possible, while being able to do good for society. But doing good for society by the idea of shared value creation does not mean sacrificing the firm’s interests or economic profits; it means making the community happy while also

making firms happy. This shared creation value leads to a more productive and happier society. Though social responsibility seems appealing in theory, it is almost impossible to determine the legitimacy of socially responsible actions, intent, and benefits. Because social responsibility comes with much uncertainty for society and for firms, pursuing social responsibility seems pointless when firms could create shared value creation instead.

The question of whether corporations can balance social responsibility and create economic profits is therefore answered: firms cannot do both simultaneously. Social responsibility comes at an economic loss for firms, which means that profits are not maximized. If firms pursue social responsibility with the intent of increasing profits, then it is not social responsibility because firms are extrinsically motivated to continue being socially responsible to create profits, corrupting the definition of social responsibility. In other words, pure social responsibility always comes at an economic loss, and intent to create profits from social responsibility is not actually social responsibility. Therefore, corporations cannot balance economic profits with social responsibility and must look to other methods, like creating shared value, in order to create an equitable and profitable balance between the two corporate goals.

## References

- Banton, Caroline. "Shareholder vs. Stakeholder: What's the Difference?" *Investopedia*, Investopedia, 28 Aug. 2020, [www.investopedia.com/ask/answers/08/difference-between-a-shareholder-and-a-stakeholder.asp](http://www.investopedia.com/ask/answers/08/difference-between-a-shareholder-and-a-stakeholder.asp).
- "Corporate Responsibility: Factsheets." *CIPD*, [www.cipd.co.uk/knowledge/strategy/corporate-responsibility/factsheet](http://www.cipd.co.uk/knowledge/strategy/corporate-responsibility/factsheet).
- "Creating Shared Value." *Harvard Business Review*, 22 Oct. 2020, [hbr.org/2011/01/the-big-idea-creating-shared-value](http://hbr.org/2011/01/the-big-idea-creating-shared-value).
- "Creating Shared Value: A Strategic Imperative for Companies." *Creating Shared Value: A Strategic Imperative for Companies | Center for Strategic and International Studies*, 30 Dec. 2021, [www.csis.org/analysis/creating-shared-value-strategic-imperative-companies](http://www.csis.org/analysis/creating-shared-value-strategic-imperative-companies).
- Friedman, Milton, et al. "Rethinking the Social Responsibility of Business." *Reason.com*, Reason, 21 Aug. 2020, [reason.com/2005/10/01/rethinking-the-social-responsi-2/](http://reason.com/2005/10/01/rethinking-the-social-responsi-2/).
- Friedman, Milton. "A Friedman Doctrine-- The Social Responsibility Of Business Is to Increase Its Profits." *The New York Times*, The New York Times, 13 Sept. 1970, [www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html](http://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html).
- Ganti, Akhilesh. "Social Responsibility." *Investopedia*, Investopedia, 22 Dec. 2020, [www.investopedia.com/terms/s/socialresponsibility.asp](http://www.investopedia.com/terms/s/socialresponsibility.asp).
- Greenwood, Chelsea. "10 Of the Companies That Give the Most to Charity in the US." *Business Insider*, Business Insider, 14 Nov. 2018, [www.businessinsider.com/companies-that-give-the-most-to-charity-in-the-us-2018-9](http://www.businessinsider.com/companies-that-give-the-most-to-charity-in-the-us-2018-9).
- "Impact." *Whole Kids Foundation*, [www.wholekidsfoundation.org/impact](http://www.wholekidsfoundation.org/impact).
- Mark Kramer Co-Founder and Managing Director, et al. "CSR vs. CSV - What's the Difference?" *FSG*, 17 July 2018, [www.fsg.org/blog/csr-vs-csv-what%E2%80%99s-difference](http://www.fsg.org/blog/csr-vs-csv-what%E2%80%99s-difference).
- "Our Approach." *Whole Cities Foundation*, 2 Oct. 2020, [www.wholecitiesfoundation.org/what-we-do/our-approach](http://www.wholecitiesfoundation.org/what-we-do/our-approach).
- "Our Impact." *Whole Planet Foundation*, 16 Nov. 2020, [wholeplanetfoundation.org/our-impact/](http://wholeplanetfoundation.org/our-impact/).

- Saad, Lydia. "Farming Rises, Sports Tumbles in U.S. Industry Ratings." *Gallup.com*, Gallup, 14 Jan. 2021, [news.gallup.com/poll/319256/farming-rises-sports-tumbles-industry-ratings.aspx](https://news.gallup.com/poll/319256/farming-rises-sports-tumbles-industry-ratings.aspx).
- Staff, Investopedia. "Milton Friedman Definition." *Investopedia*, Investopedia, 25 Jan. 2021, [www.investopedia.com/terms/m/milton-friedman.asp](https://www.investopedia.com/terms/m/milton-friedman.asp).
- "United States Securities and Exchange Commission." *Document*, [www.sec.gov/Archives/edgar/data/865436/000086543616000366/wfm10k2016.htm](https://www.sec.gov/Archives/edgar/data/865436/000086543616000366/wfm10k2016.htm).



# Bioethics: To What Extent Should Humanity Perceive Eugenics as Applicable?

---

*Bangjun Yue*

**Author Background:** *Bangjun Yue grew up in China and currently attends Institut Le Rosey, in Rolle, Switzerland. Her Pioneer research concentration was in the field of philosophy and titled “Bioethics.”*

## Abstract

Through this paper, I will analyze the case of growth hormones involving positive and negative eugenics, consider the stances various authors take in regard to those cases, and confront the question: to what extent should eugenics be perceived applicable to humanity? Initially, I will embrace the topic of eugenics by introducing the concept and types of eugenics. Then, my aim is to discuss the implications of eugenics in relation to humanity by looking at its historical applications, finding common ground in which features of eugenics are perceived as being inapplicable due to its unethical or unrealistic nature. Following this, I will study multiple authors’ opinions on positive and negative eugenics; I acknowledge the complexity in distinguishing between positive and negative eugenics in certain cases. Thus, after engaging in the case study, I aim to present solutions to the distinction between positive and negative eugenics. Consequently, I will analyze the pros and cons of eugenics and conclude that humanity should perceive eugenics as applicable to the extent that it is within the capability of technology and the boundary of ethics. However, regulations regarding eugenics still ought to be explored and established. Finally, I will consider the significance of eugenics to humanity.

## 1. Introduction

### 1.1 Eugenics

Eugenics – literally meaning “good birth” (Goering) – typically refers to the manipulation of human genes to improve future generations by identifying desirable characteristics to inherit and omitting undesirable gene traits. The term *Eugenics* was coined in 1883 by Sir Francis Galton: influenced by Charles Darwin’s theory of natural selection (P. Wilson), Galton founded meliorative movements of science and social



implications to promote the study of human improvement across generations, and translations of the study into social policy; the underlying concept is that humans can actively step into the process of reproduction by choosing whom to reproduce, which, in turn, enables us to influence inheritable family traits – both desirable and undesirable – within the next generation (R. Wilson).

Eugenics is typically understood as a concept injected with negative historical significance, primarily due to Germany's Nazi regime in World War II: inhumane exterminations, breeding programs, and medical experiments. However, intuitively, eugenics is loaded with positive intentions to promote good births, improving the lives and conditions of future generations by equipping them with good genes, all of which are the natural aims of parents to provide the best possible environment for their children. This leads to the discussion of the moral obligation to ensure our future generations' well-being in face of the problems that arise. Given that eugenics, with its intended optimistic impact, could fulfill parents' moral obligations for the security of their future children's well-being, to what extent should such obligations be met (Goering)? More specifically, to what extent – if morally permissible – should humanity perceive eugenics as applicable, especially when considering, according to Goering, whether or not contemporary practices are capable of being distinguished from past eugenic practices – be it in relation to shifts in their goals, modes, justifications, and even consequences?

## 1.2 Distinguishing Kinds of Eugenics

In their attempts to distinguish contemporary practices of eugenics from their related historical parts, philosophers began exploring the different kinds of eugenics present and past – one of them being “liberal,” which contrasts with “authoritative” eugenics (Agar). The differences lie in the fact that liberal eugenics, as its name suggests, is dependent on individual free choice, open to pluralist values, and based on conversant scientific understanding; the advocates of this nature aim to be sensitive towards the impact of deeply-rooted social problems, such as racism, sexism, and heterosexism upon individual choice. On the other hand, authoritative eugenics relies on coercive state programs based on heritability assumptions to promote social goods (Goering). Liberal eugenicists substantially help renew our understanding of contemporary eugenics by contributing efforts to distinguish the practice from its morally unacceptable predecessors. Such shifts in the concept of eugenics can be seen among many published papers, with one example being “Eugenics and the New Genetics in Britain: Examining Contemporary Professionals' Accounts” by Kerr et al. The article encompasses accounts of an impressive group of professional British scientists and clinicians, and summarizes the five rhetorical boundaries they drew between the new eugenics, or liberal eugenics, and the old eugenics: old eugenics was described as being unscientific, unfeasible, coercive (totalitarian regimes), abusive, and involved altering the gene pool of the population. Today, the emphasis placed upon individual rights has allowed modern genetics to distinguish itself from its predecessors – individual or family decisions take precedence instead of the government-dictated, historical widespread screening associated with authoritative eugenics. The shifted focus of new eugenics points to hereditary and environmental influences on the disease rather than on heredity and behavior, documenting the swing of the nature-nurture pendulum (Kerr et al.).

Understanding and practices of eugenics have undergone continuous development in recent decades following the establishment of the Human Genome

Project, advancements in prenatal and genetic testing, the creation of the National Human Genome Research Institute (NHGRI), and more, not only enabling more precise exploration of eugenics but also enlarging possible ways in which human can intervene in the reproduction process. Arising from this developing field are problematic boundaries and controversial research into how far we should carry out the potential of eugenics (Goering).

### **1.3 Positive vs. Negative Eugenics**

Eugenics can be categorized into two types: positive and negative. Positive eugenics aims to improve future generations' conditions by encouraging the reproduction of those with the best traits and capabilities so that heredity plays its role in enhancing the next generations' health and performance. Negative eugenics, on the contrary, aims to improve by eliminating genetic "defects," such as genetic diseases, disabilities, and disorders, and discourages reproduction amongst those with the least appealing and least capable traits (Buchanan 104). Although Buchanan et al. depict a distinct line that separates positive eugenics from negative eugenics, they point out that real medical cases challenge such clear distinctions and paint out the grey areas between the two, on which their views and those of others will be summarized and analyzed in later sections.

## **2. Brief History of Eugenics**

Eugenics today is commonly perceived as something that dates from the late 19<sup>th</sup> century, but in fact, the idea of selecting more appealing mating partners in order to have offspring with more desirable traits dates back to ancient times. In Plato's *Republic*, for example, a society in which selective breeding takes place to improve the conditions of the human population is depicted (P. Wilson). The modern eugenics was initiated by a cousin of Charles Darwin, British scientist Francis Galton, in the 19<sup>th</sup> century. Galton was intrigued by how scientific intervention into mating selection could "improve human stock" and pursued his goal of "creating better humans," which gained acceptance among scientists and policymakers in the early 20<sup>th</sup> century in the United States and Britain (Goering). In the United States, positive eugenics programs created certain fairs in which families competed for prize the quality of being "fitter," much like the way livestock is judged on its quality at fairs (Stern). On the other hand, negative eugenics programs promoted the sterilization of males and females regarded as "unfit" to reproduce, often referring to people who are poor and suffering from mental illnesses. This is due to the fact that many eugenicists held the assumption that inheritable traits include social and behavioral conditions, which will be inherited by children from their parents, rather than the idea that these conditions occur commonly throughout society (Goering).

Following the forced sterilizations in the United States, Adolf Hitler took a step further with his horrific eugenics experimentations during WWII. Hitler referred to the eugenics practices in America in his book, *Mein Kampf* (1934) and created the concept of the Aryan race, a race that is superior with traits such as being athletic, blond, and blue-eyed. In the book, Hitler identifies races such as Jews and Roma as inferior to the Aryan race, and promotes extreme measures - even genocide - to

ensure that the German gene pool is as pure and as free from inferior races as possible (History.com Editors). Another devastating implication of eugenics lies in the concentration camps under the hands of Josef Mengele, who carried out horrific and unscientific experiments on adults and child twins in the name of assisting Hitler in creating the superior race.

## **2.1 Historical Implications to Modern Eugenics**

One entrenched negative connotation of the concept of eugenics is rooted in the Nazi regime of World War II. The term “eugenics” was nearly equated with the inhumane and terrifying “eugenics” programs of Nazi Germany. Following the termination of WWII, eugenics organizations were careful to rename themselves, replacing the term “eugenics” and moderating their aims. For instance, the American Eugenics Society was renamed to the Society for the Study of Social Biology (Goering). In this way, historical implications redirected the aims and practices of modern eugenics, especially raising sensitivity towards coercive selective breeding, as well as the impact of past eugenics upon racism, sexism, and more.

The obligation to remember the painful history also enables eugenics to reform along with advancements in medical technology. Modern eugenics, or human genetic engineering, manipulates genes to alleviate and eliminate diseases, and harbors great significance in its potential to cure lethal, inheritable illnesses. However, problems arise when considering undesirable traits, as it is a subjective and controversial matter upon which people’s perspectives differ. Some people argue that, disease-free or not, all humans have the right to be born as nature intended, and that this should not be actively intervened in by human force (History.com Editor).

The open interpretation of negative traits again calls into consideration the careful delineation of inner boundaries between positive and negative eugenics, especially with their seemingly distinct goals of treating and enhancing, which will be discussed in section 3 below.

## **3. Case Study**

Recall that the distinction between positive and negative eugenics is that, whereas the former aims to enhance the population, the latter aims to treat diseases. Such a distinction can be summarized according to the enhancement/treatment distinction (Buchanan et al. 108). Buchanan and his team point out that this distinction requires more careful analysis for two reasons: first, it must be considered whether or not the treatment/enhancement distinction provides insight into the boundary between permissible and impermissible medical practices; second, the fact that the treatment/enhancement distinction is difficult to draw and may seem arbitrary in certain cases.

### **3.1 The Case of Growth Hormone Treatment**

Johnny is a short, 11-year-old boy with documented GH deficiency resulting from a brain tumor. His parents are of average height. His predicted adult height without GH treatment is approximately 160 cm (5 feet 3 inches).

Billy is a short, 11-year-old boy with normal GH secretion according to current testing methods. However, his parents are short, and he has a predicted adult height of 160 cm (5 feet 3 inches). (Allen and Fost)

According to Buchanan et al., this case may seem “morally arbitrary” when considering the disadvantages caused by disease and impairment (Johnny) and those caused by the disadvantageous allocation of conditions and characteristics by nature (Billy), which is regarded as normal. One reason that the enhancement/treatment distinction, in this case, seems arbitrary, is that both boys suffer equal disadvantage, regardless of the differing causes of their shortness. Both suffer social disadvantage equally as “heightism,” a common prejudice underlying job markets, undermines their competitiveness due to their deficiency in height. Another reason is that, despite the different causes of their shortness – with one being dysfunction and the other being the genotype that is regarded as normal – neither boy deserves to be disadvantaged, as neither is responsible for their shortness, which is instead the result of natural probability that is free from human intervention. In addition to this, the reason behind both cases’ “needs” for growth hormone treatment is rooted in the social prejudice of “heightism,” which depicts them as disadvantaged (115).

Buchanan et al. effectively analyze the nature of the two cases and clearly convey the grey area between the enhancement/treatment distinction. In both cases, regardless of our judgment on disease-required treatment or normal-traits enhancement, Johnny and Billy suffer from the same disadvantage. Although they seem divergent in origin, they are, in fact, identical in nature: both boys suffer from shortness to the same extent, which hints at the same extent of social discrimination due to the “heightism” prejudice that the boys are going to endure.

From this perspective, I view the enhancement/treatment distinction as being limited to social context. The definition of enhancement and treatment in this distinction indicates that our society maintains a precise definition of deficiency and disease, which leads to the moral dilemma distinguishing between the enhancement/treatment distinction of positive and negative eugenics. In contemporary society, it has become a common phenomenon that deficiency is treated comparably to disease, as a condition that has to be treated, be it in relation to height, perception of beauty, or body shape. We see people requesting facial surgery because their appearance may not satisfy society’s common perception of beauty— which for certain societies, beauty is defined as having big eyes, light-skin, and so on – again creating social disadvantages, such as “heightism.” This leads me to question: since when has beauty had a definite standard, and since when has not meeting the standard constituted a “deficiency?” I put “deficiency” in quotation marks as, initially, I viewed having an unappealing appearance as a disadvantage, but then realized it might as well fit into the common definition of deficiency, just like diseases.

In today’s society, appealing appearances denote greater advantage: “good-looking” job seekers get employed much more easily than those who are not as attractive; handsome and beautiful people attract more friends and are more popular; attractive faces even win more favors from their surroundings. Because of this, those with normal or even modest looks suffer relatively as much disadvantage as those who are considered “good-looking” are able to enjoy their advantage. This could be regarded as social inequality in our society due to resource allocation, but the problem is that such allocation is wholly determined by nature. If we take a closer look at such inequality, grey areas arise between the definitions of deficiency and disease. We aim to treat diseases because people suffer from pain or the disadvantage resulting from

that particular disease. By the same logic, an inherited feature that creates pain or disadvantage relative to other people can be regarded as a disease. Thus, deficiencies could be regarded as diseases because they create disadvantages for a minority within the majority. It is hard to distinguish between deficiency and disease when both impose disadvantages on individuals. But certainly, we do not equate disease with deficiency. Diseases are independent negative conditions that are treated more seriously, whereas the deficiency of an individual is a disadvantage dependent on other people – it is not independent, but rather relative to its surroundings. In the case above, disease is a natural prenatal defect, whereas deficiency is regarded naturally as being normal, but relatively defective postnatally.

If the distinction between disease and deficiency is drawn clearly in this way, the second problem is whether or not such a distinction applies to Johnny and Billy's case, as treatment and enhancement. Johnny's case involves treatment with the application of negative eugenics as he is suffering from a natural prenatal defect, which we categorize as a disease that needs treating for the sake of distinguishing Johnny and Billy's cases. However, is Billy's case considered an enhancement because he is regarded as normal without any prenatal defect? Billy suffers from a deficiency relative to the average height of the general population, but he is normal because there is no defect in his genes. Hence, would not treating him with a growth hormone to improve his height be considered as an enhancement – taking a step further from being normal?

The distinction is problematic here as, although according to the testing methods Billy is normal without any prenatal defect - thus harboring no "disease" - he is considered abnormal within the society into which he was born. He may be normal medically, but not socially. In fact, he does not match the standard of being normal in society – normal men tend to be above 170 cm. Enhancement makes the normal better, but how should we consider if Billy is normal or not, since there is a certain social standard of normal that seems to conflict with that of medical practices? Normal is defined by the average in the society, and positive eugenics, in broad terms, enhances the average to make individuals better in various aspects. So, since Billy is treated with a growth hormone to become average height, he may be comparable with those harboring a disease treated to become normal. Because of this, his case is treatment with negative eugenics. Yet, intuitively, this distinction does not seem to match his condition precisely. He is short because of his parents' height instead of any natural defect, so improving his height at this standard would be enhancing his normal height, breaking the natural rule of being normal. So, how should treatment and enhancement be distinguished between in this case? Defining a standard of what is normal is the key.

### **3.2 The Delineation Between Positive and Negative Eugenics in Their Distinction**

Contemporarily, it is commonly viewed that the distinction between positive eugenics and negative eugenics delineates a moral boundary for medical practices (Buchanan 106). Many people presume that negative eugenics is morally permissible, and as such genetic intervention aims to alleviate and cure pain and diseases. On the contrary, positive eugenics practices are morally unappealing, leading to complex ethical considerations (Anderson; Kitchoer). This highlights the significance of drawing clear

distinctions between positive and negative eugenics and the treatment/enhancement distinction.

We see that to solve the distinction between Johnny and Billy's cases, a precise standard for normal is required. "Normality" acts as the stanchion for the balance, with treatment on the one hand and enhancement on the other. What being normal really helps to distinguish is the difference between disease and deficiency. Medical obligations further establish this distinction; a prime example is the case of the "Plain Hero" compared with the case of someone who suffers from a congenital deformity (Daniels; Macgregor; de Beaufort et al.). A "Plain Hero" is a person of average appearance who desires to be better looking. Buchanan argues that, for the case of the "Plain Hero," cosmetic surgery is not a medical obligation but an individual satisfaction of preferences, since the Plain Hero is normal looking but is unsatisfied with his mismatch with the mainstream perception of handsomeness. Whereas for the deformity case, plastic surgery is a medical need (113). The real question is: what determines if medical practices applied to a scenario are considered treatment or enhancement?

In this case, the decisive factor seems to be the root cause for the cases needing treatment, but we have already discussed and analyzed Buchanan's point on how both Johnny and Billy are not personally responsible for their conditions; they both are dominated by natural allocation and selection. Therefore, the reason behind our concentration on the origin of the need for treatment is no more than trying to establish whether the cause meets the criteria for being "normal," or within our standard that is between disease and deficiency.

In essence, the point of the enhancement/treatment distinction in positive and negative eugenics is, in fact, a matter of how far we expect our generation to go in the future, and how far away we expect to be distant from our past. That is, being normal means that the individual meets the current standard of being average and does not suffer from congenital diseases inherited from the past generations; neither is the individual developed to reach an above-average standard. The point of negative eugenics is to apply medical practices on individuals to be distant and free from inherited diseases, whereas, for positive eugenics, the goal is to exceed the "normal," directing our future in the way we favor. To be without fault is normal; being average is normal as well. We already see how normal is not definitively defined, as it varies depending upon the context. Thus, under medical circumstances, if there is no flaw in the genes, regardless of whether the function of the gene is better or worse relative to others, it is normal. Yet, in the social context, if the gene functions worse or better than the average, the individual is considered abnormal.

Therefore, unless we reach a definitive standard of being normal that can and will be accepted in all contexts, the treatment/enhancement distinction, as Buchanan pointed out, is arbitrary. An alternative solution would be to apply medical practices according to one of the standards of being normal within many contexts. For instance, when drawing the treatment/enhancement distinction between positive and negative eugenics, one only applies medical practices to determine whether it is a case of positive or negative eugenics. If the gene is unharmed or defective in any way, regardless of the quality of the function of the gene, it is normal. Thus, under these circumstances, Billy's case should not be treated because he is genetically normal, and applying further growth hormone treatment would be considered an enhancement that exceeds the normal standard.

Despite the two presented solutions, problems remain, with one of them being that the standard of "normality" constantly changes. Recall that in the GH cases above, Billy's condition is his "normal GH secretion according to current testing

methods.” That is, his GH secretion was considered normal at the time when he was tested using the testing methods at that time. This hints that his normal GH secretion may not be considered normal at another time, since people may have derived new knowledge in studying GH secretion. Does this mean that medical practices applied to Billy’s case can be considered both positive and negative at different times? If so, there will be no exact moral boundary for the medical application of eugenics, since it is constantly changing. What is morally permissible becomes morally impermissible at another time. One may argue that it is only worth evaluating the “normality” standard at the current time. However, it is significant to explore and understand how this fluctuation in “normality” standards shifts the distinction between positive and negative eugenics, which in turn gives us a general understanding of the nature of our moral values and ethics. This problem will be further discussed in later sections, as it is of great concern to the general view on eugenics.

### **3.3 Conclusion to the Case of Growth Hormone Treatment: Enhancement vs. Treatment**

After analyzing the case of growth hormone treatment to clearly distinguish between enhancement and treatment, or the distinction between positive and negative eugenics, the following conclusion can be reached: although the enhancement/treatment distinction seems in itself clear, grey areas remain when distinguishing between enhancement and treatment. In the case of growth hormone treatment, this paper focused on the core of the distinction between treatment and enhancement – disease and deficiency - and claims that the blurry boundary between enhancement and treatment underlies the commonalities between disease and deficiency, especially when considering both within a social context. When considering how both lead to identical disadvantages for Johnny and Billy presented in the case above, it is complicated to determine if medical treatment for both cases genuinely differs, as both will be alleviated from the same disadvantageous situation from the treatment. However, when considering both cases within a medical context, it may seem that the distinction between treatment/enhancement is obvious. For the sake of precise comparison, disease is defined in this context as an independent negative condition that is treated more seriously. In contrast, the deficiency of an individual is a disadvantage dependent on other people – it is not independent, but rather relative to its surroundings. In the cases above, disease is a natural, prenatal defect, whereas deficiency is regarded naturally as normal, but relatively defective postnatally. Thus, in a purely medical context, Johnny’s case would be considered treatment of negative eugenics because it is categorized as a disease according to the definition above, while Billy’s case would be considered as enhancement of positive eugenics because he is regarded as normal from a medical perspective.

Despite the different determinants of the treatment/enhancement distinction under differing circumstances, what must be taken into account for all contexts is the extent to which our current standard is considered an enhancement, and the extent to which the opposing direction is considered treatment. Thus, a definitive standard of “normality” is needed. In essence, it is crucial to understand what defines normal, or the middle ground between being enhanced and harboring a disease. Only by acknowledging what is normal can we be capable of developing a deeper understanding of what counts as a treatment for a disease or the enhancement

of a normal individual. As discussed, the standards of being normal also vary and may even conflict with each other if considered within the context of different scenarios: Billy is normal and requires enhancement within a purely medical context, whereas he is abnormal in his short stature, which comparably is regarded as a disease in society, and arguably meets a medical need. In other words, Billy's case could be considered either way depending on what defines "normal": since Billy is treated with growth hormone to become average height, he may be comparable with those suffering from a disease but receiving treatment to become normal; or, Billy is naturally short, having inherited it from his parents, and he does not harbor any prenatal disease that caused his shortness. Thus, treating him with growth hormone would be enhancing his normal state, departing from what nature granted him. Considering Billy's genes medically, he is normal as there is no defect in his gene, regardless of the poor "performance" in building his height. However, in a social context, no one considers how genes function. Rather, society takes into account the actual "outcome" of these genes – for example, height. Billy is considered abnormal, and thus, improving his height would be comparable to treating a disease in order to achieve "normality."

To summarize the measures that clarify the treatment/enhancement distinction, it is either developing a definitive standard of being normal that can be followed and agreed upon in all contexts, or applying medical practices according solely to one definition of normal within a specific context – that is, to define Billy's case as positive eugenics because, from a medical perspective, he harbors no disease and is absolutely normal. However, it is also of importance for us to be aware that defining "normal" is an intuitively subjective matter and is constantly changing, especially as our society becomes increasingly pluralistic.

#### **4. Should we endorse Eugenics for the Sake of Human Civilization and Evolution?**

As pointed out when discussing the GH case, clarifying distinctions between positive and negative eugenics aims at drawing the moral boundary between what should be considered morally permissible and impermissible. While the idea of negative eugenics has been embraced by many for its function of alleviating people's pain, for many others, positive eugenics has been labeled a threat to humanity, partly because it is unknown whether moral considerations hinder further exploration. The following sections will encompass essential factors to be considered for eugenics as a whole, and consider the varying opinions of professionals specializing in different areas. Further opinions and reasoning will also be presented, and eventually, a conclusion will be drawn with an overall summary. Note that although the sections address positive eugenics specifically, they also refer to eugenics as a whole since the distinction between positive and negative eugenics has not been firmly established, as discussed above.

##### **4.1 Standards**

Considering positive eugenics and its goal of enhancement brings the discussion back to the GH case. As argued, to draw a clear line between positive and negative eugenics, a definitive standard for normality is required, which enables us to refer to it and to determine which medical practices count as a treatment for a disease and which count as the enhancement of normal individuals. This discussion focuses on the controversy



over the application of positive eugenics. That is, what changes are considered as enhancement? For instance, positive eugenics could be applied to enhance intelligence. But what exactly is our standard for being more intelligent than others? Perhaps, decades earlier, highly intelligent people were those skilled at arithmetic, responding to problems faster, and adapting better to new environments. However, the more pluralistic our society becomes, the more our understanding of intelligence shifts substantially. We no longer refer to intelligence as one feature, but as one that encompasses many areas without one definitive standard. One may be highly intelligent when it comes to art – or even crafts – but not when it comes to sciences. Because of the different definitions of intelligence, there is no one standard to follow that makes one definitively more intelligent compared to others if positive eugenics is applied. Another typical example would be attractiveness. Of course, in the past, it was commonly accepted that blue eyes, light-colored skin, and blonde hair were the standard for beauty in many countries – influenced by colonialism and the white Eurocentric standard of beauty. But is this still true today? People tend to be beautiful/handsome in their own unique way. Another important consideration is that one may be beautiful in some people's eyes, while being unattractive to others. Subjectivity plays a vital role in thwarting any clear definition of enhancement for an individual. This is problematic as we have no clear understanding if we are enhancing or not. Changing traits that would previously have been considered enhancing may not be regarded as such in the future, as our values and beliefs shift over time. Since positive eugenics is about enhancements, we need to establish firm standards of normality on paper. However, such standards are both unlikely and unfeasible to delineate.

#### **4.2 Responsibility**

Another concern regarding positive eugenics is responsibility. Since the area of positive eugenics is still unknown and requires further exploration, enhancing individuals by changing traits through technology today or in the future may be disastrous – just consider *Frankenstein*. Aside from the dangerous and adverse effects positive eugenics might impose on individuals, the impact it might pose on the gene pool of the whole population is also a risk. Enhancing one individual means enhancing his/her children, which will affect their grandchildren, and eventually exponentially change the gene pool of a generation. If anything were to happen, who is to take responsibility? Or, who is capable of shouldering such significant responsibilities - responsibility not only of taking the blame but also of shouldering the task to fix the situation? If anything were to go wrong with positive eugenics practices, no one would be able to rein in the wild horse, especially when descending recklessly down a very steep “slippery slope” of positive eugenics.

#### **4.3 Playing God's Role**

This corresponds to the discussion about responsibility. No one is capable of shouldering the responsibility of the consequences positive eugenics could potentially bring about, because there is so much uncertainty surrounding positive eugenics. We are not omniscient and wise like God, and we certainly cannot fix the situation if it goes wrong. Specifically, we do not have God's knowledge of the whole and how we affect the whole by making one small change; we have neither God's knowledge of the future nor God's power to deal with the consequences. Consider the “God” mentioned above as nature and the application of positive eugenics as the role that nature plays

in designing the traits and the conditions with which we are born. Humans derive from nature, and a common belief is that since nature is good, we are good. Thus, many argue that positive eugenics is neither positive nor good as it messes with nature's work.

#### **4.4 Social Implications**

If positive eugenics is put into practice, who will obtain the most advantage? Undoubtedly the wealthy and powerful, thus increasing inequality. Eugenics could even potentially be exploited by one group over another, similar to racial discrimination. Assuming that positive eugenics works well and safely, and that future generations become more competitive and are genuinely brighter than past generations, people will want further enhancement. Enhancing a child's traits will not be cheap and accessible to the whole population, considering the advanced technology and practices. Eventually, the rich will have the most enhanced children, and the poor will fall way behind. Economically speaking, the rich will become richer and the poor poorer, due to their gap in capabilities given this unequal access to this technology. Those more advantaged will have more power over those less advantaged due to their overpowering strength as a result of positive eugenics. Another concern is that people will lose their diversity and become more alike. If positive eugenics enhances an individual, it must possess a standard of enhancement, which every individual will want to follow to become so-called enhanced. If this is the case, then people will lose their uniqueness and develop similarly. Positive eugenics raises many concerns, especially in society, as it directly affects our population in the ways we live and behave. Perhaps our heterogeneous society will become homogenous.

#### **4.5 Impact on Humanity**

Another question for positive eugenics concerning humanity is whether it leaves a space for us to breed humanity. For as long as humans have existed, we have thrived and continued to live through the natural process of reproduction. If we can modify the traits of our babies through positive eugenics, what is the purpose of mating and reproducing? During a conversation about the impact of eugenics on humanity, Professor Sheldon of Northwestern University points to the significance of parental love for their children and the meaning of family. Specifically, he states that a mother loves a child because it is her child, regardless of magnificent qualities or serious deficiencies. How would parents feel about genetically enhancing their children? Can the newborns still be considered their reproduction if they are not natural? Given that family is one institution in which people are valued because of the way they are, while in the society people are constantly graded and evaluated based on their "qualities", how would eugenics reshape our ways of valuing family if we are capable of altering traits/qualities of our family members? There also exist abundant essential issues to consider about humanity besides Sheldon's point on the value of family. If newborns can be genetically enhanced, how would we define humanity in that case? Would humanity mean the best possible being we are able to create using technology? Would there even be the concept of evolution? How far would positive eugenics enable us to go? Where are its limits? These are all essential matters to consider before we can apply positive eugenics to reality.

## 5. Addressing Arguments against Eugenics

Some professionals are strongly against eugenics due to the devastating consequences it might bring to humankind. For instance, Brosius and Kreitman argue that eugenics is “evolutionary nonsense” for the following reasons:

The process of evolution requires that genes mutate, individuals are selected, and populations evolve. Without errors in the replication of genomes and the production of genetic variation among individuals, populations would not be able to respond to a new environmental setting. Evolution does not involve an active drive toward progress.

Just as evolution has no foresight, neither do *Homo sapiens* (although our highly developed central nervous system often deceives us into believing so). For our species, environmental and social conditions drastically change, even over the course of a few generations. Thus, the eradication of traits that are considered undesirable today would be a vain, if not ill-fated, attempt to ‘improve’ the human race.

Their arguments in these two paragraphs could be summarized as a response to new challenges and adaptation, and a continued lack of any standard measure for improvement. Brosius and Kreitman argue that, in changing the traits of the population, eugenics would undermine the diversity of the population, therefore exposing the population to new challenges at a higher risk and increasing the difficulty to adapt as people tend to have similar traits due to eugenic practices. However, a counterargument would be that undermining the diversity in the genetic traits of the population would, in fact, protect nearly the whole population from a particularly devastating disease or sickness. Coronavirus is a prime example. No one is immune to this virus, and if the whole human population were to be free from the danger of COVID-19, then each individual would need to have the same change in gene traits so that the virus would not affect anyone. However disadvantageous that undermining the diversity in gene traits seems, there would nevertheless be some advantages. This is a case of trade-offs, determining the priority of the population rather than a sole argument against eugenics.

As for adapting to new challenges, eugenics could also be continuously applied. Since we have applied eugenics to alter the gene pool of the population, we could alter the gene pool again when we approach new challenges. Some may argue that our eugenics practices are not advanced enough to deal with new challenges. However, consider how we have come up with medicines for unknown diseases and saved lives – eugenics could work in the same way.

For their second argument, I would argue that the purpose of eugenics is not to improve but to develop. Brosius and Kreitman argue that any attempt to improve our race would be in vain because there is no standard measure for improvement, given how rapidly our conditions change. However, if we apply the same logic Brosius and Kreitman employed for eugenics to this case, we might argue that our entire civilization has hardly improved because conditions constantly change, and what might have been regarded as an enhancement – for example, increased average height compared to the past – would be in vain for our development. Applying eugenics to develop desirable traits for the population works similarly to

enhancement in technology and arts and all areas: they help us to more efficiently and effectively address the current problems we encounter.

Eugenics or not, the knowledge of the future remains non-existent. Thus, it is more significant that we focus on the present and understand how applying eugenics will enable us to tackle current dilemmas and adapt more efficiently and effectively. As long as we are moving in a specific direction, we are making progress and developing. There is nothing done in vain for improving gene traits in the population. Therefore, eugenics should be taken into consideration on top of the two central arguments that Brosius and Kreitman made against eugenics. Yet, we must also not forget the potential hazards of applying eugenics to real life as it is still an area in need of further exploration. Carrying out experiments on human bodies is not ideal because we cannot afford the price of any consequences.

## **6. Unforeseeability: The Major Problem of Eugenics Application**

Historically, the thought of applying Eugenics as a techofix approach to our existing problems has resulted in the natural bias of merely fixating on the overstated positive consequences. As such, the negative consequences are generally neglected due to the fact that they are less identifiable and unforeseen - perhaps unforeseeable. Humans have invested in many similar cases: climate change as a result of using fossil fuels, overpopulation and climate change as a by-product of Rx improvement, and further starvation and wealth concentration to the minority as an eventual consequence of the Green Revolution. These past, profound examples should act as a warning for us to rethink the unknown consequences of eugenics on humanity. In any case, the introduction of eugenics into human society is comparable to the launch of a nuclear weapon. Humanity has witnessed the unpredictable horror of the Nazi's, and even present day the genocidal tendencies of white supremacists. Another consequence is proposals of genocide as a solution to national problems, practiced already in the U.S., Armenia, and other contemporary countries. Eugenics as a tool has all played its part in these.

In general, we should exercise caution in evaluating the degradation of human dignity by technological change, and reassess the ways in which we view nature, other humans, and ourselves, especially with the ongoing exploitation of humans as standing reserves to fulfill the desires of the few at the expense of the majority while creating an unimaginably inhumane and unsustainable world, as warned against by Heidegger and environmental phenomenologists. The unforeseeable unknowns of eugenics are truly beyond any human estimation and pose potential threats to humanity.

## **7. Conclusion: Best is Yet to Come**

Perhaps the most significant concern for eugenics is that it is a "slippery slope," and we cannot risk our whole population falling down because one "fly" in the pot could potentially ruin the whole "soup." Many are thwarted by the thought that eugenics is unknown; it is scary and threatening. However, the final point I am going to argue is that eugenics may seem unknown at first, but it could, in fact, make us more certain of our future.

I have already discussed that we neither have God's omniscient knowledge of how we are affecting the whole by making one small change, nor do we have God's knowledge of the future. Again, I will replace God with nature, with nature being the authoritative force and law governing our being. Before the concept of eugenics, we never considered the possibility of directing our future, instead obeying how things occurred naturally. Thus, many diseases were inherited and still occur today. However, when we consider eugenics, although we are uncertain of the consequences it will lead to, we can at least control a part of our future. The future we prayed for before eugenics that is free from suffering or disease can now become a reality. In a way, applying eugenics creates knowledge about the future that we will never gain if we ignorantly follow things naturally. There is also no guarantee that nature will not unintentionally eliminate us one day. In a sense, investing in eugenics is no more dangerous than fully trusting the natural law. We have adversely affected our environment and have seen some "retaliations," the most powerful and up-to-date example being coronavirus, of which we are unable to locate the exact source. Every individual within our civilization is at risk of dying from the disease.

Mathematically speaking, the future is infinity to us, and whatever we do to infinity, it is still infinity. By this logic, if we apply eugenics, with the consequence of a plus or minus, multiplication, or division to our future, our future will remain infinity. It seems that eugenics is in vain, as mentioned by Brosius and Kreitman, because we do not know what the future will be like. Broadly speaking, eugenics will not determine our future. However, we gain one essential advantage from eugenics: the solution to current problems, specifically to disease. At least, with the application of eugenics, we are sure that diseases, such as inherited obesity, will no longer be present in future generations as we eliminate them from our gene traits. This is the only thing that we are certain of in our future, and it is the most effective change we can make to our population. This certainty is enough, and whatever consequences that derive from it will not be a concern of ours at present, as the future is always unknown. It may be that, even without altering the gene pool, the same consequence will take place after a few generations. Thus, the most impactful and beneficial contribution we can have on our civilization is eugenics, as we are capable of directly addressing the sufferings and unsatisfactory aspects of our being and accomplish our goal, supported by advanced technology and technique.

As for the responsibility of shouldering the consequences, no one is or will be able to shoulder such responsibility because no one is responsible for the consequences. Instead, people are fulfilling their medical obligations in alleviating and eventually eliminating the sufferings of disease. Perhaps it is the terminology of eugenics and its historical connotations that have presented eugenics as a threat to humanity. However, if I denote eugenics as DNA for the sake of removing negative connotations, it would be much more obvious that eugenics is a natural feature that we constantly pursue and apply to reality. We change our genetic features all the time. For instance, continuously sleeping late alters our DNA. Also, viewing from past to present, we are always pursuing the best lives for, or the best versions of, ourselves, be that in the selection of one's other half, trying to be more intelligent through various mental exercises, and so on. The marriage between the higher ranks in the past was to maintain blood purity and breed the best next generation possible with the "good" genes of royal parents. We still see this today: when people look for their other half, they always tend to choose more good-looking ones, smarter ones, and taller ones. We are also eager to become better, be that in terms of intelligence, height, or beauty. These phenomena indicate that we have an intuitive desire for eugenics or

improvement amongst our next generations. The above are phenomena of eugenics, all of which are common, and deeply rooted in our lives.

The practices of so-called eugenics have no goals divergent from those of the above phenomena. The only difference is the employment of more advanced techniques and technology that helps target our aims more accurately and accomplish them more effectively and efficiently. This is no more than the relationship between a pencil and a pencil sharpener. Our intuitive desire to enhance is the pencil; the pencil sharpener sharpens the pencil, just as eugenics helps us crystallize our goals and accomplish it more easily.

Of course, advocating for the application of eugenics does not mean that eugenics is all positive without any drawbacks or limitations. One potential hazard is the existence of a standard measure for desirable traits. With the current evaluation of enhancement, the application of eugenics will modify genes of different individuals to harbor a similar trait. The loss of diversity in genes is very likely to take place as supported by Brosius and Kreitman, and the example of the “herd effect” (a whole group of genetically engineered cows wiped out by a particular disease as their identically engineered genes are at higher risk of dying from that specific disease) with its potential impact on human population must be considered carefully and seriously. This is where regulations and evaluations can be implemented to decide the extent to which eugenics should be applied. This regulatory component is one area that requires more discussion and exploration.

Another concern of eugenics is the enhancement of individuals. Our population will lose diversity not only in gene traits, but also in individuality, or characteristics that render us different from each other. If people are enhanced, they are being enhanced according to a certain standard at that time, which indicates that people will be enhanced in the same direction. We might just become closer to being identical and lose the pluralist features of our communities. We should be able to delineate how we identify and define ourselves as individuals, which is constantly changing in our societies, before moving on to considering enhancement. Would we like to see ourselves becoming homogeneous by being enhanced in the same way?

Last but not least, eugenics is merely a matter of how far we expect our generation to go in the future and how far away we expect to be from our past. It could be considered solely as a compass that indicates the direction in which we favor, thus directing our civilization. Eugenics is applicable to the extent that such a choice will be within our control, but not until we have explored to a sufficient level the ethical boundaries within eugenics, and have established measures to follow in order that eugenics is applicable to the reality that we want to achieve.

## Works Cited

- Agar, Nicholas. *Liberal Eugenics: in Defence of Human Enhancement*. Blackwell Publishing, 2004.
- Allen, David B., and Norman C. Fost. “Growth Hormone Therapy for Short Stature: Panacea or Pandora’s Box?” *The Journal of Pediatrics*, vol. 117, no. 1, 1990, pp. 16–21., doi:10.1016/s0022-3476(05)82438-2.
- Anderson, W. French. “Genetics and Human Malleability.” *The Hastings Center Report*, vol. 20, no. 1, 1990, p. 21., doi:10.2307/3562969.

- Beaufort, Inez de., et al. In the Eye of the Beholder Ethics and Medical Change of Appearance. *Scandinavian University Press*, 1997.
- Brosius, Jürgen, and Martin Kreitman. "Eugenics—Evolutionary Nonsense?" *Nature Genetics*, vol. 25, no. 3, 2000, pp. 253–253., doi:10.1038/76994.
- Buchanan, Allen E., et al. "Positive and Negative Genetic Interventions." From Chance to Choice: Genetics and Justice, by Allen E. Buchanan, *Cambridge University Press*, Cambridge, Cambridgeshire, 2000, pp. 104–155.
- Daniels, Norman. Just Health Care. *Cambridge Univ. Press*, 2001.
- Goering, Sara. "Eugenics." *Stanford Encyclopedia of Philosophy*, Stanford University, 2 July 2014, plato.stanford.edu/entries/eugenics/.
- History.com Editors. "Eugenics." *History.com*, A&E Television Networks, 15 Nov. 2017, www.history.com/topics/germany/eugenics.
- Kerr, Anne, et al. "Eugenics and the New Genetics in Britain: Examining Contemporary Professionals' Accounts." *Science, Technology, & Human Values*, vol. 23, no. 2, 1998, pp. 175–198., doi:10.1177/016224399802300202.
- Kitcher, Philip. The Lives to Come. *Simon & Schuster*, 1998.
- Macgregor, Frances Cooke. After Plastic Surgery: Adaptation and Adjustment. *Praeger*, 1979.
- Sheldon, Mark. "Eugenics and Humanity." 9 Sept. 2020, <https://philosophy.northwestern.edu/people/emeriti-faculty/sheldon-mark.html>
- Stern, Alexandra Minna. "Making Better Babies: Public Health and Race Betterment in Indiana, 1920–1935." *American Journal of Public Health*, vol. 92, no. 5, 2002, pp. 742–752., doi:10.2105/ajph.92.5.742.
- Wilson, Philip K. "Eugenics." *Encyclopædia Britannica*, Encyclopædia Britannica, Inc., 19 Feb. 2019, www.britannica.com/science/eugenics-genetics.
- Wilson, Robert A. "Eugenics and Philosophy." *Oxford Bibliographies*, 26 July 2017, www.oxfordbibliographies.com/view/document/obo-9780195396577/obo-9780195396577-0350.xml.



# What is Wolf: Narratives about Human-Wildlife Conflict in Tibetan China

---

*Chuhan Zhang*

**Author Background:** *Chuhan Zhang grew up in China and currently attends Nanjing Jinling High School International Department, in Nanjing, China. Her Pioneer research concentration was in the field of anthropology and titled “Communication and Culture.”*

## 1. Introduction to Research

“You’ve never fully understood the ‘wolf culture’ of Huawei” was the title of a news article published on the Sina website on July 17, 2020. The article analyzed the growth of Huawei from a startup to a prominent tech firm in the past decade, and argued that the “wolf spirit” embedded in corporate management and developmental strategies is a key reason for Huawei’s success. Here, the symbol of the wolf represents teamwork, tenacity, and even a degree of fierceness, as appropriate. A similar search about “wolf culture” in the Baidu search engine yields over ten million results.

It is striking to observe the prevalence of “wolf culture” in the register of pop culture, since most Chinese people grow up listening to fairy tales about wolves’ evilness. The classic plot would be about wolves trying to kill their prey at first, but always being tricked by some wiser animals in the end. When the image of the wolf abruptly transitioned from a despicable villain to an object of worship, it not only left the question “why,” but also created confusion about the true nature of the relationship between wolves and humans. There is thus an urgent need to hear from people in close contact with wolves for a long enough period, and Tibet became the center of my research.

My preliminary interest in wolves emerged from random chats with Tibetans during a two-week anthropological field trip. When asked about their daily lives or major concerns, Tibetans almost always answer something about wolves. Especially for pastoralists, encounters with wolves have been an essential component in their nomadic lifestyle. The origin of wolf and human conflict is easy to comprehend: wolves prey on livestock that pastoralists raise, and since livestock is the primary source of income for pastoralists, wolves’ predation threatens people’s livelihood. According to Tibetans, wolves kill dozens of sheep every year per household, though the number varies depending on pastoralists’ vigilance against attacks.

The conflict between wolf and man stood as an objective fact for the people in Tibet, yet people’s responses to and interpretation of such conflict differ, thereby forming distinctive patterns of narratives. Thus in this paper, I wish to compare and contrast Tibetan narratives with the established ones in Chinese literature. I will first attempt to distinguish the said narratives in traditional tales, Chinese eco-literature, Buddhism converts, and pastoralists; I will then discuss how certain images of the wolf are being socially constructed as people narrate. In the following analysis, I seek to explore the reasons behind such a construction, namely, how people identify with the conflict situation and align themselves with nature. Finally, I intend to analyze



the significance of this construction with a refutation of the ontological perspective. The complex “webs of significance” that people have spun about wolves challenge the existing stereotype of wolves that is shaped by contemporary literature, media, and marketing strategies.

## 2. The Wolf as a Literary Image

Section two probes into the literary construct of wolves through a brief overview of the development of animal narratives in China. Section 2.1 is devoted to analyzing the similar characteristics of wolves among traditional animal stories. Here, “traditional” is referred to in a broad sense, covering several tales commonly heard in Han Chinese and Tibetan culture, as well as a sketch of the animal narratives in the context of political turbulence, shortly before the introduction of ecological criticism. In section 2.2, I will then introduce how the image of the wolf is sharply reversed in light of ecological literature since the 1980s, with reference to several case studies.

### 2.1 The Wolf in Traditional Tales

Though the truth of the matter cannot be verified, documentation of the interactions between humans and wolves emerged as early as the Spring and Autumn Period in ancient China. One story is about a kind-hearted man Dong Guo who lent a helping hand to a wounded wolf he met. When the wolf was cured, however, it did not hesitate to kill its savior. The tale has been endowed with allegorical meaning after centuries of oral inheritance. Nowadays, countless children are warned by this story at an early age, not to let down their guard easily, or be fooled by what appears on the surface. This story, just like many others commonly told in Chinese and Tibetan societies, includes the wolf as the central villain in the plots. When interviewing a group of school children at Ganjia, they showed us several picture books about wolves with Tibetan and Chinese characters in them. One of them, *Pony in the Mud*, described how a pony trapped in the mud tricked the hungry wolf to pull it out of the pond and cleaned its body. When the wolf finally prepared to eat it, the pony then used its hoof to kick it away. In another story, a clever lamb stalled the hungry wolf inside his house by asking it to do all kinds of housework. Eventually, the mother sheep returned and scared the wolf away.

It suffices to understand the characteristics of wolves in folktales through these examples alone, since they can hardly escape from the “fierce, greedy, and foolish” stereotypes. Particularly in the texts for younger audiences, the image of the wolf scarcely varies, so that most children eventually learn to associate these characters with wolves in an instant reflex. To them, the wolf is not so much a living animal with flesh and blood as it is a signpost for wrongdoings and negative personalities. To parents, on the other hand, the image of the wolf serves as an educational tool to help children discern right from wrong, thus developing a mental scheme of people’s consensual perception of the world.

This applies not only to children, however. Until the ecological turn in the development of Chinese literature, the majority of animal narratives for adults followed a similar pattern, where educational intent lurks behind the interactions between mankind and wildlife. The 1950s and 60s was an era in which the Chinese literature society unanimously endorsed political correctness over individual expression. In face of rapid political changes, collectivism, socialist ideology, and revolution became the hot themes echoed repeatedly in writings (Tang, 2010). Under

such circumstances, the image of wolves again became the symbol of evilness, but with a new layer of moral and political significance. For example, in Chinese writer Qin Mu's prose *On Wolves*, he analogized wolves' greediness, slyness, and cowardice to the nature of imperialism, who he claimed to be operating under the "rules of the jungle" (Mu, 1978).

In essence, before the awakening of ecological literature, Chinese writers still regarded nature as the wilderness, to be conquered or as a set of resources to be exploited, always subject to the domination of men. Accordingly, the wolf, as a component of nature, was also subordinate to men's willpower whose true nature is not worthwhile to learn or write about.

## 2.2 The Wolf in Chinese Ecological Literature

The fervent political reformation in China in the middle of the 20<sup>th</sup> century was accompanied by burgeoning industrial development as the government attempted to catch up with the Western world. However, expedited progress came at a cost. Just like the developed world, Chinese people soon started to suffer from the environmental damages caused by their unlimited exploitation: air pollution, deforestation, global warming, etc. As people grew apprehensive of their living conditions, they were forced to acknowledge in the meantime that nature is capable of destruction when it has been pushed to the limits, and that it was naive for human beings to disrespect the environment in which they were living. Consequently, a new school of literature—the "ecological literature"—gradually formulated in the 1980s where writers generally disapproved of industrialization and emphasized the value of environmental conservation. The wolf, as a component of nature, was thus repositioned at the center of the animal narrative, and men humbly retreated.

In sharp contrast, the once despicable wolves were suddenly extolled as the embodiment of virtues that men had lost in the progress of civilization. In Jia Pingwa's novel *In Memory of Wolves*, for instance, the narrator is an ardent environmentalist who tries to protect wolves in Shang Zhou. However, the villagers resent him because wolves pose threats to their livelihood. Eventually, all the wolves are eliminated, but people's lives turn out to be more miserable because the balance of the ecosystem is broken permanently (Jia, 2020). Another famous novel, *Wolf Totem*, includes a series of accounts of an urban youth new to the Mongolia grassland. There, he witnesses how real wolves live, prey, and coexist with men. To the author, the wolf represents freedom, team spirit, bravery, tenacity, and the lost wilderness (Jiang, 2004). A similar book, *Returning to the Wolves*, is a piece of documentary writing about how a young female artist cured an injured wolf cub in the city and brought it back to the plateau. To her, the wolf is like a child that strengthens her bond with Mother Nature, and the act of releasing it instead of captivating it showed the author's respect for freedom and liberation of mind (Li, 2018).

It is not an exaggeration to conclude that the wolf is the "noble savage" in these literary works. Some even assert that the wild nature of wolves is in many ways superior to human society, so that if men seek to rectify their societal flaws and illnesses, they should learn from the wolves. The boldness of such a claim is true for some and ridiculous to others, which exactly illustrates the polarized and conflicting views people have about wolves: traditional tales present the wolf as an evil villain, and eco-literature idealizes it as a role model. Yet, both of the narratives are animated pictures that dwell in the imagination of authors, instead of authentic images shaped by daily encounters. In order to break the myths in the literary perception of wolves, it is necessary to analyze the Tibetan narratives through an anthropological lens

### 3. Images of Wolf in Tibet

#### 3.1 Introduction to Tibetan fieldwork

From July 19 to July 29 of 2020, I followed my teacher to conduct field research in the Amdo Tibetan Region, which stretches across the land of Qinghai, Gansu, and Sichuan Province. Amdo is among the three traditional districts of the “Tibetan Area” (see Figure 1), divided according to differences in dialect. Instead of identifying with the administrative division implemented by the Chinese government, most Tibetans use Amdo, Kham, or Ü-Tsang to introduce where they belong.

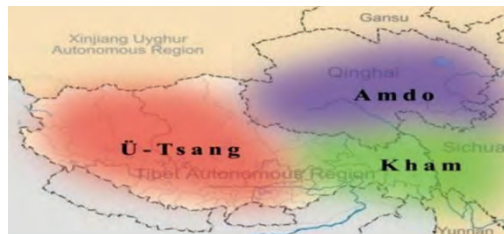


Figure 1

During my research, I visited four environmental protection organizations located in Ganjia, Maqu, Zhaqiongchang, and Nianbaoyuze, all of which are the most influential ones in their respective communities. Many of these agencies were primarily responsible for rubbish collecting and sorting, a problem made worse by tourism. However, activists soon found that waste production is an endless problem, therefore many resorted to more effective and educational approaches: the team in Ganjia began teaching local pupils traditional Tibetan craft, also filming and archiving the condition of surrounding wildlife; Nyansog (short for Nianbaoyuze Environmental Protection Agency in Tibetan) started publishing biodiversity manuals, as well as collecting plant specimen for their records; and Zhaqiongchang’s founder Damchoe dedicated his life to controlling grassland desertification.

#### Research Methodology

I, along with two other students, cooperated in interviewing members of these organizations, as well as other willing participants. My informants can be roughly divided into four social categories: monks, environmental activists, pastoralists, and school children, although these identities often overlap for different persons. With the informants’ consent, we either recorded the entire interview or typed up notes. According to our records, the total length of the interviews amounts to no less than 12 hours, the content of which forms the major source of data in this research.

#### Basic Information on Wolves

The Amdo Region is mostly plateau with an average altitude of 3000-4000 meters, and the wolf is arguably the most commonly encountered animal besides yak and sheep. The Tibetan wolf is the subspecies of grey wolf, which can be widely observed

in Middle Asia, Mongolia, Northern China, etc. It has medium body length, with long and tawny hair. Wolves hunt alone or in small herds, feeding on small animals and herbivores, such as pika, yak, sheep, hares, and animal remains (Gao et. al, 1987).

According to one of my informants Metok, a pastoralist, baby wolves live in caves under the protection of their parents, and grown-up wolves inhabit the mountain. In the afternoon when the temperature is high, they will go down the hillside to search for water (see Figure 2 for the landscape of the wolf's habitat). Unless starving for food, most wolves won't start hunting until nightfall. The pastoralists we interviewed in Nyansog also observed that wolves' attacks occurred mainly during the colder seasons, because many wild animals on the mountain actively reproduce in the summer but hibernate in winter. Therefore, the abundance of livestock kept in captivity suited their survival needs. Our driver Hu, an experienced adventurer on the Qinghai-Tibet Plateau, told us how wolves eat their prey: when successfully breaking into a herd of sheep inside the fences, wolves will typically fill their stomachs first by tearing sheep apart for their internal organs (to prevent the animals from rotting), then drag them to some secret burial spot as winter supplies.



Figure 2

### Conflicts and Dilemmas

Conflict between wolves and humans is aggravated by the fact that the distance between these two species is shrinking, which can be attributed to several factors. The most obvious issue is the overgrazing and population explosion problem in China. Over the past decades, people have been continuing to intrude into wildlife's original habitat, making it easier for wolves to trace the herds.

Political factors are also involved. After the Tibetan Liberation in 1951, many Tibetans held arms in their homes, either for self-defense or hunting. However, in the 1980s and 1990s, the Chinese government decided that gun holding fostered societal turbulence, so guns were gradually confiscated from each household. Hu recalled that he used to store a gun in his van, but then the municipality suddenly announced that all guns have to be registered and licensed before legal usage, so he handed in the gun, and never got it back. Tibetans gradually lost their guns in this way. Public security did improve, but people also lost their lethal weapon against wolves. Pastoralists' power was weakened further as stricter environmental regulation laws were passed. For example, in 1988, it was stipulated in Clause Six of the Wildlife Protection Law of the People's Republic of China that "illegal hunting of wild animals or damaging their habitat is strictly prohibited" (National People's Congress, 1988). Wolves eventually

recognized men's restraint in power, and grasped the opportunity to launch vigorous attacks on livestock.

When policymakers realized that the above regulations seemed to favor wolves' interest over men's, they did search for remedies to rebuild the balance of the ecosystem. The same wildlife protection law of 1988 also states that the country shall compensate for losses caused by wildlife under special state or local protection. However, according to Li and Gao (2019), such compensations are problematic in several respects, including tedious and redundant application procedures, insecure sources of funding, and moral dilemmas as people exaggerate or even fabricate animals' death for money. One of our informants, Gyatso, a prestigious monk in Ganjia with the title of Gexi (the equivalent of doctoral degree in the Gelug School), also revealed the plight of pastoralists:

Question:

“When wolves attack sheep and yak frequently, how does the government respond?”

Gyatso:

“I'd say there is compensation...but all these complicated procedures before filing a complaint is simply too difficult for pastoralists. They are illiterate! Some of them couldn't even settle their identity issue with the government. For example, my name is supposed to be Sangji, but the government officials thought it's Sangjie and wrote a wrong name on my ID card. The problem just remains, can't be solved! These [wrong name and livestock's death] aren't pastoralists' fault, but many times it's they who bear the consequences.”

Gyatso's student Dhondup added another reason:

“Well then, the government couldn't tell if I killed two sheep or the wolves ate them, right? But actually, many pastoralists don't bother to do so, because a sheep sells for more than 1,500 yuan, but the equivalent compensation is less than 150.”

The shortened distance between men and wolves, combined with flaws in wildlife protection regulations, make it increasingly difficult to resolve the conflict between these two species. Wolves thus became an unavoidable existence forcefully incorporated into people's lives, with which people struggle and reconcile constantly in their own narratives.

### **3.2 Wolf as Link in Causality: Buddhism**

In this section, I intend to review Buddhists' narratives of wolves. To seek connections between Buddhist philosophy and believers' ecological values, I designed a series of questions about wolves and Buddhism respectively. Some of the questions yielded intriguing responses. Due to language barriers and limited time, however, I could only develop a grasp of the basic concepts of Buddhist theory, which requires the supplement of secondary documents, as will be discussed.

**“Do you think wolves should be killed?”**

“There has been a litter of wolf cubs living at my winter pasture since the last spring. Nine of them in total. People in the village discovered traces of the wolves and attempted to kill them, so they asked me where they were. I didn’t tell them, because I pitied the wolves, I thought I should protect them. In the first year the cubs even played with me at the pasture. But the next year they ended up eating many of my yaks.”

The man who told this story to us was Lhawang. He used to be a pastoralist, but now works as an environmental activist at Nyansog, offering help on plant specimen collection. He laughed happily when recounting this story, almost as if it were the story of someone else. When asked “Do you regret keeping the wolves now that you’ve lost so many yaks?” He replied with a shake of his head, “No. Wolf cubs will grow up, it’s something foreseeable the day I chose to protect them.”

Lhawang’s case was extreme, but similar patterns of thinking were not uncommon. The other two activists/pastoralists that were interviewed together with him held the same opinion, that wolves should not be killed. The formulation of such an opinion owes much to the widespread influence of Buddhism. The three pastoralists didn’t receive an official education, but they learned when working with eminent monks at Nyansog. In fact, local Buddhist temples usually exert stronger influence on people’s way of life than the local governments do, since almost any Tibetan believes in Buddhism to a certain extent.

“Abstinence from killing” is one of the core rules of conduct for Buddhist converts. Buddhism believes that animals and men are part of “sentient beings,” therefore their life value should be respected and protected. This discipline was listed as the top of the Five Fundamental Precepts (Abstinence from killing, robbery, unwise remarks, sexual misconduct, and intoxicants) that regulates believers’ behavior as long as they have faith in Buddhism (Cering, 2007). Believers weren’t willing to hurt even the lowest and simplest form of life: I witnessed several monks rescue small bugs drowning inside their tea cups and gently putting them away.

In contrast with killing, Buddhism endorses harmonious relationships with wildlife. One religious story about Thang Tong Gyalpo that I collected conveyed similar values. Gyalpo was a Living Buddha in the 15<sup>th</sup> century with significant achievements in Esoteric Buddhism. It was said that when Gyalpo traveled to the area of Kongpo, people there pleaded with him to eliminate the disasters brought by tigers. Gyalpo accepted their request, but prayed to Buddha that people and tigers live harmoniously together, wishing the tigers longevity. People were confused, but Gyalpo’s words possessed a transforming power, making the tigers harmless. Men were then able to coexist with tigers peacefully (Jiumi, 2002). Even though the result (or “effects”) of the tale may be an imagined miracle, Tibetans never have doubts in the power of karma: as such, they are always kind to the animals with whom they share the world. In addition to their refusal to kill, Tibetan Buddhism advocates for the proactive protection of animals. For example, pastoralists will free a regular number of livestock and birds per year, putting stamps or marks on one ear to warn other people against hurting them.

To probe the ecological values behind believers’ sympathy and respect for wildlife, we interviewed the president of Nyansog Choedon, who was also identified

as a Living Buddha (widely believed as the reincarnation of a former eminent monk). We retold Lhawang's story to him for his explanation, then he smiled and answered:

Choedon:

"I would like you to think of this question: what is the difference between a wolf and a man? Wolves eat livestock because they are hungry, because they have children to feed. For men, it's the same. We also raise and eat livestock for our own survival and for our family's survival. There is no difference in nature. To a certain level, men are even luckier, because we can eat vegetables or buy groceries from stores; for wolves, there is no second choice besides meat."

("How is this related to Buddhist thinking?")

"Well, in Buddhism, we have the concept of 'being'. We regard every being as equal, so wolf and man and livestock are not that different."

Choedon's reference to "being" and "all beings are equal" entailed the ecological values of Buddhism that decides people's empathetic attitude towards wolves. Buddhism defined "being" in a broad sense: almost everything in the world ranging from creatures to landscape can be considered as beings. Those with a state of life, like men and animals, were "sentient beings"; whereas those without state of life, like plants and landscape, were "insentient beings" (Zhang, 2010). When referring to the creed "all beings are equal," most religious sects agreed that at least all sentient beings are equal in nature and intelligence with equivalent Buddha Nature, which means they all have the potential to transcend into Buddha through spiritual practice (Zhang&Ren, 2003).

In particular, all sentient beings are divided into Six Paths according to the world in which they inhabit: the Realm of Deva, Asura, Human, Animal, Pretas, and Nakaras (Nirayas). The former three paths are generally auspicious, and the latter three paths represent suffering and punishment. However, the six realms are not rigid categories: any soul can transmigrate among them and dwell inside a new body for the next life. The standard of transmigration is "karma", the accumulative impact of deeds that one has done in life (Zhang, 2010). Those who behaved in an evil manner when they are alive will consequently be deprived of the right to become a human, and fall to the ominous path of Animal, Pretas (Hungary Ghost), or Nakaras (Hell). For instance, hunters, a career that most Tibetans belittle, are considered to bear bad karma because killing innocent animals is the utmost wrongdoing one could possibly commit (Tsongkhapa, 1405, pp. 123-129).

As a result, there is no reason for men to view animals with contempt or worship, because in the broader dimension of space and time, they are similar souls stuck inside the endless loop of reincarnation, ranked according to the effects of their own actions. In certain ways, the path of humans is more beneficial to the soul's spiritual journey, because men can understand Buddhist language and practices, thus having greater possibility to transcend into Buddha than their animal counterparts (Zhang, 2009). It is also extremely rare and fortunate to step into the path of men, the possibility of which roughly equates to the chance of a cow finding its yoke when drifting up and down in a vast sea, according to one local account. Therefore, Buddhism teaches its followers that they should cherish their current state of life and

always reserve a compassionate heart for other sentient beings (Cering, 2007). Only in this way will their souls remain purified, furthering the journey of enlightenment.

**“Do you think wolves should be protected?”**

When asked whether wolves should be eliminated on the grassland, almost every interviewee answered no. So I switched to the other extreme: whether wolves should be protected by the government. Surprisingly, the answer was again negative:

(“Wolves are now regarded as animals under special conservation. Do you think the government should protect them?”)

Gyatso:

“I...(paused for a second) I think all beings should be protected. But it’s not worth it if another animal is sacrificed just because this one needs to be protected. The most proper state is let every animal grow and die by themselves in the arms of nature, without human interference. From the perspective of Buddhism, [the soul of] every animal is born into their current body because of previous karmas. They exist in this way for a reason. If we classify the endangered species into first grade, second grade, and third grade protected animals intentionally, then in the course of prioritization there must be some animals whose interests get infringed. If we were to feed wolves and snow leopards so they can live well, many sheep have to die.”

Gyatso’s student Dhondup is an eighteen-year-old Tibetan teenager who works part-time at Ganjia Environmental Protection team. He has just finished Gaokao (the Chinese College Entrance Exam) and graduated from high school. He echoed Gyatso’s opinion, with a rather scientific explanation:

Dhondup:

“The best protection is no protection. Wolf eats sheep, sheep eats grass, and when they die they turn into fertilizers. Grass then grows from it again. This is a healthy circulation that shouldn’t be broken.”

As discussed in 3.1, when government policy inclined towards the wolves, the original balance among men, sheep, and wolves is tilted. Pastoralists naturally feel unsatisfied because the government seemed to prioritize the well-being of animals before that of people. Buddhist thinking also rejects government intervention, but instead of an economical standpoint, it seeks to evaluate the impact of “overprotection” on the entire ecosystem. In Gyatso’s words, the state of being for man and animal was arranged by the circulation of cause and effect. If the government forcefully intervenes, the relationship between creatures and their environment will be interrupted with unforeseeable consequences. The monk Lhayul, leader of Nyansog and also a Kampu (highest recognition of knowledge and reputation), once made an interesting metaphor about nature (Gao, 2019):

Lhayul:



“If the nature is compared to a hotel, then men are no more than transients. Men are entitled to the services offered at the hotel, but have no right to damage its facilities and resources. That’s because there will be other guests and more guests in the future living in this hotel. The nature never belongs to human, it is a home that all creatures share.”

Both Lhayul and Gyatso disapproved of the involvement of an external hand, because it breaks the balance of the ecosystem and stirs up the conflict between men and wolves. The metaphysical explanation for the monks’ ecological ethics could again be traced back to Buddhist philosophy. The theory of “Yi” and “Zheng” provided an answer to men’s relationship with nature: “Yi” represents the living environment, and “Zheng” represents the living subject (Fang, 1986). The tenet holds that the subject and surrounding environment depend on each other and restrain each other. Therefore, men and the ecosystem shouldn’t be divided as opponents, but viewed as an inseparable whole.

Moreover, the Theory of Dependent Origination (“Pratityasamutpady” in Sanskrit), one of the founding theories of Buddhism, also discussed the balance and interdependence of things. “Dependent Origin” roughly means that existence is the compounded result of multiple conditions and causality links (Wei, 2009). Accordingly, everything is mutually conditional and interdependent, and the relationship between any two subject formed was as delicate as two pieces of straws, neither one could stand straight without the support of the other. The Huayan School of Buddhism also proposed the metaphor of “Indra’s Net” to visualize the process. According to them, the Indra’s net was a web of countless diamonds, each reflecting a portion of the image of every other diamond. Thus, the splendor of the web is the aggregation of all the diamonds and their reflection (Gautama Buddha, c.400). Similarly, the whole universe is an aggregation of every being and the strings of causality weaving them together. Man is no longer the protagonist in the narrative, and nor is any other being, so what Buddhism invites people to do is simply maintain the balance of the network with as little intervention as possible. Falling into the extremes of either over-killing or overprotection harms the ecosystem in the long run, in the lives that follow.

With belief in the equality of all beings, compassion, reincarnation among the “Six Paths”, and a broader vision of the ecosystem, Buddhist philosophy always instructs believers to think beyond the current space, time, and bodily form in which they exist, and stay humble. That’s why, in Buddhist reasoning, wolves should neither be killed nor be protected—men have no right or power to decide the fate of other beings, causality does. Causality sews all creatures together into a giant, invisible web that surpasses the limit of “this life”, so that although humans and wolves seem like completely different species, there are always certain strings pulling them together. Conflicts for dominance may yield one winner, but the karma accumulated will certainly return at some point after. Converts believe in the system of causality as well as venerating its power, so instead of conflict, they choose to narrate about wolves in an empathetic lens: a soul of the same nature but inhabiting a different body. The image of the wolf is thereby reconstructed and reincorporated into the Buddhist philosophy as “another link in causality”.

### 3.3 Wolf as Rival: Pastoralist

Most believers in Buddhism have a rather moderate attitude about wolves, which was significantly influenced by their religious indoctrination. However, another factor could be that most believers are fairly removed from real life struggles with wolves, because they have to dedicate the majority of their time to learning in local temples. The pastoralists, on the other hand, have a more aggressive stance towards wolves in general, since most of them depend on the livestock they raise for their families' livelihood. Some would even go as far as shooting the wolves in revenge for their attacks, back in the days when they still owned guns. In this section, I intend to focus on interviews with pastoralists in order to create a collage of the image of wolves from their accounts, as well as to learn how people respond to the threat of wolves with their own wisdom.

#### “What are wolves like?”

##### Sly/Clever

Wolf's slyness or cleverness is one of the most heard comments from the pastoralists. Their agility is demonstrated both in their way of hunting and in their knowledge of humans.

Wolves usually don't hunt in the daytime, because they would be easily exposed under the close monitoring of the pastoralists. So most of our interviewees lack knowledge of the exact hunting techniques that wolves employ. However, Damchoe, the founder of Zhaqiongchang, did tell us an interesting anecdote commonly talked about in the tea house.

Damchoe:

“After the guns were confiscated and fences were installed, one day, two little boys saw a wolf coming at their herd, so they wanted to scare it away. They hid themselves and waited for the shot, but surprisingly, the wolf turned around and left. By the time the two kids went back and checked their sheep, it turned out that there had been another wolf lurking behind their herd, and it had already stolen a sheep away.”

Although the truth of the story could never be authenticated, it nevertheless revealed a facet of the wolf's intelligence. The vigilant nature of wolves can be observed more clearly when there isn't a clear and present danger. According to Thubten, owner of a hotpot restaurant and a museum in Jiuzhi town, wolves usually don't return to their homes directly, but instead play a game of bluff:

Thubten:

“A wolf walked towards its cave, but suddenly turned to the other direction and ran away. A while later, it returned, but ran away again...The scenario will repeat three or four times before it eventually let down the guard and went home. The wolf pretends that it has seen people around although it doesn't, that's why it runs away. If there really were pastoralists, they would think that the wolf had recognized them and given up hunting.”

The meticulous precaution that wolves take against humans proves that they are well aware of men's common patterns of behavior. Another pastoralist, Lhami, even told us that wolves can tell people's gender, although most of us couldn't identify theirs, because "wolves ran away when men shout in alarm, but attack when women do the same." Moreover, Lhami told us that wolves can tell if the pastoralist was armed. If not, its behaviors will become significantly more provocative and aggressive.

### **Vindictive**

Wolves are eager to take revenge. When we interviewed Jampa, a pastoralist and activist at Ganjia, he told us that oftentimes, wolves can be full with a sheep or two. However, when men try to kill it and fail, they will trace the same herd until the next year and slaughter it, sometimes killing dozens. Just like an old saying of the Tibetans that basically means "wolves hunt not because its stomach is hungry, but its heart is." Thubten talked about a similar experience: his father once tried to shoot a wolf when he still had a gun, but he missed the target. The same wolf eventually returned, and killed six yaks in a night, which were almost twice the size of itself.

### **Cold-Blooded**

Indifference or loyalty is one of the key differences between wolves and sheepdogs, according to my informants. Many pastoralists raise sheepdogs nearby to their herd, to guard it when men are not present. If men are asleep at night when wolves the herd, dogs will warn its master and even fight with the wolves themselves. Even though wolves and dogs are biologically closer, Tibetans never believe they are relatives in any regard, because dogs recognize men's care and devotion to them, whereas wolves can never get close to men, not even in Lhawang's case.

### **Bold**

Wolves behave more aggressively now that men don't have access to any lethal weapon. It's said that now they even dare to look straight into people's eyes after killing, while decades ago they fled at the sight of a man carrying a gun. Dhondup told us how he was "looked down upon" by wolves in a joking tone:

Dhondup:

"Last time, I saw a group of wolves up on the hill, so I brought firecrackers to scare them away. However, some may have heard the loud sound too often that they know nothing will happen to them. So they didn't even bother to run! They just stared at me like they were looking at a fool. It's so humiliating!"

### **"How do men combat with wolves?"**

As wolves grow more reckless in their attacks, pastoralists have also devised new ways to fight against them (or to cut themselves some slack). Thubten listed five ways to prevent or kill wolves:

Thubten:

“Wolves are afraid of:

1. The smell of burnt clothes: a smell that wolves hate;
2. Donkeys: donkeys resent wolves and chase wolves whenever they see them. Donkeys are also extremely strong and enduring animals, so they run faster but longer, eventually catching up with the wolf then kicking it to death;
3. Traps: some people will place stone or steel traps within a wolf's daily routine. More violent people will just mix poison or explosives with mutton to make meatballs, so that wolves will die or be blown up the moment they chew those;
4. Dogs: a single dog can never beat a wolf, but the power of two will persuade it to abandon the hunt, and five dogs suffice to kill the wolf;
5. Barbed wires: wolves are not afraid of the barbed wire itself, so what pastoralists do is place it along the inner side of the fences, which will trap the wolf as soon as it jumps over the wall.”

However, tempting and effective Thubten's strategies sound right now, some of them will quickly lose their effectiveness, just like many of their predecessors did. Pastoralists have to constantly upgrade and refine their precaution measures because wolves learn them so quickly. Two informants, Metok and Damchoe, introduced the same story of “iteration”:

Metok/Damchoe:

“At first, wolves are afraid of men's voices alone, but later they realized men had been deprived of their guns, and started to attack. Men then put ropes around the herd at night. Wolves were intimidated in the first couple of days, but then found out there wasn't anything threatening. Men then built fences or concrete stone walls, but wolves learned to jump over it. Now, with Wechat and Taobao, pastoralists also live a modern life, you know? Whenever someone found an effective tool on the Internet, like a loudspeaker or flashlight, they broadcast it to everyone they know of on Wechat. You know what? Then every pastoralist used the same tool and wolves learned to circumvent quicker. None of these strategies lasted for long.”

There is, though, one universally applicable, and always effective, prevention technique that all pastoralists agreed to: having a man actively monitoring the livestock. Dhondup said he learned this bitter fact after several losses:

Dhondup:

“I'm a really lazy person, you know? I always get bored when grazing the livestock. Once I bought a loudspeaker and recorded a clip of myself yelling angrily towards the wolves. The speaker could magnify my voice several times, so I thought it had to be useful. One day, wolves didn't come. Three days, wolves didn't come. But on the fifth day, one sheep of mine was killed.”

(“Then how do you graze your sheep and cows now?”)

“Oh, I’ve become cleverer. The wolves are typically active in the afternoon and evening, because that’s when pastoralists become fatigued and less alert. So now I go the opposite way: I run up to the top of the hills in the morning to play with my phone—the signal’s better there. Then in the afternoon I really have to restrain myself and ran downhill, shouting along the way to chase away the wolves.

“Now I’ve realized, once you go loafing, your sheep will be carried away. That’s the rule.” He shrugged and laughed.

In the real life of Tibetans, the image of the wolf does fit some of the stereotypes portrayed in animal stories: greediness, ferocity, barbarity, etc. However, there is one important difference that sets the real struggles of pastoralists apart from the “happy ending” of tales: wolves are clever, so much so that they even outsmart man on many occasions. In the words of Thubten, “The wisdom of one wolf is equal to that of nine hunters”. Although the statement may contain some level of exaggeration, it still holds true that wolves are regarded as the smartest of all animals on the plateau. Their understanding of the human species and capacity to adapt to prevention tactics already distinguish themselves from other predators. Wolves’ existence challenges men’s authority on the pasture, thus forcing them to actively search for better strategies rather than remaining content. The combat between man and wolf is dynamic and enduring, so pastoralists sometimes invest enormous amounts of time and energy into this struggle. It is fair to conclude that the interaction with wolves is an indispensable part of a pastoralist’s life, without which life would be easier, but certainly more mundane.

Wolf constitutes the biggest threat to a pastoralist’s livelihood, yet when interviewed, the perception and feelings these pastoralists expressed towards wolves were surprisingly complex. They do hate wolves for killing their animals, but none of them wished to eradicate the species from the planet if possible. Most settled for scaring it away from their herds with a little punishment. In fact, most pastoralists are already accustomed to the fact that wolves prey by nature, so they wouldn’t even blame them if they steal a sheep for survival during winter. The pastoralists don’t want to eliminate wolves, not because they have the compassion of Buddha, but because they perceive wolves as a worthy rival, and the combat with wolves is an integral part of their pastoralist identity.

The struggle with wolves has become a ritualized game, or a rite of passage that every legitimate herdsman has to experience before becoming a man. When we interviewed several schoolchildren at Ganjia about what they would do if wolves came by, their answers were consistent: “go find daddy.” But for Damchoe, if a man hears the wolves approaching, he would keep his wife and children inside the tent and go out himself. Even if he is so scared that he hides in the bushes for half an hour, he will always return and tell his family that he has bravely chased the wolf away. The transition from fear to “disguised courage” in the face of wolf’s attack indicates how a boy grows up into a man who shoulders responsibility for the livelihood and well-being of his entire family.

Conceptualizing the wolf as a worthy rival also means that men no longer treat wolves with pure resentment, but also with an element of respect. For example, because wolves’ courage and cleverness almost always guarantee success, some Tibetans use wolf-related items to decorate their tent for a good blessing. People also

widely believed that seeing a wolf on the way out signals good luck, especially one with the direction into the opening of the Tibetan robe.

Just like Damchoe's words, "wolves make pastoralists, and pastoralists make wolves," from the macro level, the rivalry between men and wolves is wholesome to their respective evolution and the ecosystem in the long run. The two species restrict each other from unlimited development, so that the population of men, wolves, and relevant livestock are always kept in balance. Another part of his meaning is that to a certain extent, searching for loopholes in the system of order and reaping benefits from it defines wolf culture. Likewise, fighting with wolves as a loyal guardian to the herds defines the pastoralist culture. Only when exposed to constant threats in nature, the spirit and instinct of these two species are trained to improve. In pastoralists' narratives, the wolf is established as a life-long adversary, their combat with which serves as a way of proving their own meaning of existence. Damchoe was nostalgic for the strength of men back in the days with guns, but despite physical power, the intellectual challenge nowadays even increases the dynamism within wolf and man's subtle relationship.

#### 4. Conclusion

In section 2 and 3, I've presented at least four kinds of narratives that distinguish themselves in people's discussions about wolves: those of tales, of Chinese literature, of Buddhism believers, and of pastoralists. These narratives accord with the three categories of attitudes towards nature that most people fall into. In traditional tales, people dominate the narrative with an anthropocentric belief that nature is an object at their disposal. At the dawn of eco-literature, more people began to resonate with the biocentric perspective which urges man to shoulder moral responsibilities for the well-being of the ecosystem. In this particular research, however, I tend to side with the third "social nature" perspective, since the previous two stances divide men and nature in a strict dichotomy, and in effect overlook how different people define nature. People's characterization of wolves is a clear example.

Imagine the scenario where the authors of these distinct categories sit together and start a conversation: when mentioning the subject of wolf, four different sets of images will naturally surface in their heads that decide how they narrate. The sets of images are shaped by people's knowledge and encounters with wolves, as much as they are their own interpretations of it. Their subjective interpretation is closely related to their personality, educational background, life experiences, and so on. Therefore, when talking about wolves, people with similar narratives present a socially and culturally constructed image of wolves; and when interacting with them, they only reaffirm their configuration of the animal.

If put on a spectrum, Buddhism's advocacy for coexistence was an extreme of moderate stance, while hunting or killing wolves stands on the other end of violence. In reality, however, most Tibetans don't vote for pure coexistence or conflict, rather, their attitudes towards wolves are scattered across the whole spectrum. On the one hand, their faith in Buddhism convinces them of the importance of keeping the ecosystem in balance. Buddhism views wolves and humans as equal beings in the loop of causality. To accumulate good karma in their spiritual journey, many believers accordingly refuse to disrupt the rules of nature with their inept, shortsighted will. They are inclined to impose little intervention, in spite of current conflict. For pastoralists, on the other hand, threats to livelihood are among the unavoidable problems brought by wolves. Herdsmen have to engage in an intellectual competition

with wolves to prevent their attacks and give them proper warning at times. Eventually, most Tibetans figured a point of reconciliation between the two stances, with different degrees of tolerance of wolves. However, regardless of whether Tibetans portray wolves as an “another link in causality” or a “worthy rival in close combat,” their narratives about wolves were authentic responses formulated in the context of a nomadic lifestyle. In sharp contrast, wolves in the literary sense are either deemed as the embodiment of evil or hailed as role models of mankind. Arguably, these literary narratives only exist within the speculation of writers who have limited contact with wolves in their lives. The complex nature of the interaction between humans and wolves are seldom, if ever, truthfully presented in literature. Instead, the wolf is merely utilized by writers as an educational tool for the perfection of human society. From a Tibetan perspective, the wolf is viewed as a subject in society that has an equivalent status to man, which is similar to what is termed as “the ontological turn” in the broad trend development of multispecies ethnography.

#### **4.1 Development of Multispecies Ethnography**

Multispecies ethnography reconceptualizes the definition and position of humans in the world. Scholars Kirksey and Helmreich (2010) coined the term to describe an alternative epistemology concerned with nonhuman beings. While traditionally the sponsor, contributor, and recipient of the system of anthropological knowledge have been “standardized humans” (usually Anglo-Saxon males), the supporters of the multispecies perspective attempt to counter such anthropogenic and anthropocentric narratives by inviting plants, animals, and even microbes as protagonists (Ogden et al., 2013). The “ontological turn” stems from this system of belief, defining the transition from multicultural to multi natural theory.

One of the prominent and earliest scholars that champions the ontological theory is Eduardo Vivieros de Castro. He refutes Descartes’ dualism theory that poses nature against culture, contending instead that nature and culture represent a continuum. Moreover, he proposed that there is a universal set of cultural beliefs and practices applicable to different beings, which are manifested through different natures. In other words, beings perceive and behave differently because they are confined to their bodily forms, or different “worlds,” but they should share similar “perspectives” and conform to the same cultural principles (Palecek&Risjord, 2012). In Castro’s “multinaturalism” worldview, animals and men are similar souls occupying different bodies. He introduced his discourse with a famous analogy: jaguars treat as beer what humans treat as blood, meaning that although the signified item is interchangeable, the cultural reference, or signifier, is the same (Saussure, 1916).

Castro was soon regarded as the symbolic figure in Amerindian perspectivism. Following him, many posthumanist anthropologists began to regard plants as “social beings with agentive efficacy”. Rose (2011) described her aboriginal informants’ relationship with animals as intimate “kinship,” so much so that they would cry when their endeared plants were on the verge of extinction. Japanese sociologist Anna Tsing (2015) founded the Matsutake Worlds Research Group and launched their research regarding a kind of mushroom and its implications on the capitalist economy. Following its global supply chains, from foresters at Oregon to upper-class buyers from Japan, she and her team analyzed the interdependent mode of its business circulation with the matsutake mushroom as a companion species. Proceeding from the intertwined relationship between humans and floral life, several scholars delved into the microscopic world. A new wave of awareness stopped

conceiving the human body as a singular entity, but instead as a complex assemblage of countless microbial constituents, or partners. As Rose (2011) and Haraway (2008) sees it, “becoming human is an interspecies collaborative project”. The boundary between nature and culture, or human and nonhuman, is further blurred.

#### **4.2 Historical Studies of Indigenous Hunting Communities**

More or less influenced by the work of Castro, ethnographies of local hunting cultures usually seek to depict a “tilted world” compared to that in Western modernity. In Willerslev’s (2007) ethnography about Siberian Yukaghir hunters, for example, he introduced a way of interpreting nonhuman spirits and humans as “endless mimetic doubles of one another”, so that “Yukaghir hunter, when out hunting, is both hunter and animal, and with the human community he is not simply himself but also a reincarnated deceased relative” (P. 186). Likewise, Paul Nadasdy (2007) documents a reciprocal relationship between Arctic hunting peoples and their prey which they regard as intelligent existence. Sometimes those beings voluntarily give in to the hunters, while sometimes they don’t compromise easily, the latter of which resembles a lot the intellectual combat I’ve observed between pastoralists and wolves. Anthropologist Govidrajan (2015), on the other hand, explores the interspecies kinship in Uttarakhand, central Himalayas through the animal sacrifices. He observes that such kinship is established and reinforced through embodied care and violence. Animals opt to sacrifice themselves to gods through special gestures to their owners as a way to “close the cycle of kinship,” repaying the care they receive and strengthening their family’s bond to deities.

However, it is worth mentioning that ontology remains a hotly debated area in anthropology—the initial “multinaturalism” theory was heavily criticized by anthropologists like Turner, Descola, Anderson for its splitting of nature and culture in a dualistic view. Hence, the school of writers each developed their own theories in refutation, some of which concentrated on observations of nomadic societies. Tim Ingold, for instance, is a British anthropologist who researched across the Pacific Rim, and also a Castro criticizer. In his book *The Perception of the Environment*, Ingold proposed the “ontology of dwelling” though his analysis of the hunter-gatherers’ relationship with animals. He argued that that animals are conceived as “participants in the same world to which the people also belong,” and a successful hunt proved the “amicable relationship” between the hunter and his prey because the animal “has willingly allowed itself to be taken” (Ingold, 2000, p. 69).

Undoubtedly, certain elements of the ontological theory and multispecies perspective parallel with the Tibetan narratives, since they, too, value the wisdom and agency of animals in a given society. It seems that, in the Tibetan society, wolves are also portrayed as subjects to men, and men are willing to use their cultural references to understand wolves. However, it is evident from previous discussions that there exist no “multiple worlds” in Tibetans’ interpretation of their relationship with wolves. Rather, people’s alignment to Buddhist doctrines and rivalry relations shape how wolves are narrated. Furthermore, the limited discourses on hunter-gatherers and pastoralists considered them as distinct social categories of people occupied with different concerns. Yet the rough classification failed to account for the interconnection between these two roles for Tibetans: hunters were engaged in a painful taming process into pastoralists the moment their guns were confiscated, and the unresolved conflict with wolves remains their constant concern as they adapt to new identities.



The Tibetan case's apparent deviance from ontological theory in fact reveals the problematic nature of theoretical debates. Most ontology scholars swiftly and eagerly knit a web of meaning from a number of coincident observations without cross-validating the scenarios in different fields. However, without inquiry into the rich cultural context of a foreign culture, ungrounded speculation only intensifies the involution within the field of anthropology, which sets it further apart from the tangible way of life for the common public.

### References

- Centemeri, L. (2017, August 12). Review de Anna Lowenhaupt Tsing, *The Mushroom at the End of the World. On the Possibility of Life in Capitalist Ruins*, Princeton, Princeton University Press, 2015, pp. 331. Archive Ouverte HAL. <https://hal.archives-ouvertes.fr/hal-01574234>
- Cering. (2007). 藏传佛教慈悲伦理与生态保护 [The compassionate ethics of Tibetan Buddhism and ecological protection]. *Northwestern Ethnic Studies*.
- De Castro, E. V. (1998). Cosmological deixis and Amerindian perspectivism. *Journal of the Royal Anthropological Institute*, 469-488.
- Deqing, J. (2002). 汤东杰布传. [Tang Dong Jie Bu Biography].
- De Saussure, F., & Baskin, W. (2011). *Course in General Linguistics*: Translated by Wade Baskin. Edited by Perry Meisel and Haun Saussy. New York: Columbia University Press. doi:10.7312/saus15726
- Descola, P. (2013) *Beyond Nature and Culture*. Chicago: University of Chicago Press.
- Fang, L. (1987). 佛教哲学 [Philosophy of Buddhism]. Beijing, China: China Renmin University Press.
- Gama, J. (2005). 藏族传统文化中的生态观 [The ecological view in the traditional Tibetan culture]. *Journal of Kangding Teachers College for Nationalities*
- Gao, Y. (1987). 中国动物志 [Chinese Zoology]. China: Science Publishing House.
- Govindrajana, R. (2018). *Animal Intimacies*. Chicago: University of Chicago Press. doi: <https://doi.org/10.7208/9780226560045>
- Haraway, D. J. (2016). *Staying with the trouble: Making kin in the Chthulucene*.
- Ingold, T. (2000). *The Perception of the Environment: Essays on Livelihood, Dwelling and Skill*. Psychology Press.
- Jia, P. (2000). 怀念狼 [In Memory of Wolves]. China: Beijing Union Press.
- Jiang, R. (2004). 狼图腾 [Wolf Totem]. China: Changjiang Literature and Art Publishing House.
- Kirksey, S. E., & Helmreich, S. (2010). The emergence of multispecies ethnography. *Cultural anthropology*, 25(4), 545-576.
- Lin, W. (2007). 佛教“众生”概念及其生态伦理意义 [The concept of "all beings" in Buddhism and its ecological and ethical significance. Academic Research]. China: *Academic Research Journal*.
- Locke, P. (2018). Multispecies ethnography. *The International Encyclopedia of Anthropology*, 1-3.
- Ma, C. (2010). 论苯教和藏传佛教对高原藏族生态文化的影响 [On the influence of Bon

- religion and Tibetan Buddhism on the plateau Tibetan ecological culture]. China: Northwest Normal University.
- Martin P. & Risjord, M. Relativism and the Ontological Turn within Anthropology, 2013. *SAGE Journals*, pos.sagepub.com/content/43/1/3.
- Nadasdy, P. (2007). The Gift in the Animal: The Ontology of Hunting and Human-Animal Sociality. *American Ethnologist* 34, no. 1: 25–43.
- Ogden, L. A., Hall, B., & Tanita, K. (2013). Animals, plants, people, and things: A review of multispecies ethnography. *Environment and Society*, 4(1), 5-24.
- Qing, M. (1961). 艺海拾贝 [Picking Shells from Sea of Art]. China: *Shanghai Literature*
- Rose, D. B. (2014). *Wild Dog Dreaming: Love and Extinction*. Charlottesville: University Press of Virginia.
- Shan, A. (2007). 佛教的生态伦理思想与可持续发展 [Buddhist Ecological Ethics and Sustainable Development]. China: *Qi Lu Academic Journal*.
- Tallbear, K. (2019, January 29). Why Interspecies Thinking Needs Indigenous Standpoints. *Society for Cultural Anthropology*. <https://culanth.org/fieldsights/why-interspecies-thinking-needs-indigenous-standpoints>
- Tang, K. (2010). 中国现当代文学动物叙事研究 [Research on Animal Narrative of Modern and Contemporary Chinese Literature]. China: Nankai University Press.
- Tsing, A. (2015). *The Mushroom at the End of the World: On the Possibility of Life in Capitalist Ruins*. Princeton; Oxford: Princeton University Press. doi:10.2307/j.ctvc77bcc.
- Wei, D. (1999). 佛教的生态观 [The Ecological Views of Buddhism]. China: *Chinese Social Sciences*.
- Wen, H. (2008). 藏传佛教的生态哲学 [The Ecological Philosophy of Tibetan Buddhism]. Tibet, China: *Journal of Tibet Nationalities Institute (Philosophy and Social Sciences Edition)*.
- Willerslev, R. (2007). *Soul Hunters: Hunting, Animism, and Personhood among the Siberian Yukaghirs*. Berkeley: University of California Press.
- Yangjin, Z., & Tu, Y. (2011). 浅谈藏族传统文化与生物多样性保护 [On Tibetan traditional culture and biodiversity protection]. *Tibet Science and Technology*.
- Zhang, H., & Ren, J. (2003). 论中国佛教的生态伦理思想 [On the ecological ethics of Chinese Buddhism]. China: *Journal of Jishou University (Social Sciences Edition)*.
- Zhang, Y. (2010). 论佛教生态伦理的层次结构 [On the hierarchical structure of Buddhist ecological ethics]. *Journal of Southeast University (Philosophy and Social Sciences Edition)*.



# Exploring Neurotrophin Therapy as a Treatment for Rehabilitating Victims of Childhood Abuse

---

*Dojin Choi*

**Author Background:** *Dojin Choi grew up in South Korea, Spain, and three states in the US and currently attends Coppell High School, in Coppell, Texas in the United States. His Pioneer research concentration was in the field of neuroscience and titled “Understanding the Sense of Touch.”*

## **Abstract**

Childhood abuse is a traumatic experience that has severe developmental consequences reaching into adulthood. An unhealthy relationship between parent and child may impair a victim’s ability to cope with stress and act according to social norms and expectations, leading to aggressive and irresponsible behavior that hinders future success and lasting relationships. In addition, abuse exposes children to unhealthy amounts of stress and overwhelms the neuroendocrine systems that regulate stress hormone levels. Excessive hormone levels are toxic to brain tissue, and as a result, child abuse victims have degenerated brain areas compared to healthy individuals, particularly in the hippocampus. To devise a treatment for child abuse, we must first define the psychological and behavioral problems and then identify the physiological issues that can be addressed to reverse them. To that end, recovering the degenerated hippocampus will be the focus of this paper.

In recent times, neurotrophin has emerged as a potent neuroregenerative tool. This proposal seeks to apply the neuroregenerative properties of neurotrophin to address degeneration in the hippocampus caused by childhood abuse. Whether neurotrophin can rehabilitate stress hormone regulation and reduce abnormal behavioral tendencies will also be explored. Abuse will be simulated on infant rhesus macaques (*macaca mulatta*), and the extent to which neurotrophin therapy can reverse the adverse effects will be determined. Based on the results, further studies may extend neurotrophin therapy for use in human rehabilitation.

## Introduction

Working as a psychiatrist in a clinic that treated emotionally disturbed children, John Bowlby observed that infants became intensely distressed when separated from their mothers. In fact, separation from their mothers early in life was a common experience for problem children. This led him to associate a healthy relationship between mothers and the social, emotional, and cognitive development in children, a phenomenon he called *attachment theory*. Attachment theory reasons that children who fail to form attachments or have reason to regret becoming attached in their early years will become naturally aggressive and have difficulty with social interactions because they are afraid of and do not know how to form attachments with others. He defined attachments as “a deep and enduring emotional bond that connects one person to another across time and space” (McLeod, 2017).

According to Bowlby, the behavior of forming parental attachments is an evolutionary trait passed down and amplified because offspring that were able to form attachments with parents had a better chance of survival than offspring left to fend for themselves (McLeod, 2017). Indeed, animals that are dependent on their parents as newborns have innate behaviors that help form attachments, such as how ducklings imprint on the first animal they observe and how human babies smile and cry to attract their parents’ attention. Furthermore, some animals form attachments with individuals other than their parents. Societal animals, like humans, form tribes wherein members depend on one another to survive. Within a society, the ability to form attachments and manage relationships would be an advantage and evolutionarily selected. Considering this, it would be unnatural for children to prefer living without attachments or avoid developing social skills; perhaps children are not antisocial by choice.

Mary Ainsworth, another pioneer of attachment theory, surmised that the first attachment a child forms with his or her primary caregiver during infancy is the most crucial, as it serves as a blueprint for all future relationships. Ainsworth theorized that the personalities of adults were largely shaped by how they were raised and the degree of attachment they had with their caregivers. For example, a baby that was complimented frequently and shown much love would be more trusting of others and the world, thereby being more inclined to form new relationships and explore their abilities. On the other hand, a baby who would be sternly punished when acting against a parent’s expectations would be wary of attempting new things and expressing themselves in fear of being criticized (Odhayani et al., 2013). It appears that childhood is an opportunity to provide children with happy and healthy lifestyles, but conversely, also a period of great vulnerability in which a lack thereof could have serious developmental consequences.

Certainly, a nurturing and trusting relationship with their primary caregiver is pivotal in the development of a healthy personality. From infancy to adolescence, a child’s ability to regulate emotions, form relationships, and develop independence are some of the foundational skills that they acquire through interactions with parental figures. When children experience stress, a trusting relationship allows them to seek comfort and support from their caregivers. This is essential for the development of healthy coping mechanisms to stress, where children learn to place trust in identities other than themselves. Being able to trust and depend on others assures children that the world is a safe place to explore and helps them recover from failures, knowing that there are people who will support them (Odhayani et al., 2013).

The foundational importance of a healthy relationship between a child and his or her parents is the reason the widespread presence of child abuse cannot be ignored. In 2013, it was estimated that nearly 40 million children worldwide are abused each year, significantly affecting all age groups in most cultures and regions (Odhayani et al., 2013). In America, the CDC states that 1 in 7 children experience abuse every year, and that this is likely an underestimate. The World Health Organization defines child abuse as “all forms of physical and/or emotional ill-treatment, sexual abuse, neglect or negligent treatment or commercial or other exploitation, resulting in actual or potential harm to the child’s health, survival, development or dignity in the context of a relationship of responsibility, trust or power.” The Mayo Clinic further specifies five categories of child abuse:

**Physical abuse:** Physical child abuse occurs when a child is purposely physically injured or put at risk of harm by another person.

**Sexual abuse:** Sexual child abuse is any sexual activity with a child, such as fondling, oral-genital contact, intercourse, exploitation or exposure to child pornography.

**Emotional abuse:** Emotional child abuse means injuring a child's self-esteem or emotional well-being. It includes verbal and emotional assault — such as continually belittling or berating a child — as well as isolating, ignoring or rejecting a child.

**Medical abuse:** Medical child abuse occurs when someone gives false information about illness in a child that requires medical attention, putting the child at risk of injury and unnecessary medical care.

**Neglect:** Child neglect is failure to provide adequate food, shelter, affection, supervision, education, or dental or medical care.

Children of abusive or neglectful parents can show signs of *attachment disorder*, where the child develops abnormal patterns of response to human interactions while trying to adapt to their stressful and often harmful attachment to their parents (Odhayani et al., 2013). These abnormal developments may lead to disturbed thought processes and behaviors that are harmful to the child and to those around them.

The negative effects of childhood abuse can impact victims for the rest of their lives. Children growing up in abusive environments where caregivers struggle with mental health or substance abuse problems are more likely to develop learning and behavioral problems compared to children of households without abuse (Odhayani et al., 2013). Child abuse also deteriorates lifelong mental and physical health, causing the underdevelopment of neural pathways for language and higher cognitive functions. These problems directly lead to problems with “emotional regulation, self-concept, social skills, and academic motivation, as well as serious learning and adjustment problems, including academic failure, severe depression, aggressive behavior, peer difficulties, substance abuse, and delinquency later on in life” (Odhayani et al., 2013).

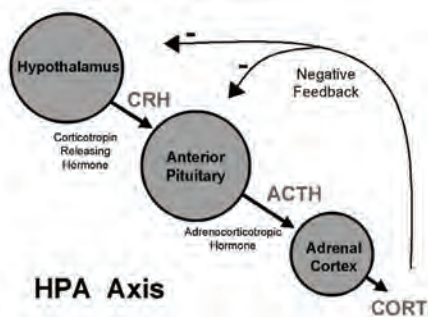
**Table 1.**  
*Symptoms of Attachment Disorder*  
(From Odhayani et al., 2013)

- **An aversion to touch and physical affection:** The child might flinch, laugh, or even say “ouch” when touched; rather than producing positive feelings, touch and affection are perceived as threats.
- **Control issues:** The child might go to great lengths to prevent feeling helpless and remain in control; such children are often disobedient, defiant, and argumentative.
- **Anger problems:** Anger might be expressed directly, in tantrums or acting out, or through manipulative, passive-aggressive behaviour; the child might hide his or her anger in socially acceptable actions, like giving a high-5 that hurts or hugging someone too hard.
- **Difficulty showing genuine care and affection:** The child might act inappropriately affectionate with strangers while displaying little or no affection toward his or her parents.
- **An underdeveloped conscience:** The child might act like he or she does not have a conscience and might fail to show guilt, regret, or remorse after behaving badly.

Furthermore, this issue threatens the wellbeing of the entire society. Beginning as toddlers, abused children can exhibit harmful behaviors toward their peers, having difficulty reacting to stress and venting their frustrations through expressions of anger and aggression such as verbal and physical bullying. Because observing and interacting with peers is also a central part of personality development during early childhood, other children could be negatively influenced and become infected with violent and aggressive behaviors. By teenage years, abused adolescents tend to engage in dangerous activities such as becoming addicted to substance abuse, participating in early sexual activity and prostitution, carrying guns, and joining gangs because they perceive these situations as safer or more welcoming than their environment at home (Odhayani et al., 2013). And unfortunately, child abuse doesn't just end with one generation. 20-30% of child abuse victims will abuse their own children, and early abuse may raise that probability to 70% (Maestripieri, 2005). Childhood abuse intervention exists to prevent these tragedies from repeating.

In addition to educating parents on the importance of childhood development, preventing abuse, and providing nurturing environments for abused children, determining how childhood abuse alters brain development could provide insights into how to clinically treat its many negative effects. There are many types of abuse, and people are affected by them in differing ways. These symptoms are the result of numerous alterations in the body that cannot be examined all at once. Therefore, this proposal will explore how specific regions of the brain are damaged by child abuse. If the damage cannot be reversed with the brain's natural repairing processes, methods to stimulate recovery will be explored.

In order to examine the chemical and neuronal changes of the brain in response to childhood stress, the concept of stress must first be defined. Corticotropin-releasing factor (CRF) and the hypothalamic-pituitary-adrenal (HPA) axis system are central figures in the brain's response to stress (Bremner, 2006). When the body psychologically interprets a stimulus as a threat, as children do when they are abused, the HPA axis is activated. The neurons in the hypothalamus are stimulated to excrete CRF (a.k.a. corticotropin-releasing hormone, or CRH), which stimulates the pituitary to release adrenocorticotrophic hormone (ACTH). This then stimulates the adrenal gland to release stress hormones such as cortisol and adrenaline. Cortisol then has a negative feedback on the axis to ensure that cortisol is not overproduced. In response to stress hormones, the body will become alert and metabolism will speed up; bodily processes such as breathing, blood flow, and blood sugar will increase in preparation for a "fight-or-flight response." Although the stress response is a natural system that is needed for survival, excessive stress and elevated hormone levels can be toxic (Bremner, 2006).



**Figure 1.**  
*The HPA Axis*

The hippocampus is the brain area most sensitive to adrenocortical steroid levels, as it has the highest concentration of glucocorticoid (a term interchangeable with cortisol) receptor sites in the brain. A study by Sapolsky et al. concluded that the glucocorticoid-sensitive hippocampal neurons play a significant role in terminating adrenocortical steroid production in the HPA axis (Sapolsky et al., 1984). Further research by Sapolsky et al. found that sustained elevated levels of corticosterone (the cortisol molecule in rats) would considerably reduce the number of corticosterone receptors in the hippocampus, while the receptor affinity of the hormone did not change. Within one week of cessation in elevated cortisol levels, the receptor numbers returned to normal levels (Sapolsky et al., 1984). The diminished receptor numbers may have occurred because excessive cortisol is toxic to the cortisol receptors or the cells to which the receptors are attached.

Another study by Meaney et al. showed that rats who had been exposed to profound stress in the form of maternal deprivation in their early lives had decreased numbers of glucocorticoid receptors in the hippocampus and other brain areas compared to rats that were raised without the stress (Meaney et al., 1988). According to Conrad (2008), excessive corticosterone levels sustained by chronic stress, but not brief periods of elevation caused by acute stress, have adverse effects on the hippocampus. These include decreasing neuronal regeneration, altering dendritic spine density, hindering synaptic plasticity, impeding long term potentiation, and

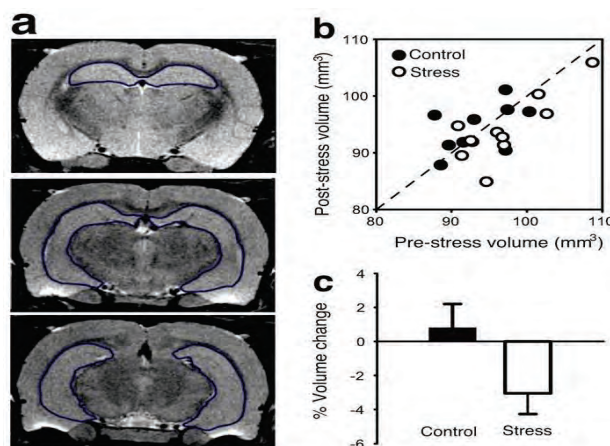
reducing dendritic complexity that may lead to the death of hippocampal neurons. More specifically, stress causes dendritic retraction, the reduction of dendritic complexity and total dendritic lengths, in the CA3 hippocampal region (Conrad, 2008).

These findings suggest that chronic stress causes the following sequence of events:

1. Corticosterone is produced as a natural response to stress.
2. The corticosterone levels are then elevated and sustained to the point that receptor numbers decrease in the hippocampus neurons.
3. The hippocampal neurons become unable to help activate the termination of corticosterone production because they have lost the ability to sense corticosterone levels.
4. The corticosterone levels continue to rise if receptor levels do not return to normal and direct the termination of cortisol excretion.
5. The long term high corticosterone levels lead to the degeneration and death of hippocampal cells.

At least in rats, profound chronic stress appears to produce a cascading effect of degeneration in the hippocampus. If similar patterns take place in the brains of abused children, there may be permanent brain damage.

There is more evidence to support the findings that cases of chronic stress, such as child abuse, lead to difficulty regulating cortisol levels and hippocampal degeneration. Stressed rats display deficient fast feedback in the HPA axis in response to corticosterone levels (Bremner, 2006). Research on primates that had experienced stressful events early in development showed long term behavioral changes compared to the control, in addition to elevated levels of CRF in the cerebrospinal fluid (Bremner, 2006). Rats that were restrained from moving for 6 hours a day for 21 days, a source of great distress, showed on average a 3% decrease in hippocampal volume according to MRI measurements before and after the experiment (Lee et al., 2009).



**Figure 2.**

(a) MRI of Coronal Sections of Hippocampus. (b) Pre and Post-stress Hippocampal Volumes. (c) Percent Change in Hippocampal Volumes.

(From Lee et al., 2009)



However, the most conclusive evidence comes from the studies with human research. It was found that patients with Cushing's disease, a condition of excessive cortisol secretion, have decreased hippocampal volume and alterations within their memory (Maheu et al., 2008). A study performed by Bremner on 17 adult patients around the age of forty with childhood abuse-related PTSD compared hippocampal volumes with a control group made up of non-abused, healthy people. These patients had a 12% reduction in their left hippocampal volume compared to the control (Bremner, 2006). Post traumatic stress disorder patients are similar to patients of severe child abuse in that in both, their HPA axis systems have had long term problems with dysregulation. Chronic stress from long periods of traumatic experiences and acute stress from sudden traumatic events can both lead to PTSD (Bremner, 2006). Severe stress could break down the body's neuroendocrine response to stress similarly to the rat models, leading to the misregulation of stress hormones such as cortisol, norepinephrine, and serotonin. Early dysregulation of the neuroendocrine systems involving stress is likely connected to the development of problematic behavior observed in abused children; an overly excited state of mind and a tense body can inadvertently lead to the anxiety and impulsivity seen as aggression and antisocial behavior by others (Bremner, 2006).

Depression is another mental disorder common in child abuse victims. Using MR imaging to measure hippocampal volume, Bremner et al. concluded that depressed women who had experienced sexual abuse as children had 18% smaller left hippocampal volumes compared to depressed women who had not experienced abuse, and 15% smaller left hippocampal volumes compared to healthy women who had neither depression nor experience of childhood abuse. Although the relationship between child abuse and depression is unclear, it appears that the factor of childhood abuse was the major cause of hippocampal shrinking, not the depression. Surprisingly, the right hippocampi of the three groups had similar sizes (Bremner et al., 2002). In another study, Bremner et al. examined the hippocampal volumes of patients with panic disorders and found that patients who had a history of sexual abuse during childhood had significant decreases in magnitude compared to the control. They did not find decreased hippocampal volume in patients without a history of abuse (Vythilingam et al., 1997).

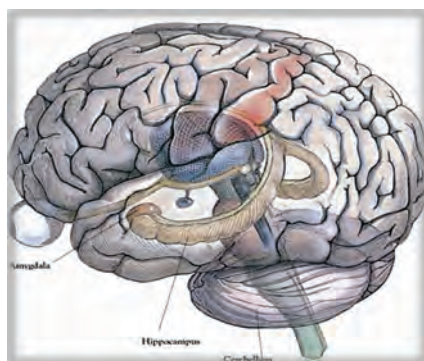
Degeneration of the hippocampus is also a major cause of diseases involving memory loss, such as dementia, epilepsy, and Alzheimer's disease. A statistical study performed on elderly Finnish men who had experienced childhood trauma, such as living in an orphanage or having childhood experiences of war, found that childhood stresses significantly increase the chance of developing AD and dementia (Donley et al., 2018). Child abuse may have more severe consequences than most people are aware of; in addition to decreasing quality of life, duration of life can be shortened as well.

Extensive research has been done on the effects of traumatic events, such as child abuse, on the human brain. The amygdala, the hippocampus, and the prefrontal cortex are the central brain areas that respond to stress. In particular, the hippocampus is "very sensitive to the effects of stress" (Bremner, 2006).

The hippocampus has been connected with significant roles in learning, memory, and spatial navigation. It is embedded subcortically to the temporal lobe and is part of the limbic system. The limbic system produces the brain's "primitive" emotions, such as hunger, sex, pain, pleasure, and memory. There are two hippocampi, one in each hemisphere of the symmetrically bifurcating limbic lobes, posterior to the amygdala. The amygdala is known as the emotional center of the brain that triggers the "fight-or-flight" response. Considering how the amygdala has

extensive connections with the hippocampus, the hippocampus plays a direct role in generating emotional reactions (Anand et al., 2012). Therefore, treating the hippocampus may be an avenue to directly addressing the source of aggressive and impulsive behaviors in addition to restoring stress hormone regulation.

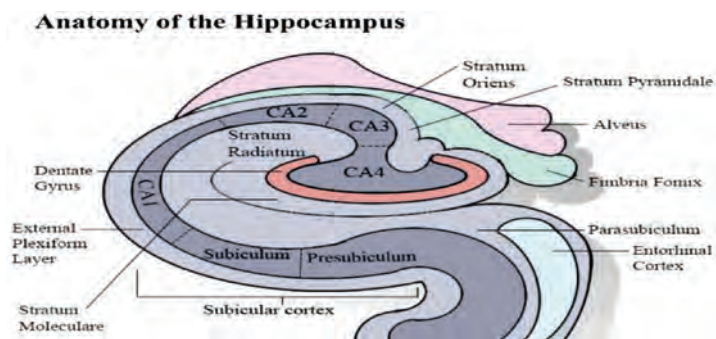
The hippocampus is one of the phylogenetically oldest parts of the brain, so the hippocampi and their functions in humans are thought to be similar to those of other mammals (Anand et al., 2012). Therefore, animal testing of the hippocampus may be able to reveal insights into the human hippocampus through tests that are allowed in animals but not on humans for ethical reasons.



**Figure 3.**

*Illustration of the Hippocampus  
(From Yamagishi, 2015)*

The hippocampus consists of two parts: the cornu ammonis (CA) and the dentate gyrus. It is part of the hippocampal formation that includes the CA, dentate gyrus, subiculum, and entorhinal cortex. The cornu ammonis is further subdivided into CA 1 through 4. The different sections of the hippocampus are interconnected with each other and other parts of the brain, and the firing of these interconnecting neuronal circuits is predicted to be involved in memory formation and spatial navigation. For example, the input the entorhinal cortex sends to the dentate gyrus is thought to play a role in pattern recognition and encoding memories (Anand et al, 2012).



**Figure 4.**

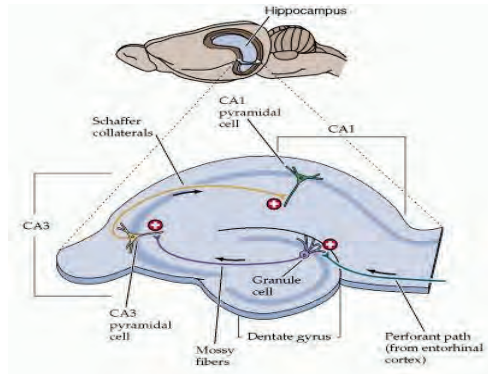
*Coronal View of the Human Hippocampal Formation and its Subsections  
(From Pang et al., 2019)*

Anterograde amnesia in patient H.M. after having significant portions of his hippocampus and its surrounding area excised is a famous case that exemplifies the importance of the hippocampus in forming memories. Patient H.M.'s memories from before his surgery remained intact, but he was incapable of storing his experiences as memories thereafter (Lynch, 2004).

Memory formation and learning can be neurologically explained by long term potentiation. As an experience is processed by the brain, certain neurons will fire together. Modifications in the synapses between sets of neurons make the set of neurons fire together even without the experience. Although an oversimplification, this connecting of neurons essentially stores a memory. Long term potentiation, or LTP, describes the synaptic modifications that involve synaptic receptors such as NMDA and AMPA (Lynch, 2004).

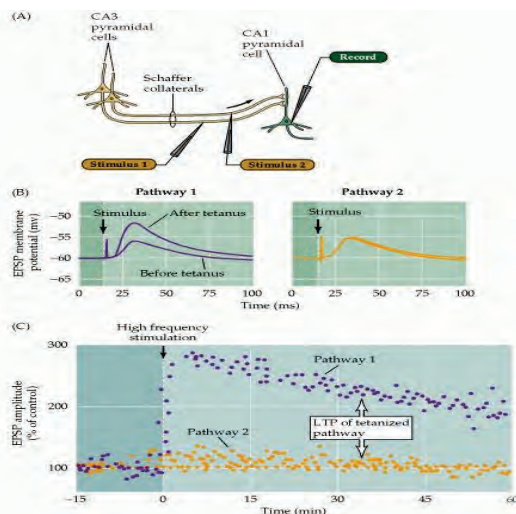
Because the acquisition of new memories requires the formation of new sets of neurons, hippocampus neurons are incredibly synaptically plastic compared to neurons in other brain areas, and are constantly forming new synaptic connections. Of the many divisions and subsections of the hippocampus, the CA3 region has the most connections with itself and with the other sections. The axons of CA3 pyramidal neurons, called Schaffer collaterals, form extensive connections with inhibitory and excitatory neurons in the other sections of the hippocampal formation. Neuronal circuits originating in the entorhinal cortex, called the perforant pathway, and the axons of neurons in the dentate gyrus, mossy fibers, are also known for having extensive connections with this subsection. The CA3 region has been extensively researched for its role in memory formation and susceptibility to neurodegeneration (Cherubini and Miles, 2015).

The Schaffer collaterals display essential mechanisms of information storage. For instance, the strength of excitatory postsynaptic potentials (EPSP) in the CA1 neurons innervated with Schaffer collaterals can be altered by the frequency of action potentials sent by the Schaffer collaterals. In an experiment using electrodes and voltmeters, EPSP in a single CA1 neuron was measured in response to excitation from a single Schaffer collateral. When receiving multiple action potentials within a few milliseconds, a high frequency, the CA1 neuron becomes sensitive to all successive action potentials from that Schaffer collateral for more than an hour, regardless of their frequencies; the EPSPs will have a higher amplitude than standard action potentials. But when the excitation occurs two to three times per minute, a low frequency, successive excitations from the Schaffer collateral will produce constant EPSPs without elevations (Purves et al., 2001). The elevated EPSP in the CA1 neuron from the high frequency excitation by the Schaffer collateral is a brief LTP, where synaptic receptors are modified to create a temporary neuronal circuit involved in information storage. Certainly, the absence or degeneration in the CA3 region would reduce the hippocampus' ability to form memories.



**Figure 5.**

*Diagram of Rodent Hippocampus and Major Structures  
(From Purves et al., 2001)*



**Figure 6.**

*Stimulation Frequency and Strength of ESPS  
(From Purves et al., 2001)*

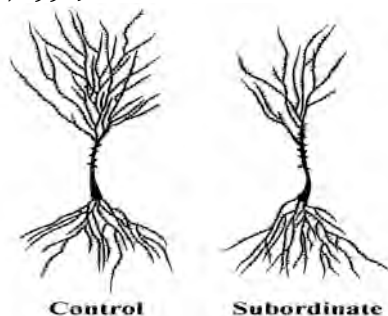
Deficiencies in verbal declarative memory have been attributed to a loss of neurons or damage in the CA3 region of the hippocampus. A study of epileptic patients who had hippocampal resections in their CA3 regions showed a loss of verbal declarative function after taking the Wechsler Memory Scale and Selective Reminding Test memory tests before and after surgery (Sass et al., 1995). Verbal declarative memory consists of episodic memories (memories of experiences, specific events, the time they occurred, and the contextual surroundings, like the circumstances and emotions involved in) and semantic memories (storage of acquired facts, meanings, and concepts that are not associated with the personal experiences and context of the memories) that can be stated verbally (Bremner, 2006) (Lynch, 2004). Patients of childhood abuse who underwent memory tests such as the WMS and verbal components of the SRT were also found to have deficient

verbal declarative memories. In fact, a deficit in verbal declarative memory is widely known to be a symptom of childhood abuse (Bremner, 2006).

What these facts cannot conclude is that damage in the CA3 region is solely responsible for impaired acquisition of verbal declarative memory. The hippocampus, and the brain in general, is too interconnected to designate a specific function to a specific area. The deficiencies in verbal declarative memory in the epileptic patients who had undergone CA3 resections could be explained by the fact the CA3 region plays an important part in the acquisition of declarative memory. However, what these facts can suggest is that the relationship between child abuse and degeneration in the CA3 regions of the hippocampus is worth investigating.

Magariños et al. performed a study that examined dendritic atrophy/retraction in the CA3 region as a result of chronic psychosocial stress. Because most studies on the hippocampal response to glucocorticoids were done on rodents, Magariños et al. selected a non-rodent species, the tree shrew, in order to determine the similarities in stress response between different mammalian species. Adult male tree shrews are extremely territorial, so placing two shrews in a confined space evokes psychosocial stress. After aggressively competing to establish separate territories, one shrew will claim a larger area and become dominant, while the other shrew becomes subordinate. During the 28 days that they were in contact, the subordinate shrews avoided aggression against the dominant shrews. Urine samples were collected throughout the period to measure cortisol levels in the dominant and subordinate shrews. The urine of subordinate shrews showed an increase in urinary cortisol levels as time passed, while the urine of the dominant shrews displayed stable cortisol levels that were similar to the control group. The control group had tree shrews living alone without competition stress. This suggests that the psychosocial stress of being subordinate caused the HPA axis to activate and continuously produce cortisol (Magariños et al., 1996).

The study also examined the changes in the structures of pyramidal neurons in the CA3 region of the hippocampus. After the confrontation period was over, the shrews were killed using a method that did not evoke additional stress, and their hippocampal regions were cut into thin slices for examination. After staining the slices with a variation of the Golgi stain, pyramidal neurons of the same subclass (subclasses have different apical branching patterns) were examined under a light microscope. The pyramidal neurons of subordinate shrews had significantly decreased apical dendritic branch points and apical dendritic lengths compared to controls (Magariños et al., 1996).



**Figure 7.**

*Effects of Chronic Stress and Elevated Cortisol Level on Dendritic Branching  
(From Magariños et al., 1996)*

However, morphology of dendritic “trunks” and dendritic spine density were not affected. In addition, neurons in the CA1, CA2, or Dentate Gyrus regions of the hippocampus were not physically affected in any way. Furthermore, 10 to 20 days after the end of the confrontation, the atrophy in the CA3 pyramidal dendrites was reversible. Magariños et al. suggested that the dendritic atrophy is a mechanism to protect the pyramidal neurons during periods of long but manageable stress. Elevated cortisol levels could excessively stimulate the pyramidal neurons, so reducing dendritic branches would reduce the number of incoming synaptic potentials, preventing the overexcitation of the neuron and its destruction (Magariños et al). However, dendritic retraction in the CA3 region could only be effective when the duration and severity of stress is manageable. In Conrad’s 2006 study on dendritic retraction, rats that were exposed to unpredictable stress for over 4 weeks began exhibiting dendritic retraction in areas in addition to CA3, such as in the CA1 and the dentate gyrus. They also began to show decreasing dendritic spine density, a symptom that was not present in the tree shrew experiment (Conrad, 2006). Possibly, intimidation from the dominant shrew during the 28 day period did not cause stress levels severe or long enough to begin altering dendritic spine density. Even more serious stress may be unmanageable and lead to neuronal death. In any case, degeneration in the CA3 region of the hippocampus could be an early and prevalent symptom of chronic stress.



**Figure 8.**

*Model of a Dendritic Spine  
(From Bear et al., 2016)*

In another study by Conrad (2006), mice raised in enriched environments and mice raised in depravity were assessed on their ability to solve mazes. Mice from enriched environments exhibited superior performance. The deprived rats may have performed poorly because their stressful rearing environments made their hippocampi vulnerable to degeneration. Maze solving requires the formation of memories, learning, and spatial navigation, which are all functions of the hippocampus. Therefore, having a healthy hippocampus may have been the reason for better performance by enriched rats. Conrad went on to dissect the subjects and confirmed that the mice raised in depravity had dendritic retraction and decreased dendritic spine density (Conrad, 2006).

If moving a child from an abusive environment to a more nurturing one could lessen the deleterious consequences of child abuse, that would be phenomenal. There are systems in place to accomplish this, when child welfare agencies send abused children to new homes, for example. However, what if abuse is not noticed and there is no intervention? What if the new foster homes or orphanages are also abusive and uncaring? Even if a caring home is found, what if the children keep suffering from the

effects of abuse? Experiencing continuous stress throughout childhood may inflict brain alterations that could become more severe as time passes. Assuming that irreversible damage occurs — recall that the left hippocampi of victims of sexual abuse during childhood had significantly decreased volumes in adulthood — a treatment for providing neurorehabilitation is needed.

The hippocampus “demonstrates an unusual capacity for neuronal plasticity and regeneration” (Bremner, 2006). This may be because “the hippocampus expresses the highest levels of Brain Derived Neurotrophic Factor mRNA in the brain” (Conrad, 2008). BDNF is a type of neurotrophin, naturally produced molecules that regulate neuronal “proliferation, migration, differentiation, survival, apoptosis, and synaptic plasticity.” The high expression of BDNF in the hippocampus makes researchers think it might be a central mediator in the process of learning and forming memory (Duan et al., 2018). When BDNF binds to tropomyosin receptor kinase B (TrkB), the BDNF-TrkB complex participates in numerous signaling systems, such as phosphorylation cascades and secondary signaling systems, and is able to significantly influence neuronal function (Liu and Nusslock, 2018). Nerve Growth Factors are another type of neurotrophin involved in synaptic plasticity. They regulate the “development, differentiation, growth, regeneration, and expression of functional traits” in neurons. NGF is also associated with enhanced memory function (Duan et al, 2018).

However, the production and excretion of neurotrophins is impaired by the presence of excessive cortisol levels. Stress decreases BDNF mRNA levels in the hippocampus (Anand et al., 2012). Looking back at Magariños’ tree shrew experiment, a resurgence of neurotrophic factors may have been the reason that dendritic atrophy in the subordinate tree shrews was reversed 10 to 20 days after the confrontation period.

On the other hand, it has not been ascertained whether neurotrophins can initiate neurogenesis. Research so far has only been able to determine two predominant areas in which neurogenesis occurs naturally throughout life: the subventricular zone (SVZ) of the lateral ventricles and within the dentate gyrus of the hippocampus (Barlow and Targum, 2007). Glial cells like myelin cells and astrocytes are different from neurons and are known to replicate. In 1998, Ericsson et al. determined that no other area in the hippocampus showed DNA markings of neurogenesis (Baptista and Adrade, 2018). But interestingly, increased levels of BDNF have been shown to significantly increase net neurogenesis in the dentate gyrus of the hippocampus (Liu and Nusslock, 2018). More research has been done to discover new areas where neurogenesis takes place, but conclusive evidence has not been found. Therefore, researchers are now attempting to discover the conditions that stimulate neurogenesis in the specialized regions and artificially amplify them in other brain areas in order to develop treatments that can repair deceased brain tissue. Increased neurotrophin levels may be one of those conditions.

However, the role of neurotrophins in stimulating synaptic plasticity has already been experimentally proven. Conner and Varon conducted a study that examined whether NGF was continuously required to maintain synaptic connections. After observing new synaptic connections form because of exogenous NGF infusion into rat hippocampi, they ceased NGF infusion and blocked the expression of endogenous NGF. As a result, the synaptic innervations detached without the presence of local NGF. This suggested that NGF levels may regulate the activation of axonal remodeling (Conner and Varon, 1996). Hypothetically, the absence of NGF could initiate a signaling system that leads to the detachment of axons from their synapses, while axons form and maintain innervations if NGF is present. If this is the



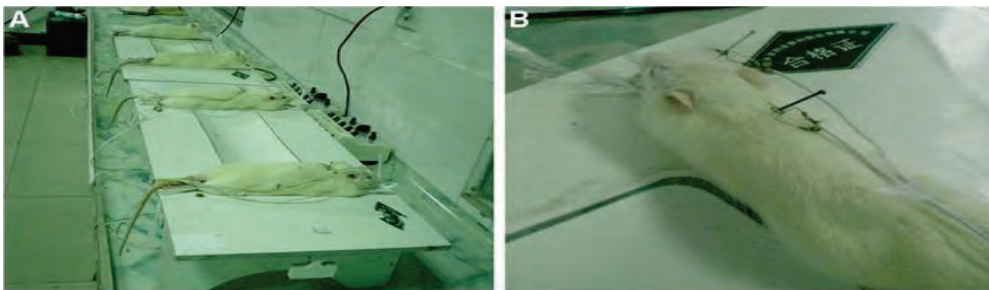
case, NGF may contribute to neuronal plasticity in the processes of learning and memory formation.

Furthermore, Mamounas et al. performed a study on rat brains to observe how BDNF infusions affected axonal regeneration in the CA3 region of the hippocampus that had been damaged with a neurotoxin called *p*-chloroamphetamine (PCA). PCA specifically targets axons of serotonergic (5-HT) neurons while sparing the soma. The PCA caused rapid degeneration of most 5-HT axons, but BDNF infusions led to significant axonal resprouting and “complete recovery” was observed near the infusion site. The concentration of BDNF in the infusions was proportionate to axonal sprouting, suggesting that more BDNF leads to more recovery. There was no “critical period” wherein BDNF needed to be administered before the damage became permanent, and the BDNF stimulated significant axonal sprouting even a month after the PCA was introduced.

Admittedly, numerous studies found neurotrophins like BDNF and NGF were unable to trigger neurogenesis. However, their effects on axonal sprouting and synaptic plasticity suggest that they could be used to combat the negative effects of chronic stress. Furthermore, they appear to be able to promote the regeneration of damaged neurons, preventing apoptosis and allowing neurons to survive and regenerate.

In a remarkable rehabilitation study, Duan et al. examined the neuroregenerative properties of electroacupuncture on ischemic rat brains. Electroacupuncture has been widely used as a treatment for stroke in human patients, and Duan et al. hypothesized that electroacupuncture was effective because it stimulates the release of BDNF and NGF. They wanted to see if neurotrophins could also rehabilitate the rats’ learning and memory formation abilities that had been damaged by ischemia. In order to induce the effects of stroke, the rats underwent a surgical procedure. After surgery, they showed signs of ischemia in their right hemispheres and right hippocampi; their left limbs, ipsilateral to the ischemic side, were deficient when they moved, and Horner’s signs were observed. Rats that survived the procedure and showed these symptoms were given electroacupuncture treatments (Duan et al., 2018).

The rats were treated with electroacupuncture (EA) on the Baihui (right midpoint of the parietal bone) and Dazhui (middle of the back between the C7 and T1 vertebrae) acupoints. Continuous electrical stimulation of 50-Hz frequency was conducted through the needles for 15 minutes once a day, for 14 successive days. The electricity was not intense enough to cause visible muscle contraction in the rats (Duan et al., 2018).



**Figure 9.**

*Electroacupuncture in Ischemic Rats*  
(From Duan et al., 2018)



In order to test for whether EA improved memory ability and spatial learning, the Morris water maze test was used. The Morris water maze test begins by training a rat to swim to a platform when placed in a circular tank filled with water. The platform is visible at first, but then the water is made cloudy and the water level is increased until the platform is no longer visible. The rat's ability to learn and form memories of the platform's location is quantitatively measured by the time it takes to find the submerged platform, the escape latency. In addition to the original version, Duan et al. incorporated a variant of the Morris water maze test. In this version, the platform was removed entirely, and the number of times that a rat swam across the location where the platform had been located in 2 minutes was counted (Duan et al., 2018).

The rats that had received EA had statistically significantly shorter average escape latencies ( $44.21 \pm 5.24$  seconds compared to  $70.17 \pm 5.12$  seconds), and the number of times that they swam across the platform location was also significantly greater than the ischemic rats that had not received EA ( $3.12 \pm 0.58$  crossings compared to  $1.41 \pm 0.32$  crossings). These results suggested that EA improved the memory and learning abilities in the stroke-induced rats (Duan et al., 2018).

Next, the researchers looked for signs of neurotrophin expression and neuronal recovery in the CA3 region to see if increased neurotrophin levels had caused recovery from ischemia. The rats were anesthetized, and their brains were harvested. The hippocampi were treated with several chemical solutions and sectioned coronally into  $5\mu\text{m}$  thickness. After immunohistochemical staining to reveal expression of BDNF and NGF, sections from the CA3 region were viewed under a light microscope. The control group that had not received EA showed low levels of BDNF and NGF and had degenerated neurons. The signs of degeneration included nucleus shrinkage, cytoplasmic contraction and deformation, and the absence of Nissl bodies. However, in the EA treated samples where BDNF and NGF were strongly expressed, "there was no obvious degeneration" other than a few neurons that showed disorganized structures and nerve fibers (Duan et al., 2018).

Curiously, the extent of brain damage and neuronal loss due to the ischemia and the extent of recovery in the EA treated rats was not mentioned. Testing of neuronal densities that could have revealed whether more neurons were saved because of neurotrophin repair was also not performed. Because the study neglected to mention whether new neurons were observed, this study will assume that neurogenesis did not occur (Duan et al., 2018).

Regardless of this, it appears that neurotrophins BDNF and NGF have rehabilitative qualities as well as the capacity to improve memory. Supplementing BDNF and NGF treatments with other forms of therapy could be a beneficial treatment for child abuse victims.

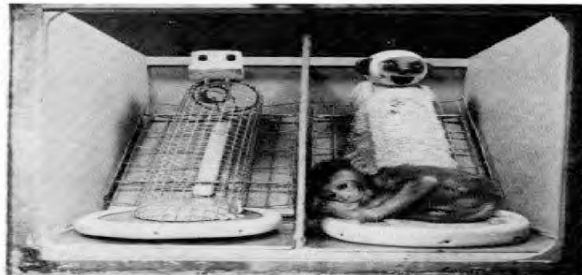
However, these studies have compatibility issues with this proposal's focus on researching brain damage in child abuse victims. In Magariños et al.'s tree shrew experiment, chronic stress was not delivered during the subjects' early developmental phase, and the cause of the stress, the dominant male, was not of parental relations to the subordinate shrew. Furthermore, brain damage from ischemia may not be similar to brain damage from chronic childhood stress — and, of course, the test subjects were not human children. Therefore, the question remains of whether prolonged child abuse can cause irreversible damage to the brain and whether rehabilitation at the neuronal level is possible.

Thus, Harry Harlow's monkey experiments may be useful in bridging the gaps. An advocate of attachment theory, Harlow also believed that a person's first attachment is towards his or her mother as an infant. He objected to the prominent

behavioral psychologists of his time who assumed attachments, such as love, were motivated by bestial needs — psychologists who believed romantic love was simply an expression of sexual desire and that the love an infant felt for its mother existed because of its dependence on her breast milk. Harlow began experimenting on macaque monkey babies to search for the origin of maternal love. According to Harlow, the macaques were ideal test subjects because “the basic responses relating to affection, including nursing, contact, clinging, and even visual and auditory exploration, exhibit no fundamental differences in the two species. Even the development of perception, fear, frustration, and learning capability follows very similar sequences in rhesus monkeys and human children” (Harlow, 1958).

In his famous surrogate mother experiment, he linked mother-infant attraction to the importance of affectionate physical contact and reaffirmed how essential a nurturing childhood was in the developmental process.

Macaque infants were separated from their mothers and raised in a cage with two inanimate objects that looked like mothers. One of the mothers was made out of wire and had a milk bottle attached that served as a breast. The wires of the wire mother were kept warm. The other mother was made out of foam and a terry cloth, but had no bottle (Harlow, 1958).

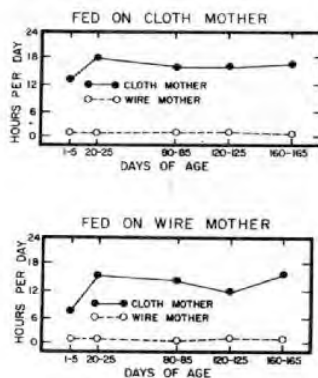


**Figure 10.**

*Wire Mother (left) and Cloth Mother (right)  
(From Harlow, 1958)*

As Harlow had predicted, the monkey infants remained hugging onto the cloth mother for the majority of the day and only sought the wire mother to drink from the bottle. The monkeys also displayed behaviors that indicated dependence on the cloth mother. When a scary object, a toy that made loud sounds and sudden movements, was introduced to the cage, the monkeys screamed and ran to the cloth mother to hold onto it. But after regaining composure, the majority of the subjects reapproached the toy and threatened it. On the other hand, when a cloth mother was not present, the monkeys would remain cowering in a corner.

In another experiment, when infant macaques were left alone in an unfamiliar room with unfamiliar objects, they would huddle into a ball and show signs of emotional distress. The same occurred when they were placed with a wire mother. On the other hand, when a cloth mother was present, the monkeys began exploring the room and ran to the cloth mother when startled. (Harlow, 1958). The behaviors observed in Harlow’s experiment are examples of how the comfort and support of a mother encourages babies to interact with the world; a hug from a mother can make monsters less scary and make unfamiliar places look safe to explore.



**Figure 11.**

*Graphs of Hours per Day Spent with Cloth Mothers Compared to Wire Mothers  
(From Harlow, 1958)*



**Figure 12.**

*Scary Toy Used to Frighten the Infant Macaques in Harlow's Experiments  
(From Harlow, 1958)*

In another of Harlow's experiments, macaques that had been raised without a mother and in isolation from other macaques during the first six months of their lives displayed inappropriate and abnormal social behavior such as aggression, emotional instability, and stereotype and self injurious behavior. Harlow called this phenomenon *Isolation Syndrome*. "Stereotypies are repetitive, seemingly functionless activities that often interfere with an individual's typical behavioral repertoire" (Lutz et al., 2003). Stereotype behaviors include pacing, digit sucking, hair pulling, bouncing, body flipping, rocking, huddling, and self-clasping. Common self injurious behaviors included self-biting, hair pulling, and head banging that were sometimes severe enough to require veterinary attention (Lutz et al., 2003).

Combining elements of the previously mentioned studies with Harlow's experiments may help us design the most effective experiment for this study.

## Hypotheses

1. If child abuse inflicts severe stress on victims, then their stress hormone levels at rest will become elevated. This is because the stress from being abused raises stress hormone levels too high for too long and breaks down the feedback

- mechanisms in stress hormone regulation. If this hypothesis is not substantiated, it could suggest that there are unknown mechanisms in the body that prevent the neuroendocrine system from going out of control.
2. If child abuse elevates resting stress hormone levels, then abused children will exhibit anxious and impulsive behaviors because their state of mind and body remain hypersensitized even in unthreatening situations. If this hypothesis is not substantiated, it could suggest that stress hormones do not have a direct effect on behavior control.
  3. If child abuse elevates resting stress hormone levels, then the hippocampal volumes in abused children will decrease because hippocampal neurons degenerate and die when stress hormone levels are elevated too high for too long. If this hypothesis is not substantiated, it could suggest that excessive hormone levels are not the cause of hippocampal degeneration or that there are unknown mechanisms that protect the hippocampus from excessive hormone levels.
  4. If child abuse causes hippocampal neurons to degenerate and die, leading to a decrease in the hippocampal volume, then the victims' memory ability will be impaired because the hippocampus plays an important role in memory formation and learning. If this hypothesis is not substantiated, it would suggest that hippocampal volume is unrelated to hippocampal functioning.
  5. If abused children are treated with neurotrophin therapy, then the extent of physiological alterations caused by child abuse will be lessened because neurotrophin has regenerative qualities capable of restoring the systems that child abuse breaks down. If this hypothesis is not substantiated, it would suggest that the neuroregenerative properties of neurotrophin are limited.

## Methods

This experiment will attempt to examine the neuronal alterations and degeneration in the hippocampus caused by child abuse and trace potential neuronal recovery from neurotrophin therapy. Within the hippocampus, we will focus on neurons in the CA3 region. Furthermore, how neuronal recovery translates to improvement in memory and how a return to normal resting stress hormone levels terminates abnormal behavior will be observed. In essence, this proposal seeks to discover the extent of recovery neurotrophin therapy can catalyze.

There will be 8 groups with 12 rhesus macaque (*Macaca mulatta*) subjects in each. The groups will simulate different childhood environments and the effects on the macaques will be observed. There will be 6 pairs of mother and infant in each group, and 3 of the infants will be male and 3 will be female, as hormonal differences due to gender may impact the results. While the groups will receive different treatments, certain characteristics in the macaques will be kept the same in order to attribute changes to the experimental variables.

Although this proposal attempts to examine the relationship between the brain and child abuse, raising a mother and infant by themselves would create a faulty and inconclusive study design. Rhesus macaques are pack animals, so their development is heavily influenced by interactions with members of their troop. Being isolated from other macaques may introduce entirely new variables other than child abuse, so it would be unreasonable to consider such an experiment as a model that could be related to humans. Rhesus macaque troops may number from 10 to 80 members, and there are usually 4 times as many females as males. To factor out the effects of having too many or too few troop members, the macaque subjects will be

from troops numbering between 30 and 50 and a 4:1 ratio of females to males will be maintained (Maestriperi and Hoffman, 2011). There will be 6 troops involved in this study, where one mother and infant pair from each treatment group will be raised in the same troop. Any unpredicted variables from a single troop will affect the treatment groups equally. This form of cluster sampling will decrease variability and make data more reliable to the general population of rhesus macaques. To accommodate this strategy, each troop will be from a different primate research center with enclosure settings set as similar as possible. Sources of stress like availability of food, sanitation, health, limited space, or competition within the troop may also impact parenting (Maestriperi and Hoffman, 2011). It is important to keep the macaques healthy and free of disease, as ill subjects might not be able to participate in the study. For food, the amount, type, and time fed will be kept the same, and standard maintenance practiced by zoos should suffice.

The infant is entirely dependent on the mother for the first 2 to 3 months and will stick close to her for 1 to 2 years even after weaning, up until another sibling is born. Mothers increasingly reject the infants' attempts to make contact in order to foster independence, and children will begin to engage in sex-specific behaviors and lifestyles. In the wild, most males will leave their natal group in search of a new troop around this age. Males rough play with other males, and females spend time with other females, but their behaviors are also dependent upon their hierarchical rank. (Maestriperi and Hoffman, 2011). Females reach sexual maturity at age 3 and males at age 4, when an increase in sex hormones develops secondary sexual characteristics and behaviors, much like puberty in humans. In the wild, females begin mating 3 to 4 years after birth, while males usually need to reach full body size at around 8 years to compete with other males for mates. 8 macaque years correlates to the end of puberty and late teen years in humans, the end of adolescence, when personality is said to be fully developed and solidified (Maestriperi and Hoffman, 2011). To determine whether the damage caused by child abuse cannot be reversed by the body's natural healing processes, the macaques' hippocampi will be monitored throughout 8 years. If mature macaques abused in childhood exhibit deficiencies compared to non-abused macaques, it would suggest that natural recovery is unlikely to occur. Various tests will be carried out to measure physiological changes caused by child abuse, and neurotrophin therapy in the form of electroacupuncture will be administered as a potential method of recovery.

Female rhesus macaques mate several times in their lives and give birth every 1 or 2 years. The quality of motherly care improves with age and more offspring raised, so the mothers selected for the experiment will be at least 8 years old and have raised at least 3 offspring (groups 3 and 4 will be exceptions). Although there exists a social hierarchy in every pack, where higher ranking members display more dominant behavior and lower ranked members are submissive, child abuse will most likely degenerate their hippocampi regardless of social standing; therefore, the hierarchical rank of the mothers will not be a variable in this experiment. Another thing to consider is that macaques in the same treatment group may have contrasting parenting styles. Rhesus macaques are highly intelligent animals and indeed have different personalities, like humans (Maestriperi and Hoffman, 2011). But parenting in humans is also incredibly varied, depending on the distinct circumstances of individual lives. Perhaps this variable makes comparing the study to human parenting more realistic. Hence, the small differences between the macaque parenting styles will be ignored and only the presence of abuse will be considered.

Group 1 will be the control group. Macaques will be raised by caring mothers and will exhibit the characteristics of a healthy childhood. Infants will spend their

days nursing and sleeping while their mothers carry them on their stomachs and backs. For several months after birth, this close physical contact and attention from their mothers is essential in their developmental process. After being weaned towards the end of their first year, they become increasingly independent until they no longer seek their mothers. Over time, juveniles develop social skills and find their place in the hierarchical structure of the troop (Maestriperi and Hoffman, 2011). This natural rearing will display what a healthy macaque brain should look like. The hippocampal volumes, neural compositions, and memory ability of group 1 will be compared throughout the 8 year period with the other groups to determine the effects of child abuse and neurotrophin therapy.

Group 2 will be identical to group 1, except that they will begin receiving neurotrophin therapy after they become separated from their mothers. If group 2 macaques show defects compared to group 1 macaques and other groups receiving neurotrophin therapy also exhibit the same signs, the defects can be attributed to neurotrophin therapy and not child abuse. This will help determine whether electroacupuncture or neurotrophin therapy is safe for use in human rehabilitation. On the other hand, perhaps group 2 will be more healthy and have better hippocampal functioning than group 1. This may reveal the potential for neurotrophin therapy to be used clinically or supplementally in humans.

Group 3 macaque infants will be separated from their mothers a few days after birth and will be isolated with the cloth mothers from Harlow's experiments. Seeking a soft and embraceable physical contact, they are expected to become attached to the inanimate cloth frames and think of them as their real mothers. A bottle with formula will be attached to the cloth mom, so the infants will also become dependent on it for food (Harlow, 1958). They will not come into contact with their biological mother or other macaques until they are 18 months of age, the average age when macaques become independent. They will be separated from the cloth mothers and put back into the troops that they were born in. This scenario simulates parental neglect and the lack of attachments early in life. The unresponsiveness of the cloth moms will most likely severely stress the infant macaques, as mothers are normally animated and affectionate. Furthermore, infant macaques play with other infants and receive affection from other members of the troop. With the absence of this kind of experience, group 3 macaques will most likely be socially inept and will be unable to assimilate into the troop (Harlow, 1958). Such happens in human victims as well. How the macaques are impacted will be observed.

Group 4 will be identical to group 3, except they will continuously receive neurotrophin therapy after they are separated from their cloth mothers and are introduced back to their troop. This is a model to see if abused human children can be treated with neurotrophin therapy after being separated from neglectful parents. Perhaps neurotrophin therapy can lessen the negative effects of childhood stress compared to group 3.

Group 5 will be made up of mothers that abuse their infants. Research has shown that 5-10% of all rhesus macaque infants are abused every year. Abuse is a behavior passed down in matriline, meaning that female monkeys that were abused as infants are more likely to abuse their own children (Maestriperi, 2005). The mothers in group 5 will be selected if they have displayed the following abusive behaviors to their previous infants: infant dragging, crushing, throwing, stepping or sitting on, hitting, or biting (Maestriperi, 2005). Rhesus macaque mothers tend to repeat abuse with all of their children and alternate between short periods of abuse and long periods of caregiving. If no mothers in a troop are abusive, they can be transferred in from somewhere else. Abuse is likely to happen when a mother is

feeling stressed, so having to adapt to a new troop may stress the mother and increase abuse intensity. In monkeys as well as humans, child abuse is correlated with parental vulnerability to stress and emotional problems (Maestriperi, 2005). During the 1 to 2 years that the macaque children live with their abusive mother, the stress will likely elevate their cortisol levels, and their hippocampi will shrink. The hippocampal volumes will be measured throughout the 8 year period to see if natural recovery occurs.

Group 6 will be identical to group 5, except the macaque children will begin receiving neurotrophin therapy after becoming independent from their abusive mothers. If they show improved memory abilities or healthier hippocampi compared to group 5, neurotrophin therapy could be explored as a treatment for the damage caused by child abuse.

Group 7 is a modeling a scenario in which child abuse prevention succeeds in separating a child from their abusive parent and finds them a more nurturing environment. After 6 months of living with their abusive biological mother, the child will be taken away and introduced into a different troop, where they will be adopted by a non-abusive mother. Cross-fostering in macaques is a common practice used in primate research centers to prevent inbreeding (Maestriperi, 2005). Because infants are most dependent on and closest with their mothers during the first 2-3 months, the abuse inflicted by the biological mother will still have a considerable effect. This strategy examines the extent to which foster care or adoption can heal and prevent hippocampal degeneration.

Group 8 is identical to group 7, except neurotrophin therapy will be administered once they are fostered by their new non-abusive mothers. This scenario attempts to see if nurturing environments can be more effective in healing and rehabilitating abused children when paired with neurotrophin therapy.

**Table 2.**  
*The 8 Treatment Groups*

<b>Group</b>	<b>Treatment</b>
1 (control)	Raised by non-abusive mother
2	Raised by non-abusive mother + neurotrophin therapy after childhood
3	Raised by cloth mother
4	Raised by cloth mother + neurotrophin therapy after childhood
5	Raised by abusive mother
6	Raised by abusive mother + neurotrophin therapy after childhood
7	Raised by abusive biological mother but adopted by non-abusive foster mother after 6 months
8	Raised by abusive biological mother but adopted by non-abusive foster mother after 6 months + neurotrophin therapy after starting adoption

### **Electroacupuncture**

Electroacupuncture will be administered similar to the treatment the rats received in Duan et al.'s study. After macaque children in the groups 2, 4, 6, and 8 become independent from their mothers, at around 18 months, they will receive neurotrophin therapy in the form of electroacupuncture. The treatment will be administered for 15 minutes every day for the entire balance of the 8 year period. After the subjects are anesthetized, metal acupuncture pins will be inserted into the Baihui and Dazhai acupuncture points, and a continuous 50-Hz current will be delivered. This is intended to increase BDNF and NGF levels as it did in the rat subjects. Although the macaque brains are larger than the rat brains and electroacupuncture may not be as effective in the macaques as in the rats, they are also receiving the treatment for a significantly longer time. The current will still be set at 50-Hz because excessive amounts of neurotrophin may cause adverse effects.

### **Blood Hormone Level Test**

All of the macaque children in the 8 groups will have their blood tested for stress hormone levels at rest. Every 3 months starting from birth, CRH (which reflects cortisol levels), 5-H<sub>1</sub>AA (which reflects serotonin levels), and MHPG (which reflects norepinephrine levels) levels will be measured. Because testing needs to reflect basal hormone levels without elevation caused by a stressful event, some precautions will have to be taken not to stress the subjects before testing. Testing macaques in groups 5 and 6 will only happen if abuse has not taken place in the last 3 hours. Capturing the animals for testing will be done causing as little stress as possible. In order to separate mother from child during the first 1-2 years for testing, both will be trained to voluntarily walk inside a transfer box where they will be anesthetized via intramuscular injections and moved to separate squeeze cages. Blood samples will be quickly collected from the femoral vein within 3 minutes to ensure the samples can reflect the resting hormone levels. The hormonal levels will be compared between groups to verify that changes to the hippocampus are caused by chronic stress during childhood. These procedures are based on Sanchez et al.'s study that measured CRH and ACTH levels in abusive rhesus macaque mothers (Sanchez et al., 2010).

### **Behavior Assessment**

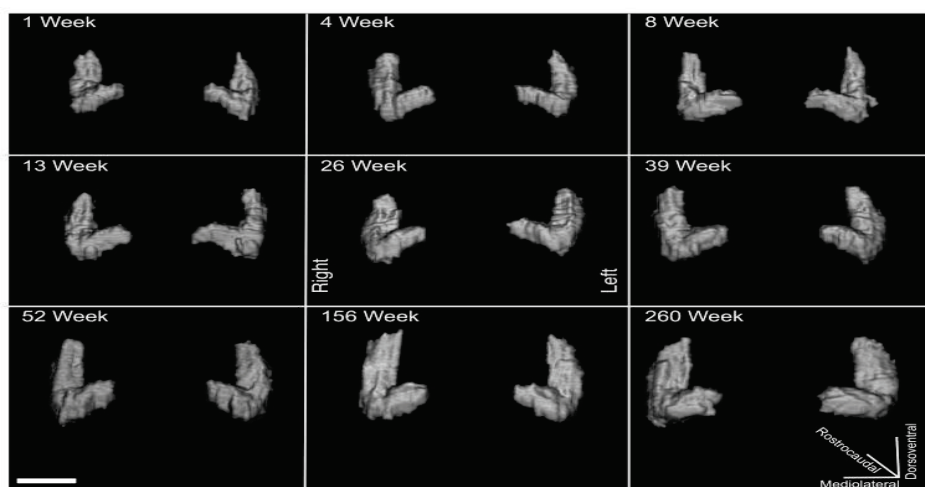
For 1 hour every 3 months, highly trained observers familiar with natural macaque behavior will look for inappropriate social behavior such as events of aggression with other troop members and stereotype/self injurious behavior in the subjects. Stereotype behaviors include pacing, digit sucking, hair pulling, bouncing, body flipping, rocking, huddling, and self-clasping. Common self injurious behaviors include self-biting, hair pulling, and head banging. Specific behaviors like biting and digit sucking will be counted separately, and the number of times a specific behavior was observed on one testing day will be averaged from the sum of the entire group. A line graph will be plotted with the number of abnormal behaviors observed at each testing as a function of successive testings every 3 months. This test will be performed on the subjects for 8 years starting from birth. In addition, chi square statistical analysis will be used to compare the frequency of abnormal behaviors between the different treatment groups. These methods will reveal the extent to which abuse and neurotrophin therapy influences abnormal behaviors.



## Hippocampal Volume Measurement Using MRI

While the subjects are still unconscious from anesthesia, their hippocampal volumes will be measured in an MRI machine. If an MRI is not present, they may have to be transported to a research center that has one. Scanning procedures are based on a 2014 study by Hunsaker et al. that measured hippocampal volume in rhesus macaques as they aged. The MRI will have the following settings while scanning: slice thickness = 0.70 mm; number of excitations (NEX) = 1; repetition time (TR) = 2200 ms; echo time (TE) = 4.73 ms; inversion time (TI) = 1100 ms; flip angle = 7°; field of view (FOV) = 180 mm; matrix = 256 × 256 (Hunsaker et al., 2014). The data will be compiled with the ANALYZE 11.0 program developed by the Mayo Clinic that traces the hippocampus using sagittal and coronal slice images. Hunsaker et al.'s study averaged the hippocampal volumes in rhesus macaque subjects as they aged and accounted for statistical variation and gender. They determined that hippocampi in rhesus macaques reach full size in 260 weeks (5 years). The hippocampal volume measurements in this study will be taken every 3 months and referenced to Hunsaker et al.'s data to determine if child abuse can cause a decrease in hippocampal volumes and whether the decrease can or cannot be stopped or reversed after childhood. It is possible that the hippocampus will keep shrinking due to a debilitated stress response system. Group 1 hippocampal volumes need to mirror Hunsaker et al.'s data in order for their study to be applicable to this proposal.

The data collected from groups 1 and 7 will be compared to determine if an environmental change can cause hippocampal growth or prevent volume decrease. Although unlikely to occur, hippocampal volumes might increase with neurotrophin therapy or over time, meaning that neurogenesis has happened. Every group's hippocampal volume averages will be graphed as a function of age. The graphs of the 8 groups will be compared.



**Figure 13.**

*MR imaging of the hippocampus  
(From Hunsaker et al., 2014)*

**Table 3.**

*Male Rhesus Macaque Hippocampal volumes  
(From Hunsaker et al., 2014)*

<b>Age</b>	<b>Left Hippocampus Volume (mm<sup>3</sup>)</b>	<b>Right Hippocampus Volume (mm<sup>3</sup>)</b>
1 week	162.84 +/- 25.19	159.40 +/- 26.44
4 weeks	232.42 +/- 23.87	225.64 +/- 24.98
8 weeks	260.51 +/- 33.43	264.42 +/- 32.34
13 weeks	310.52 +/- 24.97	305.08 +/- 25.06
26 weeks	341.42 +/- 40.55	342.15 +/- 44.19
39 weeks	367.93 +/- 29.21	365.17 +/- 42.65
52 weeks	398.84 +/- 40.89	396.37 +/- 42.60
156 weeks	458.06 +/- 32.86	458.67 +/- 36.10
260 weeks	457.36 +/- 51.32	453.31 +/- 52.03

**Table 4.**

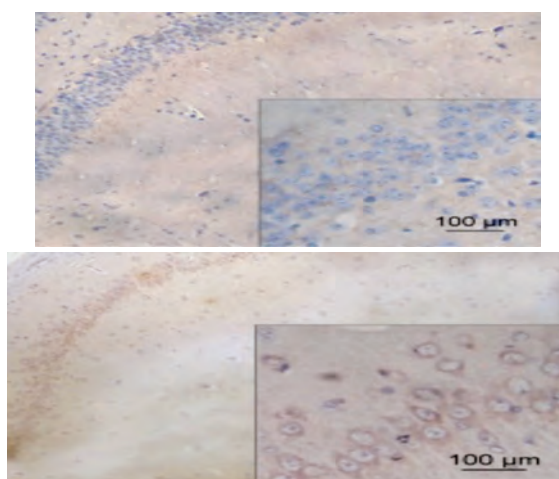
*Female Rhesus Macaque Hippocampal volumes  
(From Hunsaker et al., 2014)*

<b>Age</b>	<b>Left Hippocampus Volume (mm<sup>3</sup>)</b>	<b>Right Hippocampus Volume (mm<sup>3</sup>)</b>
1 week	168.04 +/- 28.95	159.88 +/- 27.24
4 weeks	233.44 +/- 20.57	232.10 +/- 22.44
8 weeks	256.95 +/- 32.04	261.49 +/- 26.39
13 weeks	293.00 +/- 31.28	286.86 +/- 30.55
26 weeks	312.95 +/- 37.48	311.10 +/- 44.42
39 weeks	350.78 +/- 36.55	353.08 +/- 47.31
52 weeks	387.15 +/- 43.64	385.17 +/- 60.14
156 weeks	444.08 +/- 47.59	441.91 +/- 43.22
260 weeks	445.42 +/- 46.44	439.27 +/- 47.65

## Hippocampal Dissection

Hippocampal volume decrease in the abused groups means that there are fewer neurons in the hippocampi of abused children compared to healthy ones. Unfortunately, it also means that neurons have died. In order to observe the extent of brain damage and the neuronal health differences in the treatment groups, the brains will be dissected and examined in detail with microscopy. 8 years after birth, all 48 macaque children will be sacrificed. In accordance with ethical guidelines, the subjects will be killed quickly and painlessly with a method that does not cause further brain damage. The hippocampus will be harvested and thin slices of the CA3 region will be prepared with chemical treatments and staining for viewing under light microscopes. The slices will consist of neurons within the CA3 region, as in Magariños et al.'s tree shrew hippocampus dissection, and slices containing the neurons that innervate pathways to and from the CA3 region, such as the Schaffer collaterals, mossy fibers, and perforant pathways. Primarily, signs of cell death and degeneration will be looked for in the child abused groups. These include comparatively low neuronal density, presence of neurons undergoing apoptosis, or the remains of already ablated cells. Nucleus shrinkage, cytoplasmic contraction and deformation, and absence of the Nissl body can indicate that neuronal degeneration occurred (Duan et al., 2018). Loss of dendritic and axonal branches and decreased dendritic spine density between neuronal innervations is also expected (Magariños et al., 1996).

In the groups that received neurotrophin therapy, whether neurotrophins BDNF and NGF were expressed will be verified. Once the slices are treated with immunohistochemical stains, BDNF and NGF will react and stain areas of high neurotrophin expression with a reactant product of brownish color. A darker stain indicates that more neurotrophins were expressed. If the groups that received electroacupuncture exhibit evidence of neurotrophin expression and have improved hippocampal CA3 region health, this would suggest that neurotrophin therapy can be applied to rehabilitate victims of chronic stress such as child abuse.



**Figure 14.**

*Immunohistochemical Staining Marking BDNF Expression in Rats that Received Electroacupuncture (bluish hues indicate scant expression, while brownish hues indicate increased expression). (From Duan et al., 2018)*

## Memory Test

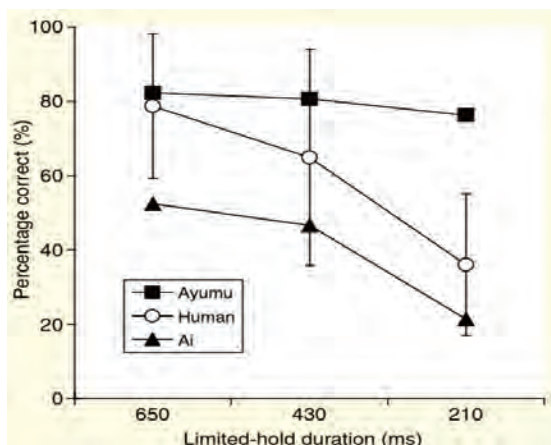
To quantify changes in hippocampal function, a memory test devised by neuroscientist Tetsuro Matsuzawa will be used. Chimpanzees were the test subjects, but rhesus macaques should be able to take it as well, although performance ability may not be as high. At the end of the 8 year period, before the macaque children are sacrificed, this memory test will be given to see if the different treatment groups demonstrate alterations in memory ability. The health of the hippocampus likely impacts memory ability, so the abused macaques will likely show impaired memory ability. If the neurotrophin therapy is revealed to cause regeneration in hippocampal neurons in addition to significantly improving memory, it will be evidence that neuronal regeneration was effective; the treatment was able to translate to improved hippocampal functioning. If macaques show improved cognitive functioning and no side effects, this form of neurotrophin therapy might be deemed safe enough for human testing and eventually for clinical use in child abuse victims.

The macaques will be trained to recognize the sequence of Arabic numerals from 1 to 9, 1-2-3-4-5-6-7-8-9, on a computer touch screen. For further details on how they were trained, see Inoue and Matsuzawa, 2007. Treats will be given for correct responses to motivate learning. Once they show mastery in tapping the numbers in the correct order, the macaques will be tested on tapping randomly arrayed sets of numbers in the correct sequence, but the numbers will be covered up with white squares after a brief glance. The macaques will have to tap on the numerals in the correct sequence based on their memory of where each number was. This tests their memory ability and the spatial recognition abilities of their hippocampi. The chimpanzees had less than a second to memorize the positions of the numerals before they were covered, 650, 430, and 210 milliseconds, yet some were able to correctly complete the sequence in 80% of 500 trials (Inoue and Matsuzawa, 2007). For other chimpanzees, the percentage of sequences completed correctly decreased with shorter durations. However, even macaques with a healthy hippocampus might not have the ability to memorize the positions almost instantly. After all, not even humans were as good as the chimpanzees at this task. Therefore, the 3 time periods used in this proposal will be increased if the macaques in group 1 cannot keep up. Whether the neurotrophin therapy or the mid-childhood elimination of abuse improved memory ability will be determined by higher percentages of sequences tapped correctly.



**Figure 15.**

*Picture of Matsuzawa's Memory Test  
(From Inoue and Matsuzawa, 2007)*



**Figure 16.**

*Graph of Percentage of Memory Tests Completed Correctly.  
Ayumu and Ai are Chimpanzees.  
(From Inoue and Matsuzawa, 2007)*

## Discussion

In the blood hormone level test, the hormone levels of the different groups taken every 3 months will be graphed alongside one another to observe how the different environmental settings of the groups generated stress and elevated hormone levels. We predict that abused monkeys in the abuse-only groups 3 and 5 will consistently have higher amounts of CRH, 5-H1AA, and MHPG compared to the control group throughout the 8 year testing period. Possibly, hormone values might return closer to control values because of natural repair, but this is unlikely because the chronic stress will have caused severe dysregulation in the neuroendocrine system. Hormone level differences could become more severe with time because of cascading degeneration in the HPA axis. On the contrary, groups 4 and 6, who received neurotrophin therapy after the childhood abuse, are predicted to have values closer to the control group than their abuse-only counterparts. Perhaps neurotrophin will directly or indirectly stimulate the genesis of cortisol receptors in the hippocampus and begin restoring hormone homeostasis. It is unknown whether neurotrophins can repair the HPA axis, so reduced hormone levels in these groups could lead to further research on the relationship between neurotrophins and the neuroendocrine system. Group 2 will most likely have hormone values similar to group 1. If group 2 values diverge significantly from group 1 values, it will mean that neurotrophin therapy could have harmful side effects. For group 7, it is predicted that the switch to a more nurturing environment will lower stress levels and cease further degeneration or maybe even repair the stress response system, bringing hormone levels closer to healthy levels. The addition of neurotrophin therapy in group 8 might cause this group to show better results than group 7, or perhaps speed up the repair process. Furthermore, an equation that predicts the average extent of hippocampal degeneration based on stress hormone levels could be created by finding the relationship between blood

hormone levels and decreases in volume. This formula could help indicate what stress levels are too high and when clinical intervention would be required. Replications of this study could verify its accuracy, and this equation could be referenced when an equation that applies to humans is formulated.

Abnormal behaviors are predicted to be more numerous in the abused groups. The chi square analysis will most likely show that the variable of abuse significantly increases the frequency of abnormal behaviors. These tendencies can be explained by elevated resting stress hormone levels. Elevated stress hormones would make the macaques more sensitive to environmental stimuli and make them perceive benign stimuli as threats. Thus, their behavior will be aggressive and impulsive. Possibly, they may engage in stereotype and self injurious behavior as a calming behavior to lessen their anxiety caused by their chronic stress. I expect that neurotrophin therapy administered in groups 4 and 6, after they become independent, will decrease resting stress hormone levels and the number of abnormal behaviors compared to groups 3 and 5. If so, the results will support the hypothesis that neurotrophin therapy can aid recovery in the neuroendocrine stress regulation system, leading to the reduction of abnormal behavior that had been caused by excessive stress hormones.

Neurotrophin therapy in this proposal was administered as electroacupuncture, but other methods could be used instead, such as direct infusion of neurotrophins into the hippocampus or bloodstream. It is not clear how electroacupuncture leads to neurotrophin production, so there may be more efficient methods to control the amount and type of neurotrophins expressed.

In the hippocampal volume measurements from MR imaging, it is predicted that the subjects will show the same pattern of results as in the blood hormone testing. Hippocampal volume and hormone levels are directly correlated — elevated stress kills neurons and decreases volume. Abuse-only groups 3 and 5 will have decreased volumes compared to Hunsaker et al.'s measurements, while hippocampal volumes in neurotrophin therapy groups 4 and 6 will be closer or within normal limits. The volumes will still increase over time because the hippocampus grows with age, but they will be less than reference values. Plotted on a line graph, the volumes might show decline after reaching full size at 5 years because of continued elevated stress hormones and cascading degeneration in the regulation mechanisms. This would mean that natural repair is unlikely. A graph may also show abused hippocampal volumes being consistent but lower than the control because the HPA axis and hormone regulation mechanisms were repaired naturally and did not cause further hippocampal degeneration. The most desirable outcome is that neurotrophin therapy would return volumes to within normal limits, or volumes would keep increasing after 5 years. If volumes increase, it will mean that neurogenesis occurred.

Group 1 and 2 data will most likely be within normal limits. Group 7 volumes will likely be higher than group 5 volumes, and group 8 values are hopefully even closer to the control. If neurotrophin therapy can consistently increase or prevent further decline in hippocampal volumes, further studies could be conducted to learn more about how neurotrophins function and how they can be applied.

The condition of hippocampal neurons in the CA3 region are also predicted to improve with neurotrophin therapy and cross-fostering. On the other hand, abused hippocampal CA3 regions are likely to display signs of cell death and degeneration. The decreased number of healthy neurons and the degenerated conditions of the remaining ones will likely lead to decreased brain functioning and poorer performance on the memory test. A close up view of the prepared slices under a microscope may reveal dead and dying neurons with evidence of apoptosis and

deformed neuronal structures; axons dendrites in the Schaffer collateral, mossy fiber, and perforant pathway innervations will likely show branch retraction and there will be decreased dendritic spine density. If immunohistochemical staining reveals neurotrophin expression in groups 4 and 6 along with hippocampal CA3 neurons in better condition compared to groups 3 and 5, neurotrophin's capacity to initiate regeneration will be supported.

If signs of neurogenesis are observed, research on neurotrophin therapy can be extended to design treatments for degenerative diseases such as Multiple Sclerosis, Parkinson's, Huntington's, or Alzheimer's disease.

The hippocampal conditions created in the different treatment groups should also influence performance accuracy on Matsuzawa's memory test. The primary purpose of the memory test is to quantify the regenerative properties of neurotrophin therapy. If groups 2, 4, 6, and 8 have higher percentages of sequences completed correctly than groups 1, 3, 5, and 7, neurotrophin's ability to improve brain functioning will be supported. In addition, poorer performance by abused groups will suggest that child abuse decreases brain functioning and reconfirm how dangerous child abuse is.

The limitations of this proposal must be acknowledged. A predominant one is that it cannot use human children as subjects. Experimenting with abuse on humans is unethical, so animals that respond similarly to humans must be used instead. However, although Harlow believed infantile behaviors "exhibit no fundamental differences in the two species," there are glaring differences in the two species' childhoods. For one, human children are much more dependent on their parents than rhesus macaques. While macaques become independent from their mothers at two years of age, modern human society is structured so that children must depend on their parents or guardians until 18 years of age. Furthermore, human child abuse can come from fathers as well as mothers, even teachers, neighbors, or foster parents. And since childhood is much longer in humans than in macaques, abused human children must bear abuse for a critically longer duration. It is then possible that physiological degeneration in human children is more severe.

Another area of concern is that there may be inconsistent links among the symptoms considered in this proposal. For example, abnormal behaviors like aggression or self harm may persist even when stress hormone levels and the hippocampus recover. Such a result would weaken our assumption that stress hormone levels and the hippocampus play a major role in behavior control. If so, perhaps another or several other brain regions not taken into account may be responsible. Unexpected results that confound the connections established in this paper would suggest that the brain regions are much more interconnected and less specialized than previously thought, serving as a reminder of the complexity of the brain, molecules, and behavior.

If the hypotheses are proven correct, this proposal should be extended to human trials. Considering the prevalence and severe consequences of child abuse, neurotrophin therapy at its experimental stages needs to be designed into a clinically applicable treatment as soon as possible. Testing in humans is a necessary step in this process. An extension of this proposal would verify if the changes in the macaques can also occur in humans and would investigate more variables that could improve the effectiveness of the treatment. Instead of stimulating abuse on the subjects, human studies would involve testing on children who have already suffered abuse. The extension would look similar to this animal model: blood hormone tests, clinical behavioral observations, hippocampal volume measurements with MR imaging, and memory tests can be replicated in humans, and after gathering samples with similar

characteristics, such as type of abuse, duration of abuse, and severity of symptoms, existing therapies for child abuse would be paired with neurotrophin therapy to see if neurotrophin could enhance results. Pharmacotherapy, counseling with a psychiatrist, and nurturing foster care are additional treatments that would be included. In human studies, perhaps gathering subjects for samples would be more efficient if they were recruited from foster homes. Furthermore, neurotrophin therapy should be tested at different intensities. Electroacupuncture at varying electrical intensities or number of times administered could alter the amount of neurotrophins expressed and impact results. Forms of neurotrophin therapy other than electroacupuncture such as infusions or genetic editing could also be more potent. And in addition to the hippocampus, extended studies could test other major brain areas affected by child abuse, such as the amygdala and prefrontal cortex, to see if neurotrophin can induce recovery in those regions as well.

## Conclusion

The consequences of childhood abuse are far reaching and profound. Parents are responsible for forming a trusting and nurturing relationship with their children, and abuse is an action that disregards a child's well-being. Growing up, abused children may find themselves disadvantaged compared to their peers and blame themselves for conditions that were forced on them by their abuse. Considering attachment theory, aggressive and antisocial behaviors are psychological defense mechanisms to prevent further physical and emotional trauma; hostility and impulsiveness are a result of heightened sensitivity from elevated stress hormone levels; poor performance in academic settings is possibly due to deficient hippocampi; mental illnesses such as depression and anxiety disorders develop because the children cannot cope with trauma. It's not their fault.

In addressing this issue, this proposal took the first steps towards better understanding the physiological deformations caused by child abuse and how to remedy them. There are many types of abuse and people are affected by them in differing ways. The symptoms are a result of numerous physiological alterations that are difficult to examine all at once. In other words, research in this field must take the approach of addressing specific problems, a few at a time. This proposal aimed to rehabilitate stress hormone regulation and the CA3 region of the hippocampus using neurotrophins with the intention of repairing memory and decreasing unnatural behaviors in abused children. In much the same way that individual neurons form neuronal circuits, neuronal circuits make up specialized brain areas, and specialized brain areas collectively create consciousness, addressing many small issues will eventually solve the bigger ones. Hoping that even the smallest of discoveries can make a difference, research in this field progresses.

## References

- Anand, K. S., & Dhikav, V. (2012). Hippocampus in health and disease: An overview. *Annals of Indian Academy of Neurology*, 15(4), 239–246. <https://doi.org/10.4103/0972-2327.104323>
- Barlow, C., & Targum, S. D. (2007). Hippocampal Neurogenesis: Can it be a Marker for New Antidepressants?. *Psychiatry (Edgmont (Pa.: Township))*, 4(5), 18–20.



- Baptista, P., & Andrade, J. P. (2018). Adult Hippocampal Neurogenesis: Regulation and Possible Functional and Clinical Correlates. *Frontiers in neuroanatomy*, 12, 44. <https://doi.org/10.3389/fnana.2018.00044>
- Bremner, J. D. (2006). Traumatic stress: effects on the brain. *Dialogues in Clinical Neuroscience*, 8(4), 445–461.
- Bremner, J. D., Randall, P., Vermetten, E., Staib, L., Bronen, R. A., Mazure, C., Capelli, S., McCarthy, G., Innis, R. B., & Charney, D. S. (1997). Magnetic resonance imaging-based measurement of hippocampal volume in posttraumatic stress disorder related to childhood physical and sexual abuse—a preliminary report. *Biological Psychiatry*, 41(1), 23–32. [https://doi.org/10.1016/s0006-3223\(96\)00162-x](https://doi.org/10.1016/s0006-3223(96)00162-x)
- Cherubini, E., & Miles, R. (2015). The CA3 region of the hippocampus: how is it? What is it for? How does it do it?. *Frontiers in Cellular Neuroscience*, 9, 19. <https://doi.org/10.3389/fncel.2015.00019>
- Conner, J. M., & Varon, S. (1996). Maintenance of sympathetic innervation into the hippocampal formation requires a continuous local availability of nerve growth factor. *Neuroscience*, 72(4), 933–945. [https://doi.org/10.1016/0306-4522\(95\)00598-6](https://doi.org/10.1016/0306-4522(95)00598-6)
- Conrad, C. D. (2008). Chronic stress-induced hippocampal vulnerability: the glucocorticoid vulnerability hypothesis. *Reviews in the Neurosciences*, 19(6), 395–411. <https://doi.org/10.1515/revneuro.2008.19.6.395>
- Conrad, C. D. (2006). What is the functional significance of chronic stress-induced CA3 dendritic retraction within the hippocampus?. *Behavioral and Cognitive Neuroscience Reviews*, 5(1), 41–60. <https://doi.org/10.1177/1534582306289043>
- Duan, X., Zhang, L., Yu, J., Wei, W., Liu, X., Xu, F., & Guo, S. (2018). The effect of different frequencies of electroacupuncture on BDNF and NGF expression in the hippocampal CA3 area of the ischemic hemisphere in cerebral ischemic rats. *Neuropsychiatric Disease and Treatment*, 14, 2689–2696. <https://doi.org/10.2147/NDT.S183436>
- Donley, G., Lönnroos, E., Tuomainen, T. P., & Kauhanen, J. (2018). Association of childhood stress with late-life dementia and Alzheimer's disease: the KIH D study. *European Journal of Public Health*, 28(6), 1069–1073. <https://doi.org/10.1093/eurpub/cky134>
- Harlow, H. F. (1958). The nature of love. *American Psychologist*, 13(12), 673–685. <https://doi.org/10.1037/h0047884>
- Hunsaker, M. R., Scott, J. A., Bauman, M. D., Schumann, C. M., & Amaral, D. G. (2014). Postnatal development of the hippocampus in the Rhesus macaque (*Macaca mulatta*): a longitudinal magnetic resonance imaging study. *Hippocampus*, 24(7), 794–807. <https://doi.org/10.1002/hipo.22271>
- Inoue, S., & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology*, 17(23). <https://doi.org/10.1016/j.cub.2007.10.027>
- Lee, T., Jarome, T., Li, S. J., Kim, J. J., & Helmstetter, F. J. (2009). Chronic stress selectively reduces hippocampal volume in rats: a longitudinal magnetic resonance imaging study. *Neuroreport*, 20(17), 1554–1558. <https://doi.org/10.1097/WNR.0b013e328332bb09>
- Liu, Patrick Z, and Nusslock, Robin. “Exercise-Mediated Neurogenesis in the Hippocampus via BDNF.” *Frontiers in Neuroscience*, vol. 12, 52. 7, Feb. 2018. <https://doi.org/10.3389/fnins.2018.00052>
- Lutz, C., Well, A., & Novak, M. (2003). Stereotypic and self-injurious behavior in rhesus macaques: a survey and retrospective analysis of environment and

- early experience. *American Journal of Primatology*, 60(1), 1–15. <https://doi.org/10.1002/ajp.10075>
- Magariños, A. M., McEwen, B. S., Flügge, G., & Fuchs, E. (1996). Chronic psychosocial stress causes apical dendritic atrophy of hippocampal CA3 pyramidal neurons in subordinate tree shrews. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience*, 16(10), 3534–3540. <https://doi.org/10.1523/JNEUROSCI.16-10-03534.1996>
- Maestripieri, D., & Hoffman, C. (2011). *Behavior and Social Dynamics of Rhesus Macaques on Cayo Santiago*. Chicago, IL: Uchicago. Retrieved September 09, 2020, from <http://primate.uchicago.edu/2011CayoBook02.pdf>
- Maestripieri, D. (2005). Early experience affects the intergenerational transmission of infant abuse in rhesus monkeys. *Proceedings of the National Academy of Sciences of the United States of America*, 102(27), 9726–9729. <https://doi.org/10.1073/pnas.0504122102>
- Maheu, F. S., Mazzone, L., Merke, D. P., Keil, M. F., Stratakis, C. A., Pine, D. S., & Ernst, M. (2008). Altered amygdala and hippocampus function in adolescents with hypercortisolemia: a functional magnetic resonance imaging study of Cushing syndrome. *Development and Psychopathology*, 20(4), 1177–1189. <https://doi.org/10.1017/S0954579408000564>
- Mamounas, L. A., Altar, C. A., Blue, M. E., Kaplan, D. R., Tessarollo, L., & Lyons, W. E. (2000). BDNF promotes the regenerative sprouting, but not survival, of injured serotonergic axons in the adult rat brain. *The Journal of Neuroscience: the official journal of the Society for Neuroscience*, 20(2), 771–782. <https://doi.org/10.1523/JNEUROSCI.20-02-00771.2000>
- Meaney, M. J., Aitken, D. H., van Berkel, C., Bhatnagar, S., & Sapolsky, R. M. (1988). Effect of neonatal handling on age-related impairments associated with the hippocampus. *Science (New York, N.Y.)*, 239(4841 Pt 1), 766–768. <https://doi.org/10.1126/science.3340858>
- Lynch, M. A. (2004). Long-term potentiation and memory. *Physiological Reviews*, 84(1), 87–136. <https://doi.org/10.1152/physrev.00014.2003>
- Odhayani, A., Watson, W. J., & Watson, L. (2013). Behavioural consequences of child abuse. *Canadian family physician Medecin de Famille Canadien*, 59(8), 831–836.
- Pang, C. C., Kiecker, C., O'Brien, J. T., Noble, W., & Chang, R. C. (2019). Ammon's Horn 2 (CA2) of the Hippocampus: A Long-Known Region with a New Potential Role in Neurodegeneration. *The Neuroscientist: a Review Journal Bringing Neurobiology, Neurology and Psychiatry*, 25(2), 167–180. <https://doi.org/10.1177/1073858418778747>
- McLeod, Saul. "Attachment Theory." *Simply Psychology*, 2017, [www.simplypsychology.org/attachment.html](http://www.simplypsychology.org/attachment.html).
- Purves, D., et al. (2001). *Neuroscience 2nd edition*. Sunderland, MA: Sinauer Associates.
- Sanchez, M. M., McCormack, K., Grand, A. P., Fulks, R., Graff, A., & Maestripieri, D. (2010). Effects of sex and early maternal abuse on adrenocorticotropin hormone and cortisol responses to the corticotropin-releasing hormone challenge during the first 3 years of life in group-living rhesus monkeys. *Development and Psychopathology*, 22(1), 45–53. <https://doi.org/10.1017/S0954579409990253>
- Sapolsky, R. M., Krey, L. C., & McEwen, B. S. (1984). Glucocorticoid-sensitive hippocampal neurons are involved in terminating the adrenocortical stress response. *Proceedings of the National Academy of Sciences of the United*

- States of America*, 81(19), 6174–6177. <https://doi.org/10.1073/pnas.81.19.6174>
- Sass, K. J., Buchanan, C. P., Kraemer, S., Westerveld, M., Kim, J. H., & Spencer, D. D. (1995). Verbal memory impairment resulting from hippocampal neuron loss among epileptic patients with structural lesions. *Neurology*, 45(12), 2154–2158. <https://doi.org/10.1212/wnl.45.12.2154>
- Vythilingam, M., Heim, C., Newport, J., Miller, A. H., Anderson, E., Bronen, R., Brummer, M., Staib, L., Vermetten, E., Charney, D. S., Nemeroff, C. B., & Bremner, J. D. (2002). Childhood trauma associated with smaller hippocampal volume in women with major depression. *The American journal of psychiatry*, 159(12), 2072–2080. <https://doi.org/10.1176/appi.ajp.159.12.2072>



# “¿Que Dijo?” Adolescent Feelings Towards Language Brokering and its Effects on Family Conflict

---

*Ethan Estrada*

**Author Background:** *Ethan Estrada grew up in the United States and currently attends the Liberal Arts and Science Academy High School in Austin, Texas in the United States. His Pioneer research concentration was in the field of psychology and titled “The Psychology of Immigration.”*

## **Abstract**

This study assessed whether ethnic identity moderates the relationship between Latino adolescents' feelings about language brokering and family conflict. The association between negative feelings about language brokering and parentification was also examined. Feelings about language brokering, family conflict, ethnic identity, and parentification were assessed in a sample of Latino adolescents (N = 384). Results show that high ethnic identity was correlated with low levels of family conflict regardless of individuals' feelings about language brokering. The most family conflict was found for individuals who reported negative feelings about language brokering and low ethnic identity. In addition, negative feelings about language brokering were correlated with parentification. These findings suggest that high ethnic identity may be a protective force against family conflict. Findings are discussed with regard to factors that affect family functioning.

*Keywords:* Latinos, language brokering, parentification, ethnic identity, family conflict, adolescents, immigrant

## **1. Introduction**

While English has become more widely spoken among Latinos as a whole, only 34% of foreign-born Latinos say they speak English proficiently (Krogstad et al., 2015). Additionally, although the 1964 Civil Rights Act guarantees individuals with Limited English Proficiency (LEP) the right to a translator and interpreter, this requirement is often unmet, with only 56% of hospitals reporting that they have language translation/interpretation services. For the children of foreign-born Latinos or other Latinos with LEP, the duty of translation and interpretation, therefore, often falls on them.

Despite the commonplace practice of translating for their parents, there

is limited research on what impact language brokering has on Latino adolescents and their family. Accordingly, this study examines feelings about language brokering in relation to family conflict for Latino adolescents in high school, using ethnic identity as a variable that moderates this relationship. Additionally, the study assesses the relationship between parentification and negative feelings about language brokering.

### **1.1. Effects of Language Brokering**

Language brokering is the process of translation, interpretation, and general mediation of information typically engaged in by immigrant children to aid their parents (Weisskirch 2005). Language brokering often occurs as a result of the child's faster adaptation to U.S. society, including their increased fluency in English and understanding of more aspects of U.S. culture in comparison to their parents (Lazarevic, 2017). For Latino immigrants, language brokering has been observed to be beneficial for youth who are positively connected or bonded to their family (Love & Buriel, 2007). Simultaneously, language brokering has been associated with an increased likelihood of depression and family conflict (Love & Buriel, 2007; Martinez, Jr. et al, 2008; Shen & Dennis, 2017). While different studies attribute these positive or negative outcomes to various phenomena, most see ethnic identity as a key variable that impacts these outcomes because of the importance of the Latino cultural value *familismo*.

However, despite the centrality of family among Latinos, most of the literature on language brokering has mainly focused on its psychological effect on the Latino individual and has found that language brokers with low ethnic identity see their parents' needs as burdensome (Monzó, 2016). Furthermore, intergenerational differences amplify these feelings. In their study of first- and second-generation Israeli immigrants, Oznobishin and Kurman (2016) found that whereas first-generation individuals felt more obligated to their parents when their help was needed to language broker, second-generation participants were less empathetic and understanding, believing their parents should have already been adjusted to the host culture. Generational status may therefore be a proxy variable for strength of ethnic identity and a variable that affects the occurrence of family conflict.

### **1.2. The Importance of Ethnic Identity**

Past research has indicated that ethnic identity can serve a promotive and protective role in Latino adolescent development. In a recent study, positive ethnic identity served as a buffer in stressful situations and was also associated with better psychosocial adjustment (Constante et al., 2019). Moreover, when parents actively promoted behaviors related to cultural maintenance, participants had increased ethnic language proficiency, and ethnic language proficiency was tied to both better academic achievement outcomes and other positive externalizing behaviors (Phinney et al., 2001).

Overall, Phinney et al.'s study underscores the importance of the role of the family in ethnic socialization. Thus, contrary to the typical belief that youth seek additional autonomy and separation from their parents to further develop their identity, for Latinos family cohesion is perhaps more important than

autonomy (Constante et al., 2019). This means that the maintenance of positive family relationships and increased family cohesion during adolescence may be instrumental in fostering positive internalizing and externalizing outcomes for Latino youth. Since language brokering has been shown through study to be related to adolescent individuals' perception of ethnic and cultural identity as well as family dynamics (Umaña-Taylor et al., 2013, Zhang et al., 2020) this study examines the relationship between these variables.

### **1.3. The Role of Parentification**

While family obligation and togetherness may unite many Latino families, parentification is a phenomenon that is still a variable worth assessing. Briefly defined, parentification is the role reversal of parents and children that occurs when children assume certain roles and responsibilities typically reserved to their parents (Hooper et al., 2012). Some research points to parentification as having positive correlates, including increased maturity, resilience, etc., but other research points to the increased likelihood of depression and other negative psychological outcomes that have also been reported alongside high levels of parentification (Hooper et al., 2012; Huq et al., 2016). Previous research shows that parentification of Latino children and adolescents challenges traditional Latino hierarchy, where parents are superior and children subordinate, and adds to the stress and negative feelings that surround language brokering (Huq et al., 2016). In Hooper et al.'s (2012) study of the differences in parentification for White and Black participants, ethnic belonging, in conjunction with parentification, was able to predict a higher percentage of depressive symptoms in Black Americans. For other ethnic minority communities, ethnic belonging or other aspects of ethnic identity alongside parentification might moderate positive or negative outcomes.

However, overall, parentification as a phenomenon affecting the Latino community is not very well studied, though the concept of role reversal is mentioned in multiple studies (Lazarevic, 2017; Love & Buriel, 2007; Martinez Jr. et al., 2009; Shen & Dennis, 2019; Umaña-Taylor et al., 2013; Weisskirch, 2005). Despite the mention of role reversal, the developing literature on the nature of parentification and its effects on minority groups has been largely inconclusive, as the populations most studied in relation to the subject are White, college-aged students (Hooper et al., 2012). Hence, this study aims to address this gap by assessing the prevalence of parentification among Latinos and its relationship to language brokering. Because parentification caused by language brokering challenges both traditional family dynamics of Latino immigrant families and, by challenging this dynamic, ethnic identity itself, parentification as a variable that contextualizes why language brokering might lead to increased family conflict is relevant to this study.

### **1.4. Goals of the Study**

This study will examine how the feelings surrounding language brokering of late adolescents may be tied to family conflict or other negative family relationships by way of a theoretical model moderated by participants' ethnic identity. In addition, it will assess parentification as a correlate of negative feelings about

language brokering. The research questions are the following: (1) Are feelings about language brokering in Latino adolescents related to family conflict and negative family relationships? (2) How does ethnic identity moderate the relationship between feelings about language brokering and family conflict? and (3) What is the relationship between language brokering and parentification?

**Hypothesis 1.** Negative feelings about language brokering, combined with low levels of ethnic identity, will be correlated with higher levels of family conflict and negative family relationships.

**Hypothesis 2.** Positive feelings about language brokering, combined with high levels of ethnic identity, will be correlated with lower levels of family conflict and positive family relationships.

**Hypothesis 3.** Negative feelings about language brokering, combined with high levels of ethnic identity, will be correlated to low levels of family conflict, presumably because the adolescents will see language brokering as their duty to the family even though it might make them uncomfortable or embarrassed.

**Hypothesis 4.** Positive feelings about language brokering, combined with low levels of ethnic identity, will be correlated to average or slightly elevated levels of family conflict due to what may be a lack of heritage culture socialization and disparate ethnic identities within the family.

**Hypothesis 5.** For individuals who have negative feelings about language brokering, negative feelings will be correlated with parentification.

## 2. Method

### 2.1. Participants

Participants for this study were recruited from Latino families in Central Texas, a state where Latinos will represent a plurality of the state's population by 2022 (Ura & Jin, 2019). The sample size was 384 participants between the ages of 15 and 18 ( $M = 16$ ,  $SD = 1$ ). Sample size was determined by using a sample size calculator (Sample Size Calculator, n.d.) with a confidence level of 95%, a confidence interval of 5, and the current estimated population of Latino youth in the U.S. In this study, a majority of the Latino participants identified as Mexican American ( $n=230$ ). This study employed purposeful sampling, and adolescents were eligible if they met the following characteristics: (1) they identified as ethnically Hispanic/Latino; (2) they were bilingual in English and Spanish; and (3) the adolescents translated for their parent(s).

## 2.2. Measures

Demographic information was obtained, including the gender, address, and nativity and generational statuses of the participants. Participants were given the choice to complete all measures in either English or Spanish. All scales were back-translated into Spanish if a Spanish version was not available. All participants chose to take the measures in English.

***Feelings About Language Brokering.*** For the purposes of this study, feelings about language brokering were measured with 3 of the 7 subscales from the Language Brokering Scale (Kim et al., 2014). These subscales were Adolescent Focused Burden, Negative Feelings, and Adolescent Independence. Each item was rated on a Likert Scale from 1-5, 1 being strongly disagree, 5 being strongly agree. Scores for the three subscales were summed together for a total score. Higher scores on all subscales were related to increased negative feelings about language brokering. To achieve this, scores for Adolescent Independence were reverse-coded, with lower scores representing increased positive feelings. Example statements for each subscale are “It is stressful to translate” (Adolescent-Focused Burden), “I am embarrassed to translate when my parent asks me to” (Negative Feelings), and “I feel competent and capable when I translate for my parent.” The Language Brokering Scale has been used with various ethnic groups in the past, including Latinos (Kim et al. 2014) and has shown adequate reliability across subscales, from Cronbach’s Alpha = .74-.77 (Love & Buriel, 2007). The scale has also demonstrated good construct validity, specifically metric factorial invariance and slope construct validity invariance (Kim et al., 2014.) In the current study, the mean score was 38.3 ( $SD = 13.69$ ).

***Family Conflict.*** Family Conflict was measured with the Asian American Family Conflicts Scale (FES). While the scale was specifically developed for Asian American adolescents and young adults, research on Latino adolescents has used this scale in the past to study family intergenerational conflict (Shen & Dennis, 2017). The measure is made up of 10 family situations, each appraised by its Likelihood and Seriousness. Each statement was rated on a 5-point Likert scale. For Likelihood, ratings ranged from “Almost never” to “Almost always,” while Seriousness was rated from “Not at all” to “Extremely.” Higher scores represent increased family conflict. In regard to the content of the 10 situations, examples include “Your parents want you to sacrifice personal interests for the sake of the family, but you feel this is unfair” and “Your parents tell you that a social life is not important at this age, but you think that it is.” The scale provided both convergent, discriminant, and construct validity (Lee et al., 2000). Cronbach’s Alpha, consistent with similar studies, was .91 (Shen & Dennis, 2019). The mean score was 55.4 with a standard deviation of 20.



***Ethnic Identity.*** Ethnic Identity was measured with the Ethnic Identity Scale, which assesses three components of ethnic-racial identity: affirmation, exploration, and resolution (Umaña-Taylor et al., 2004). All items for affirmation and one for exploration were reverse coded. Altogether, higher scores represent positive ethnic identity. For this scale, low ethnic identity is represented by a lower score while high ethnic identity is represented by a higher, more positive score. Example statements for each component are “My feelings about my ethnicity are mostly negative” (affirmation), “I have attended events that help me learn more about my identity” (exploration), and “I am clear about what my ethnicity means to me” (resolution). For this study, all subscales were used to obtain a sum score. The Ethnic Identity Scale’s items are rated on a 4-point Likert Scale from “Does not describe me at all” to “Describes me very well.” The scale has been used with diverse populations, including Latinos, and has consistently scored moderately strong alpha coefficients from .84 to .89 as well as high construct validity (Umaña-Taylor et al., 2004). The mean score for this scale was 50.7 with a standard deviation of 14.95.

***Parentification.*** To assess parentification, the Parentification Inventory was used (Hooper et al., 2011). Examples of statements include “I often help solve problems between my parent(s),” and “I really enjoy my role in my family.” The items are scored on a 5-point Likert Scale, with 1 being “Never true” and 5 being “Always true.” 5 items were reverse coded, so that higher sum scores represented higher levels of parentification, and lower sum scores represented lower levels of parentification. The Parentification Inventory has been used with diverse populations, including Latinos, and Cronbach’s alpha coefficients for the subscales range from .81 to .88 (Hooper & Wallace, 2009). The Inventory demonstrates high concurrent and construct validity (Hooper et al., 2011). The mean score was found to be 75.3 ( $SD = 18.72$ ).

### **2.3. Procedure**

Students were recruited from various high schools in and around the Central Texas area. A prepared email statement was sent by the researcher to high school administrations and teachers to be passed on to students in the 10th-12th grade who identified as Hispanic/Latino according to school records. The email statement included an overview of the study, including when and how the study would take place, through electronic survey. If the students were interested in the study, they replied to the email for more information or called the provided number of the researcher. After students demonstrated interest in the study, they were emailed consent forms in both English and Spanish. The consent and assent forms were signed by themselves and their parent(s) to ensure informed consent and assent for individuals under 18. Upon return of the signed consent form, participants were sent a link through email to an online survey, conducted using SurveyMonkey, where they completed the measures. After they completed the measures, a list of mental health resources was provided if participants were

interested in seeking further help. This study was approved by the Institutional Review Board.

### **3. Results**

#### **3.1. Analytic Strategy**

Three steps of analysis were conducted. The first step was to analyze the correlations between all the measures. The second step was a moderation analysis to assess whether feelings surrounding language brokering would predict family conflict depending on the adolescent's level of ethnic identity. There were four pathways to be analyzed: positive feelings about brokering and high ethnic identity (Pathway 1), positive feelings about language brokering and low ethnic identity (Pathway 2), negative feelings about language brokering and high ethnic identity (Pathway 3), and negative feelings about language brokering and low ethnic identity (Pathway 4). (See Figure 1.) A final and third step of analysis analyzed the correlation between negative feelings about language brokering and parentification.

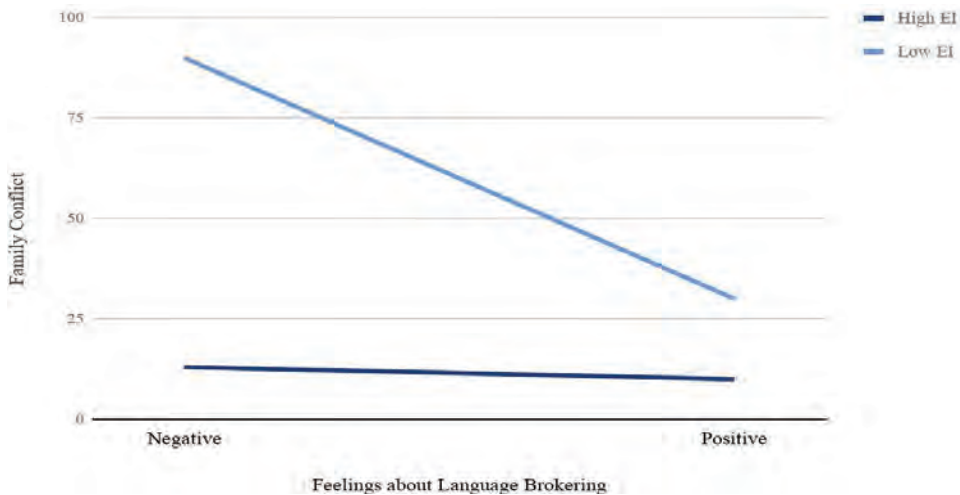
#### **3.2. Moderation Analyses**

To test the relationship between feelings about language brokering and their effect on family conflict using the moderators of high and low ethnic identity, a linear multiple regression analysis was run. Multiple regression analysis is a common method to establish moderation in an experiment (Fairchild & MacKinnon, 2008), in this case, how high and low ethnic identity influence how feelings about language brokering predict family conflict. The multiple regression analysis was completed in three steps. First, the independent variable (feelings about language brokering) was analyzed to test its predictability of the outcome variable (family conflict). Second, the moderator (ethnic identity) was analyzed to test its effect on the outcome variable. The final step was to assess whether the interaction between the independent variable, language brokering, and the moderator, ethnic identity, interacted to predict family conflict above and beyond step 1 and 2.

The findings revealed that the lowest levels of family conflict were found for individuals in pathway 1, positive feelings about language brokering and high ethnic identity. Similarly, low levels of family conflict were also found in pathway 3, negative feelings about language brokering and high ethnic identity. In regard to the pathways with higher levels of family conflict, individuals in pathway 4, where participants reported negative feelings about language brokering and low ethnic identity, reported the most family conflict. For individuals in pathway 2, where individuals reported positive feelings about language brokering and low ethnic identity, results show that participants reported both slightly elevated family conflict as well as higher and lower family conflict than hypothesized.

### 3.3. Parentification Correlation Analysis

To assess a correlation between negative feelings about language brokering and parentification, a correlational analysis was run. Findings indicated a positive correlation between negative feelings about language brokering and parentification. In other words, increased negative feelings about language brokering were correlated with higher levels of parentification ( $r = 0.8$   $p = .05$ ).



**Figure 1.** *The Effect of Feelings about Language Brokering on Family Conflict, moderated by Ethnic Identity (EI)*

## 4. Discussion

This study examined the relationship between feelings about language brokering and family conflict, assessing whether ethnic identity served as a moderator. Findings from the study add to extant literature on language brokering's impacts on the Latino family and establish a correlation between negative feelings about language brokering and parentification ( $p = .05$ ). In particular, findings indicate, as hypothesized, that individuals who had high ethnic identity and either positive or negative feelings about language brokering experienced the lowest amount of family conflict. Extant literature also supports these findings. In Weisskirch's (2005) study, individuals with positive feelings about language brokering also experienced positive, or high, ethnic identity. Indeed, positive feelings about language brokering seem to reinforce ethnic identity (Weisskirch, 2005). Furthermore, among first generation individuals, a demographic that often reports higher ethnic identity, language brokering is often seen as a duty to their family and, because of this, they enjoy the process of language brokering (Oznobishin & Kurman, 2016). High ethnic identity, therefore, seems to be a strong protective force against conflict within the family (Constante et al., 2019; Huq et al., 2016; Umaña-Taylor, 2013), and this might

be the reason individuals report less conflict within the family even though they have negative feelings about language brokering.

In contrast, as predicted in Hypothesis 1, negative feelings about language brokering for individuals with lower ethnic identity intensified family conflict, these individuals scoring the highest in family conflict. There is also support for this finding in the literature, where language brokering, especially in high frequencies, has been known to increase family conflict when the child and parent have different levels of cultural maintenance (Martinez et al., 2009; Shen & Dennis, 2017). These disparate ethnic identities enhance Latino adolescents' perception of language brokering as a burden (Kam & Lazarevic; 2014; Shen & Dennis, 2017).

However, while Hypothesis 4 predicted that positive feelings about language brokering paired with low ethnic identity would result in slightly elevated levels of family conflict, findings differed in the effects on family conflict. Specifically, while slightly elevated family conflict was the case for some participants, others reported either less or more family conflict. This departure from what was predicted in the hypothesis may be due to the smaller number of participants in pathway 2.

For adolescents who experienced lower levels of family conflict than hypothesized, these individuals' parents may also report lower ethnic identity, as conflict has been found in past studies to arise when individuals and their parents report differing ethnic identities (Shen & Dennis, 2017; Zhang et al., 2020). If both parents and adolescents report low ethnic identity, this may decrease conflict. For individuals who reported higher family conflict than hypothesized, this may mean they broker less, and the conflict that arises between themselves and their family is based solely on their lower ethnic identity, which may differ from that of their parents (Bostean & Gillespie, 2018). Future research that also examines the ethnic identity of parents in relation to their children would be needed to investigate this further.

The final step of analysis discovered a correlation between parentification and negative feelings about language brokering. To my knowledge, this study is the first of its kind to look at parentification as a correlate of feelings about language brokering. The correlation in this study was both significant and strong ( $r = 0.8$ ,  $p = 0.05$ ), supporting Hypothesis 5 and contextualizing negative feelings about language brokering. Findings suggest that parentification, or role reversal between the Latino adolescent and their parent, might be a predictor of feelings about language brokering, especially negative feelings. These findings provide a potential reason why pathway 4, or individuals who have both negative feelings about language brokering and low ethnic identity, is the most toxic. While previous research has shown ethnic and immigrant minorities like Latinos are often parentified (Hooper et al., 2012; Hooper et al., 2015; Titzmann, 2012), in Hooper et al.'s (2015) study, Latinos were found to experience positive benefit from parentification. The authors suggest collectivism embedded in Latino ethnic identity as the potential reason for these positive feelings, yet ethnic identity was not examined as a variable. Because of this, parentification, identified in this study as a correlate of negative feelings about language brokering, would be a good topic for future research to analyze whether its effects change from negative to positive based on ethnic identity being either low or high. In addition, future studies should explore how

components of parentification may be perceived differently based on the culture or group being studied. For Latinos, some aspects of parentification may be seen positively, as family obligation, while other aspects may be seen negatively, as boundary violations. This might be the reason why some Latinos report higher levels of parentification and either positive or negative outcomes.

#### **4.1. Limitations of the Study**

While the current study had many strengths, the study also had several limitations.

First, it was cross-sectional. A longitudinal study design would be able to draw the direction of hypothesized relations as well as describe findings causally. A longitudinal study would also be able to understand how feelings about language brokering change with age and developmental stage, as it has been suggested in previous studies that adolescent compared to young adult language brokers may report increased negative feelings about language brokering due to a lower sense of family obligation (Lazarevic, 2017). The second limitation of the study was that the measures used were exclusively self-reporting, which means that participants may have not either responded fully or truthfully to all measures. Finally, the current study was completed on a well-established Latino community in Central Texas. This means that data might not be generalizable to Latino populations in rural and/or developing Latino communities. In established Latino communities, the ability of Latinos to explore and partake in activities related to their ethnic identity may be greater. Future studies should examine the same relationships in less-established communities.

#### **4.2. Future Directions**

Future studies will benefit from recognizing the diversity of generational statuses among Latinos. In Umaña-Taylor et al.'s (2013) study, the strength of ethnic identity for Latinos was shown to play a protective role in Latino adolescence, but family ethnic socialization and ethnic identity were dependent on parents' nativity and/or generational statuses. Foreign-born parents were believed to be better equipped to socialize their children than those parents who were U.S. born (Umaña-Taylor et al., 2013). Further, adolescents of the first compared to the second generation have been shown in past research to report stronger ethnic identities and less family conflict as they are more able to empathize with their parents (Oznobishin & Kurman, 2016). If future studies examine nativity and/or generational status as one variable, researchers might be able to better understand the moderating role of ethnic identity. Moreover, future studies should examine the same relationships for other Latino subgroups as there is tremendous diversity within the Latino ethnic category.

Future studies should also examine the role that peer discrimination and acculturation play in contextualizing negative feelings about language brokering. Research in the past has found that discrimination related to adolescents' language brokering can worsen acculturation and associated negative behaviors (Huq et al., 2016; Kam et al., 2017). Likewise, higher levels of acculturation in adolescents have been found to make youth more vulnerable

to discrimination (Romero et al., 2018). That said, ethnic identification can attenuate these adverse effects (Kam et al., 2017). Further research would allow a better understanding of how discrimination and acculturation impact the relationship explored in this study.

Another direction for future studies would be to look at the congruence of parents' expectation that their child or adolescent translate and the child's fluency or ability to translate. Negative feelings about language brokering may arise when parents expect their child to have a greater ability to language broker than they actually do.

As for future studies regarding parentification, because the current study found a correlation between parentification and adolescents' negative feelings about language brokering, further exploration into how parentification predicts feelings or attitudes about language brokering is needed. Studies should take advantage of recent research that has identified different types of parentification in Latino families, including a recent study of mothers and their adolescent children that identified two distinct parentification profiles of adolescents. These were the reluctant confidant, one the mother confided in without the adolescent's voluntary reception of private information, and the deliberate confidant, where the adolescent voluntarily asked for their mother's private information (Shin 2019). While these parentification profiles were identified, more research is needed to determine how these distinct parentification themes affect adolescent internalizing outcomes and family dynamics.

### **4.3. Intervention and Policy Implications**

The findings in this study highlight the importance of making both more translators and interpreters available in public institutions as well as hiring more bilingual staff, especially bilingual doctors, since research shows youth may feel the most negatively about language brokering in medical settings (Kam et al., 2017).

Even more important is increasing bicultural education programs for Latinos, as biculturality may reduce negative feelings about language brokering and/or associated negative outcomes (Love & Buriel, 2007; Romero et al., 2018; Zhang et al., 2020). Indeed, interventions should be focused on increasing ethnic identity, as youth who strongly endorse their heritage culture experience less family conflict and negative feelings about language brokering (Constante et al., 2019; Huq et al, 2016 Kam et al., 2017; Umaña-Taylor et al., 2013; Weisskirch, 2005).

In conclusion, this study advances the field by offering new insights into the role ethnic identity plays in the relationship between feelings about language brokering and the corresponding effect on family conflict. The relationship where the most family conflict was found was for participants who had low ethnic identity and negative feelings about language brokering. These individuals also scored high on parentification. Future research should examine parentification and peer discrimination to contextualize these negative feelings.

The results of this study hold major policy implications that could improve the well-being of all Latino families.

## 5. References

- Bankston, C. L., & Zhou, M. (1995). Effects of Minority-Language Literacy on the Academic Achievement of Vietnamese Youths in New Orleans. *Sociology of Education*, 68(1), 1. <https://doi.org/10.2307/2112760>
- Bostean, G., & Gillespie, B. J. (2018). Acculturation, acculturative stressors, and family relationships among Latina/o immigrants. *Cultural Diversity and Ethnic Minority Psychology*, 24(1), 126–138. <https://doi.org/10.1037/cdp0000169>
- Constante, K., Cross, F.L., Medina, M. et al. (2020). Ethnic Socialization, Family Cohesion, and Ethnic Identity Development Over Time among Latinx Adolescents. *Journal of Youth and Adolescence*, 49, 895–906. <https://doi.org/10.1007/s10964-019-01139-3>
- Fairchild, A. J., & MacKinnon, D. P. (2008). A General Model for Testing Mediation and Moderation Effects. *Prevention Science*, 10(2), 87–99. <https://doi.org/10.1007/s11121-008-0109-6>
- Fok, C. C. T., Allen, J., Henry, D., & Team, P. A. (2011). The Brief Family Relationship Scale. *Assessment*, 21(1), 67–72. <https://doi.org/10.1177/1073191111425856>
- Hooper, L. M., & Wallace, S. A. (2009). Evaluating the Parentification Questionnaire: Psychometric Properties and Psychopathology Correlates. *Contemporary Family Therapy*, 32(1), 52–68. <https://doi.org/10.1007/s10591-009-9103-9>
- Hooper, L., Doehler, K., Wallace, S., & Hannah, N. (2011). The Parentification Inventory: Development, Validation, and Cross-Validation. *The American Journal of Family Therapy*, 39(3), 226–241. <https://doi.org/10.1080/01926187.2010.531652>
- Hooper, L. M., Tomek, S., Bond, J. M., & Reif, M. S. (2015). Race/Ethnicity, Gender, Parentification, and Psychological Functioning: Comparisons Among a Nationwide University Sample. *The Family Journal: Counseling and Therapy for Couples and Families*, 23(1), 33–48. <https://doi.org/10.1177/1066480714547187>
- Hooper, L. M., Wallace, S. A., Doehler, K., & Dantzler, J. (2012). Parentification, Ethnic Identity, and Psychological Health in Black and White American College Students: Implications of Family-of-Origin and Cultural Factors. *Journal of Comparative Family Studies*, 43(6), 811–835. <https://doi.org/10.3138/jcfs.43.6.811>
- Hooper, L. M., Tomek, S., Bond, J. M., & Reif, M. S. (2015). Race/Ethnicity, Gender, Parentification, and Psychological Functioning: Comparisons Among a Nationwide University Sample. *The Family Journal: Counseling and Therapy for Couples and Families*, 23(1), 33–48. <https://doi.org/10.1177/1066480714547187>
- Hooper, L. M., & Wallace, S. A. (2009). Evaluating the Parentification Questionnaire: Psychometric Properties and Psychopathology Correlates. *Contemporary Family Therapy*, 32(1), 52–68. <https://doi.org/10.1007/s10591-009-9103-9>
- Huq, N., Stein, G. L., & Gonzalez, L. M. (2016). Acculturation conflict among Latino youth: Discrimination, ethnic identity, and depressive symptoms. *Cultural Diversity and Ethnic Minority Psychology*, 22(3), 377–385.

- <https://doi.org/10.1037/cdp0000070>
- Kam, J. A., & Lazarevic, V. (2014). The Stressful (and Not So Stressful) Nature of Language Brokering: Identifying When Brokering Functions as a Cultural Stressor for Latino Immigrant Children in Early Adolescence. *Journal of Youth and Adolescence*, 43(12), 1994–2011. <https://doi.org/10.1007/s10964-013-0061-z>
- Kam, J. A., Marcoulides, K. M., & Merolla, A. J. (2017). Using an Acculturation-Stress-Resilience Framework to Explore Latent Profiles of Latina/o Language Brokers. *Journal of Research on Adolescence*, 27(4), 842–861. <https://doi.org/10.1111/jora.12318>
- Kim, S. Y., Wang, Y., Weaver, S. R., Shen, Y., Wu-Seibold, N., & Liu, C. H. (2014). Measurement equivalence of the language-brokering scale for Chinese American adolescents and their parents. *Journal of Family Psychology*, 28(2), 180–192. <https://doi.org/10.1037/a0036030>
- Krogstad, J. M., Stepler, R., & Lopez, M. H. (2015, May 12). English Proficiency on the Rise Among Latinos. *Pew Research Center's Hispanic Trends Project*. <https://www.pewresearch.org/hispanic/2015/05/12/english-proficiency-on-the-rise-among-latinos/>
- Lee, R. M., Choe, J., Kim, G., & Ngo, V. (2000). Construction of the Asian American Family Conflicts Scale. *Journal of Counseling Psychology*, 47(2), 211–222. <https://doi.org/10.1037/0022-0167.47.2.211>
- Lazarevic, V. (2017). Effects of cultural brokering on individual wellbeing and family dynamics among immigrant youth. *Journal of Adolescence*, 55, 77–87. <https://doi.org/10.1016/j.adolescence.2016.12.010>
- Love, J. A., & Buriel, R. (2007). Language Brokering, Autonomy, Parent-Child Bonding, Biculturalism, and Depression. *Hispanic Journal of Behavioral Sciences*, 29(4), 472–491. <https://doi.org/10.1177/0739986307307229>
- Machado, A. (2014, May 7). Why Many Latinos Dread Going to the Doctor. *The Atlantic*. <https://www.theatlantic.com/health/archive/2014/05/why-many-latinos-dread-going-to-the-doctor/361547/>
- Martinez, C. R., McClure, H. H., & Eddy, J. M. (2008). Language Brokering Contexts and Behavioral and Emotional Adjustment Among Latino Parents and Adolescents. *The Journal of Early Adolescence*, 29(1), 71–98. <https://doi.org/10.1177/0272431608324477>
- Oznobishin, O., & Kurman, J. (2016). Family obligations and individuation among immigrant youth: Do generational status and age at immigration matter? *Journal of Adolescence*, 51, 103–113. <https://doi.org/10.1016/j.adolescence.2016.05.005>
- Phinney, J. S., Romero, I., Nava, M., & Huang, D. (2001). The Role of Language, Parents, and Peers in Ethnic Identity Among Adolescents in Immigrant Families. *Journal of Youth and Adolescence*, 30(2), 135–153. <https://doi.org/10.1023/a:1010389607319>
- Romero, A. J., Piña-Watson, B., & Toomey, R. B. (2018). When is bicultural stress associated with loss of hope and depressive symptoms? Variation by ethnic identity status among Mexican descent youth. *Journal of Latina/o Psychology*, 6(1), 49–63. <https://doi.org/10.1037/lat0000078>
- Sample Size Calculator - Confidence Level, Confidence Interval, Sample Size, Population Size, Relevant Population - *Creative Research Systems*. (n.d.). [Surveysystem.com. https://www.surveysystem.com/sscalc.htm](https://www.surveysystem.com/sscalc.htm)



- Shen, J. J., & Dennis, J. M. (2017). The family context of language brokering among Latino/a young adults. *Journal of Social and Personal Relationships, 36*(1), 131–152. <https://doi.org/10.1177/0265407517721379>
- Shin, Y. (2019). How do mothers manage their privacy with adolescents? Exploring mother–adolescent communication in Mexican immigrant families. *Journal of Social and Personal Relationships, 36*(11–12), 3733–3751. <https://doi.org/10.1177/0265407519836282>
- Titzmann, P. F. (2012). Growing Up Too Soon? Parentification Among Immigrant and Native Adolescents in Germany. *Journal of Youth and Adolescence, 41*(7), 880–893. <https://doi.org/10.1007/s10964-011-9711-1>
- Umaña-Taylor, A. J., Yazedjian, A. & Bámaca-Gómez, M. Y. (2004). Developing the Ethnic Identity Scale using Eriksonian and social identity perspectives. *Identity: An International Journal of Theory and Research, 4*, 9–38. [https://doi.org/10.1207/S1532706XIDO401\\_2](https://doi.org/10.1207/S1532706XIDO401_2)
- Umaña-Taylor, A. J., Zeiders, K. H., & Updegraff, K. A. (2013). Family Ethnic Socialization and Ethnic Identity: A Family-Driven, Youth-Driven, or Reciprocal Process? *Journal of Family Psychology: JFP: Journal of the Division of Family Psychology of the American Psychological Association (Division 43), 27*(1), 137–146. <https://doi.org/10.1037/a0031105>
- Ura, A., & Jin, C. H. (2019, June 20). *Texas gained almost nine Hispanic residents for every additional white resident last year*. The Texas Tribune; Texas Tribune. <https://www.texastribune.org/2019/06/20/texas-hispanic-population-pace-surpass-white-residents>
- Weisskirch, R. S. (2005). The Relationship of Language Brokering to Ethnic Identity for Latino Early Adolescents. *Hispanic Journal of Behavioral Sciences, 27*(3), 286–299. <https://doi.org/10.1177/0739986305277931>
- Zhang, M., Kim, S.Y., Hou, Y. et al. Parent–Adolescent Acculturation Profiles and Adolescent Language Brokering Experiences in Mexican Immigrant Families. *Journal of Youth and Adolescents, 49*, 335–351 (2020). <https://doi.org/10.1007/s10964-019-01064-5>



# Coastal Plastic Management as a Response to Rafting Invasive Species

---

*Jiaying Liu*

**Author Background:** *Jiaying Liu grew up in China and currently attends Indian Springs School, in Birmingham, Alabama in the United States. Her Pioneer research concentration was in the field of environmental studies and titled “Our Planet’s Plastic Plight.”*

## Abstract

Plastic in the ocean is one of the largest environmental concerns in today’s world. Marine plastic is known to pose various threats to marine life. In recent years, there has been an increase in discussion about the role of floating plastic debris in transporting invasive species. However, there has not been sufficient investigation and public awareness of this risk to yield viable response mechanisms. This study reviewed nine studies of various locations that assessed biological invasion risks by sampling marine debris with fouling biota, as well as the coastal plastic management efforts of Australia, New Zealand, Hawaii (U.S.), and California (U.S.). This study found that a lack of consistent methodology across the sampling studies is hindering the risk evaluation of biological invasion via marine plastic at the target areas, and that countries and regions with high public awareness of plastic pollution have an unexplored potential to leverage citizen engagement in response to this risk. Based on these results, this study proposes that a consistent methodology for sampling studies on fouling biota needs to be developed, and that measures should be taken to involve the public in gathering data on fouling biota.

## 1. Introduction

The ubiquity of plastic products and the mismanagement of plastic waste has resulted in large amounts of plastic entering the ocean. It is estimated that between 4.8 and 12.7 million metric tons of plastic are released into the ocean each year, while approximately 5.25 trillion plastic particles (268,940 metric tons) are floating on the ocean surface (Jambeck et al. 2015, Erikson et al. 2014).

Marine plastic debris can have land-based origins – for example, litter from beach recreational activities – and ocean-based origins, such as detachment of fishing gear (Li et al. 2016). With the staggering amount of plastic and the multitude of origins, the accumulation of marine plastic has become a global issue without a clear solution.

Marine plastic poses various threats to individual marine organisms as well as marine ecosystems. Marine life's entanglement in plastic and ingestion of plastic are the most publicly known harms. Following the 2011 Japan earthquake and tsunami, the potential for marine plastic debris to become a pathway for biological invasion has gained increasing attention. The tsunami released millions of anthropogenic objects, of sizes ranging from small plastic fragments to large floating docks, which drifted across the Pacific Ocean and landed on coastlines from Hawaii to central Alaska (Carlton et al. 2017). Much of the debris arrived carrying living communities of invertebrate and fish species indigenous to Japan (Carlton et al. 2017). A study sampled 48 species rafting on Japanese tsunami marine debris (JTMD) and found that 13 species have highly matching habitats on their landing location, indicating high probabilities for establishment (Simkanin et al. 2019). A survey on 30 remote islands for colonization by marine organisms on floating debris suggests that anthropogenic debris more than doubles the dispersal opportunities for marine fauna (Barnes 2002).

The establishment of biota on floating debris is commonly referred to as “biofouling” or “rafting.” The established biota is commonly referred to as “hitchhiker” or “rafter.” Species not native to an area are referred to as “alien,” “non-indigenous (NIS),” or “invasive” species. The methods of transporting invasive species are referred to as “vectors” for biological invasion. Transportation of NIS on anthropogenic marine debris differs from common invasive vectors (e.g. ballast water and ship hulls) in having significantly lower moving speed and a one-way direction (NOAA 2017, Carlton et al. 2017). Compared to biodegradable rafts such as trees, anthropogenic marine debris, particularly plastic, endures a much longer period of drifting, which facilitates the establishment and development of biological communities (Carlton et al. 2017). Plastic debris is often positively buoyant in seawater and able to travel long distances as a result of both wind and ocean currents (Duhec et al. 2015). The studies on JTMD have shown that plastic debris could facilitate species invasion events that could not be achieved through the commonly known vectors. With the ubiquity of marine plastic, this potential vector is extremely difficult to monitor and control.

Biological invasions resulting from marine plastic could lead to the displacement of native species and cause severe, irreversible disruption to native ecosystems. Tremendous social and economic costs may also result from biological invasions. Effective management of biological invasion risks involves preventative measures, early detection of non-indigenous species, and rapid response for eradication of the invader (Simberloff et al. 2013). Some marine biological invasion vectors, such as ballast water, have been well studied and the research has led to a series of international negotiations and policy enactments (International Maritime Organization n.d.). However, international agreements specifically targeting marine plastic debris have only developed fairly recently and have not sufficiently addressed the release of plastic into the ocean and the transport of plastic debris across state territories (Wu 2020). The large-scale

patterns of community establishment on marine plastic and the drifting trajectories involved are still largely unknown (Smithsonian Institution 2019). The studies of non-indigenous rafting-species on floating marine debris have been sporadic, often conducted by individual research efforts with concerns for potential biological invasions at a specific area.

Furthermore, these studies demonstrate great heterogeneity in methodology. The development of systematic evaluation for biological invasion risks induced by marine debris is still at its early stage. Sampling of debris is the foundational step of the research, and is particularly important in cases where direct observation of an alien species is lacking. This process may involve the collection of floating debris, stranded debris, or both (Rech et al. 2018b, Ivkić et al. 2019, Campbell et al. 2017, Rech et al. 2018a). Plastic debris often accounts for the vast majority (above 70%, in some cases above 90%) of anthropogenic marine debris samples (Rech et al. 2018a, Campbell et al. 2017, Rech et al. 2018b, Shabani et al. 2019). In some studies, identification of all biofouling organisms is performed, and non-indigenous species are determined (Rech et al. 2018a, Rech et al. 2018b, Shabani et al. 2019, Carlton et al. 2017). The surveys on the taxonomy of biofouling organisms vary in ranges, and they are dependent on the purpose of individual studies. For example, Shabani et al. (2019) only identified rafting macroorganisms in the north coast of the Persian Gulf and did not account for microorganisms. Some studies also identified the origins of non-indigenous species. Sometimes predictions are made on the transportation pathways of these species based on ocean current patterns and the drifting dynamics of the substrates (Shabani et al. 2019, Campbell et al. 2017).

The environmental impacts of marine plastic have gained wide public awareness and attracted public involvement in reducing coastal plastic pollution, particularly in economically developed regions. Large scale beach cleanup and beach survey programs have led to the removal of significant amounts of litter from beaches and, sometimes, open water. For example, Sustainable Coastlines Hawaii has organized beach cleanup events since 2011, attracting 30,875 volunteers and collecting 519,286 pounds of litter. While conducting sample studies along long stretches of coastline can be time-consuming and costly, these cleanup activities often cover significant lengths of coastline. The high levels of public engagement may be a potential resource for data collection on the fouling status of marine debris and resolve the opportunistic nature of existing debris sampling studies for biological invasion.

While the risks of marine plastic as a vector of biological invasion have been recognized by scientists, sampling studies have been sporadic, and effective management efforts to prevent potential invasions are lacking. This study aims to (1) review existing sampling studies that evaluate the risk of biological invasion via marine plastic, (2) review coastal plastic management efforts of regions with high public activism against plastic pollution, and (3) provide suggestions for leveraging public engagement in response to the risks of marine plastic as a biological invasion vector.

## 2. Background

### 2.1 Impact of biological invasions

Invasive species are regarded as one of the major threats to biodiversity (Molnar et al. 2008). For example, Zebra mussels, originating in southern Russia, were transported to North America, likely by ship hulls or ballast water. By 2000, the species had spread across the eastern U.S. and Canada, smothering native mussel species, decreasing phytoplankton densities, and affecting a variety of plankton-feeding invertebrates (Simberloff 2013). Biological invasion can also have tremendous economic costs. For example, introduced invertebrate species costs Australian agriculture approximately US\$420 million per year (Pimentel 2011).

Lag times often exist between the first emergence of an invasive species and its large-scale establishment. A species could maintain long periods of relative inactivity followed by drastic changes in its dynamics (Crooks 2005). It is extremely difficult to predict the growth of an invasive population based on the status quo, so even small-scale establishments of non-indigenous species should be addressed with strong precautions. While the eradication of a small-scale invasive population could be successful, such as the removal of fewer than 20 mature bivalves *P. caniculus* from South Australia, the removal of a well-established population is almost impossible (Thresher & Kuris 2004). Due to the high environmental and economic costs invasive species may entail, as well as the irreversibility of species establishment, early intervention is crucial in managing biological invasion risks.

### 2.2 Interaction between marine debris and rafting biota

Marine biota can attach to an immobile structure before the solid detaches and becomes marine debris. Biota can also establish on floating debris. Microorganisms often form a biofilm on floating debris within hours of its submergence, and invertebrate communities can be established within three to four weeks (Kiessling et al. 2015). The diversity and size of biofouling communities are affected by the size, rugosity, buoyancy, and stability of their substrates (Thiel & Gutow 2005, Goldstein et al. 2014). Most biofouling organisms on floating debris are sessile or semi-sessile, which allows the organism to firmly attach itself to the debris (Kiessling et al. 2015). Suspension feeding (feeding on substances suspended in water) accounts for the feeding method of 72 % of rafting taxa (Kiessling et al. 2015). Some species, such as *Lepas pectinata* and *Dosima fascicularis* develop adaptation in response to the physical characteristics of their substrate (Whitehead et al. 2011).

Debris with low floating stability, such as Styrofoam, may shift positions frequently in water, causing sides to alternate between exposure to water and the air, and lead to high mortality in fouling biota (Bravo et al. 2011). Hard plastic tends to have better stability and more consistent biota colonization. A sample study in the North Pacific shows that the diversity of the biofouling community demonstrates a positive relationship with debris size (Goldstein et al. 2014). The weight of the attached biota can also influence the buoyancy of the debris,

causing it to sink, and then resurface if the rafting organisms die and detach from the debris (Kiessling et al. 2015).

### **2.3 Significance of Japan tsunami marine debris**

The debris released into the ocean from the 2011 Japan earthquake and tsunami has captured extensive media coverage. Rafting biota is only one of numerous reasons that JTMD gained large public attention. Large floating debris, such as derelict fishing boats, could be hazards to navigation. Radiation was also a major concern when the debris first entered the ocean. The enormous human costs of the disaster also gave the debris significant emotional value. The arrival of large debris on the east Pacific shores has often been followed by extensive response efforts coordinating various agencies, government and academic institutions. For instance, the removal of the Misawa dock from the Washington Coast followed a “Whole of Government” approach. Extensive personnel were involved in discrete steps of tracking, approaching, assessing, cleaning, and removing the dock (Barnea et al. 2013).

Carlton et al. (2017) documented at least 289 invertebrate and fish species that survived up to six years of drifting across the Pacific Ocean on JTMD. The report demonstrated the ability of shallow-water species to disperse transoceanically on durable, artificial substrates. The report is followed by a series of studies on the invasion risks of JTMD rafting organisms.

## **3. Review of sampling studies on fouling biota**

For this study, various academic databases were searched using the following keywords: marine plastic, anthropogenic marine debris, marine litter, rafting, biofouling, hitch-hiking, non-indigenous species, and biological invasion. Studies that 1) involve beach or ocean surface debris sampling and 2) have fouling biota assessment as the main objective were selected. Some reports of individual sightings of NIS rafting on anthropogenic marine debris appeared in the search, such as the discovery of a non-indigenous coral colony (dead) on a piece of polyurethane foam at Normandy, France (Hoeksema et al. 2018). However, these reports do not fit criterion 1 and are not considered in this study. Using these keywords and applying these criteria yielded only nine studies (as of July 2, 2020). Each of the studies was assigned a random integer for the convenience of referencing in this paper (Table 1).

The methods and results of these studies are thoroughly examined, and a list of essential parameters is developed to describe the methods and results of each study. The parameters are categorized into three parts: “sampling methods,” “sampling results: debris,” and “sampling results: rafting biota.” Each parameter was filled by data explicitly given in the literature (e.g. number of sampling sites), simple calculations based on data from the literature (e.g. deriving the percentage of total plastic by adding up the percentages of all plastic types), or marked as “not given (NG)” or “not applicable (NA).” As this study expected the data types across different literatures to vary significantly, and that there may be a significant amount of “not given” or “not applicable” data for each

parameter, it will not include numerical analysis. The study will qualitatively evaluate each parameter to find a general set of practices in evaluating biological invasion risk by marine debris sampling, and to identify areas where more consistency across studies is needed to achieve comparable results.

**Table 1.** *Study locations*

Literature by study number: (1) Shabani et al., 2019 (2) Rech et al., 2018a (3) Campbell et al. 2017 (4) Rech et al. 2018b (5) Carlton et al. 2017 (6) Goldstein et al. 2014 (7) Barnes & Milner 2005 (8) Gracia et al. 2018 (9) Miralles et al. 2018

Study No.	1	2	3	4	5
Location	Northern coast of Persian Gulf	Coast of Asturias, Bay of Biscay, Spain	Western Coromandel Peninsula, New Zealand	Lagoon of Venice, Italy & Algarve region, Portugal	Midway Atoll - Hawai'i Island & south central Alaska - central California
Study No.	6	7	8	9	
Location	Eastern & Western North Pacific Gyre	The Atlantic from 68°S-78°N	Atlantico Department, Carribean Coast of Colombia	Coast of Asturias, Bay of Biscay, Spain	

As shown in Table 2, out of a total of nine selected studies, eight involve solely stranded debris collection. Five studies collected both anthropogenic and natural debris; four studies collected fouled debris only. The minimum size of collected debris is only specified by two studies, which reported 1.5 centimeters and 5 centimeters. Excluding Study 4, which focuses only on three locations, the number of sampling sites ranged from 15 to 27. The length of coastline covered by sampling activity is provided by only three studies. The size of each sampling site varies. The sampling area may be near the shoreline, or it may stretch into the backshore area. Six studies suggest ethanol as a preservative for collected biota, although solution concentrations vary.

**Table 2: Sampling methods**

Study No.	1	2	3	4	5	6	7	8	9
Origin of debris	Stranded	Stranded	Stranded	Stranded	Stranded	Floating	Stranded	Stranded	Stranded
Types of debris	Anthropogenic + natural, fouled only	Anthropogenic, fouled only (survey 1)*	Anthropogenic	Anthropogenic, fouled only	Anthropogenic + natural, JTMD suspects, fouled only**	Anthropogenic + natural	Anthropogenic and natural (kelp is also collected)	Anthropogenic + natural	Plastic bottles = ALDFG***
Size of debris	Diameter $\geq$ 5cm	Size >1.5cm (survey 2)	Visible	NG	Length from <1m to >13m	NG	Visible	NG	NG
Number of sampling sites	16	15	27	3	NA	NA	16	26	22
Length of coastline represented	~600km	190km linear distance	NG	NG	NG	NA	NG	72km	NG
Area of sampling	100m*5m/site	"Whole area" (survey 1); 3m*3m*4*2/site (survey 2)	.50m*2m/site	8000m <sup>2</sup> ; 7000m <sup>2</sup> ; 2800m <sup>2</sup>	NA	NA	200m*4m/site	100m*50m/site	Entire beach, vegetation line to waterline
Preservation of debris	NG	80% ethanol	70% ethanol	96% ethanol	NG	5% formalin buffered with sodium borate or 95% ethanol	NG	70% ethanol or dry preserved	Ethanol

NA = Not applicable NG = Not given

\*The study involves two independent surveys at different areas of each selected site, sampling only fouled litter (survey 1) and all stranded litter (survey 2), respectively.

\*\*The majority of debris was not directly collected by researchers, but acquired from other debris-collecting efforts.

\*\*\*Abandoned, Lost or otherwise Discarded Fishing Gear.

For floating debris sampling and collection of JTMD, the total amount of sea-based litter is assumed to be 100%. Table 3 shows that, among seven remaining studies, only two provided estimations of the abundance of sea-based litter among all beach litter. It is worth noting that "sea-based" is defined by the selected literature as being released from sea-based activities, such as fishing and sailing. Surveys focusing on fouled debris only show a moderate count of items collected (22-132 items), while surveys involving both fouled and unfouled litter involve as many as 7,597 items for one study. Among studies with applicable data for "Count of all collected debris" and "Number of objects with fouling organisms," the proportion of fouled debris is relatively small, ranging from  $\approx$  20% in Study 3 to below 2% in Study 9. Studies with applicable data for "Abundance of plastic in sampled areas (by count)" show that plastic is the most ubiquitous debris throughout the study, with the exception of Study 8, in which the most ubiquitous debris is woody debris. Among the five studies with applicable data for "Abundance of plastic among fouled debris (by count),"



plastic is the single most fouled type of debris, with proportions above 90% in four studies.

**Table 3.** *Sampling results: debris*

Study No.	1	2	3	4	5	6	7	8	9
Abundance of sea-based litter	NG	NG*	39%	64%	Assume 100%	Assume 100%	NG	NG	NG
Count of all collected debris	132	94 (survey 1)	393	22	NG	NG	3200	7597	>1677
Number of objects with fouling organisms	132	94	~80	22	634	242	~259	NG	26
Abundance of plastic in sampled areas (by count)**	NA	75%-100% on 14 sites, <50% on 1 site (survey 2)	74%	NA	NA	NA	60.2% ± 6.33	30%	100% among bottles, a major material for ALDFG
Abundance of plastic among fouled debris (by count)	73%	98% ± 6	~96%	91%	NG**	≥90%	NG	NG	NG

NA = Not applicable NG = Not given

\*76% of debris (survey 2) has unidentified origins. Thus the sample size of debris with identifiable origins is too small to be representative of the true abundance of sea-based debris.

\*\*Debris was not classified by material, but by function (e.g. dock, vessel etc.).

\*\*\*In this study, polystyrene foam is considered a part of the “plastic” category.

As illustrated in Table 4, seven studies only considered macroscopic biota; Study 5 identified macroinvertebrate and protist taxa, but could not adequately assess them due to poor preservation conditions. Four studies employed genetic barcoding in identifying rafting fauna, while two studies relied on visual identification. The taxonomic levels for identification varied among studies. It is common that, especially for stranded debris, rafting biota are already dead or significantly degraded, so identification to species level is not possible (Rech et al. 2018a, Miralles et al. 2018). Seven studies identified at least one NIS.

**Table 4.** *Sampling results: rafting biota*

Study No.	1	2	3	4	5	6	7	8	9
Types of biota screened	Macroscopic	Macroscopic	Macroscopic	Macroscopic	Macroscopic & microscopic	Macroscopic & microscopic	Macroscopic	Macroscopic	Macroscopic
Taxonomic identification method	Visual	Genetic barcoding	NG	Genetic barcoding + visual	Genetic analysis	Visual	NG	NG	Genetic barcoding
Taxonomic abundance*	21 taxa	23 species	~7 phyla	30 taxa	NG	95 taxa	NG	≥3 phyla	17 species
Number of NIS	6	5	1	8	289**	≥1	NG	NG	8

NA = Not applicable NG = Not given

\*Studies vary among the lowest taxonomic level of identifying rafting fauna, this count uses the level that is reflective of all fauna in a study.

\*\*Species arriving from Japan, but not necessarily non-indigenous in the landing area.

## 4. Review of coastal plastic management efforts

### 4.1 Case selection

This portion of the study aims to review the status quo of response to marine plastic as a vector for biological invasion in countries or regions that exhibit high levels of public awareness and public engagement against coastal litter pollution. In order to be selected for examination, a country or region had to meet the following requirements: 1) there is at least one beach cleanup or survey activity that applies for the entire country or region, that relies heavily on citizen participants, and is well-coordinated by governmental or non-governmental organizations; 2) the website and reports of major government sectors and non-governmental organizations should be available in English, in order to avoid inaccuracies in translation. On this basis, the final list of regions was narrowed to Australia, New Zealand, Hawaii (U.S.), and California (U.S). Although Hawaii and California are both U.S. states, they are discussed separately due to significant differences in location and plastic management approach. Each region is evaluated under the following subcategories: recognition of marine plastic as a biological invasion vector, effectiveness of citizen-participation coastal surveys in detecting and reducing biological invasion risk, and effectiveness of citizen-participation coastal cleanups in detecting and reducing biological invasion risk.

#### 4.2.1 Overview of all selected regions

While Australia, Hawaii, and California explicitly recognize marine plastic (or marine debris, by broader definition) as a vector for invasive species, only Hawaii provides concrete steps for citizens to respond to marine debris with fouling organisms (Table 5). However, Hawaii is also the only region that lacks a

coordinated beach survey program that allows citizens to contribute to the coastal debris database. Beach survey programs available in Australia, New Zealand, and California allow citizens to select a preferred site to collect and record the type, size, and quantity of coastal debris (Table 6). However, data on biofouling organisms is not collected in any of these programs.

**Table 5.** *Overview of response towards marine plastic as a biological invasion vector*

Country/Region	Australia	New Zealand	Hawaii	California
Mentioned NIS via plastic as a threat	Yes	No	Yes	Yes
Encourages the public to report NIS on marine debris	No	No	Yes	No
Has beach surveys for citizen participation	Yes	Yes	No	Yes

**Table 6.** *Overview of beach survey methods*

Country/region	Australia	New Zealand	California
Sampling area	3-6 transects with length from water edge to 2m into backshore vegetation and width of = 2m	One 100m*20m area	Accumulation: 100m along the shoreline, from water edge to backshore Standing stock: four 5m segments randomly selected from an 100m shoreline
Data collected	Debris type, color, and size class for the first debris in every 1/10 of transect	Count and weight for each litter category; confidence level (high or low) for weight; visual assessment grade (A-D) for plastic resin pellets; separate assessment for large items regardless of stranded or floating condition.	Standing stock: Debris density by count, debris type Accumulation: Debris deposition rate (by count), debris type, debris weight.
Minimum size for audit	NG	5mm	2.5cm
Requires repeated survey	No	Repeat every 3 months	Yes
Requires volunteers to report biofouling status of debris	No	No	No
Provides alternative opportunities for younger age groups	Yes	No	Yes

As summarized in Table 7, the majority of beach cleanup projects were founded in the past decade. All of them demonstrate high levels of success in attracting public engagement and accomplishing large-scale waste removal. The majority of cleanups rely on volunteers to record data from collected litter, such as weight, debris type, debris count etc. Although it is unknown whether the response to debris with biofouling objects is included in volunteer training, no relevant guidelines are mentioned in any of these programs' websites.

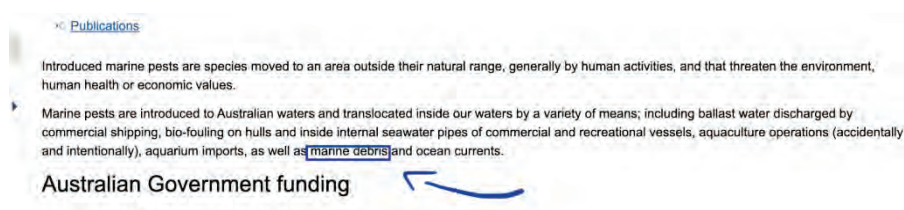
**Table 7.** Summary of nationwide and region-wide cleanup activities

Organization Name	Tangaroa Blue	Sea Cleaners	Sustainable Coastlines	808 Cleanups	Sustainable Coastlines	Coastal Cleanup Day
Country/region	Australia	New Zealand	New Zealand	Hawaii	Hawaii	California
Year founded	2004	2012	2009	2015	2011	1985
Debris collected	1,305 tons (16,115,247 items)	8.8 million liters (over 85 million pieces)	219,747 kg (1,595,651 L)	626,919 lbs (all categories, including non-marine litter)	519,286 lbs	24,663,820 lbs
Number of volunteers	168,684	NG	NG	30,871 (all categories)	30,875	1,572,328
Number of cleanup sites	3,557	NG	NG	NG	NG	1,868.5 miles of waterway
Number of cleanups	19,964	NG	1,403	1,095 group events, 5,857 adopt-a-site cleanups (all categories)	147	NA
Cleanup participation hours	436,120	Over 160,000	178,991	NG	NG	NG
Collects data from debris	Yes	NG	Yes	Yes	NG	Yes
Asks volunteer to note fouling biota	No	No	No	No	No	No

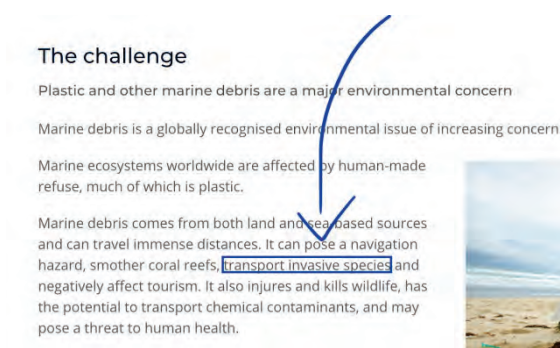
#### 4.2.2 Australia

The Department of Agriculture, Water and the Environment of Australian Government briefly mentioned marine debris as a vector for marine pest dispersal under the "Marine pest" section of its website (Figure 1). The Commonwealth Scientific and Industrial Research Organisation (CSIRO) briefly mentioned this vector on the frontpage of its website regarding marine litter:

(Figure 2). However, rafting species does not appear to be a focus in marine plastic management. In 2018, a threat abatement plan was developed to address the threat of marine debris (mainly plastic) to native species and ecological communities. Interestingly, although the plan does not mention marine plastic as an invasion vector, it proposes the potential for rafting microorganisms to graze on and degrade plastic (Commonwealth of Australia 2018). The CSIRO Marine Debris Team conducted an extensive survey at 175 sites across the entire Australian coastline between October 2011 and May 2013. Seventy-five percent of total debris was reported to be plastic, and fouling biota is not mentioned in the study (Hardesty et al. 2017). Entanglement and ingestion are still the dominant problems put forward as rationales for Australia's coastal plastic management efforts.



**Figure 1**



**Figure 2**

Citizens can volunteer in marine litter surveys and contribute to the National Marine Debris Database. For each survey, beach litter in three to six two-meter-wide transects from the shoreline to two meters into backshore vegetation is collected. Type and color of each item is documented, while the size class of the first item encountered in every 1/10 of the transect length is recorded

(National Marine Debris Database n.d.). A modified survey activity, Emu Parade, allows students and young adults to engage in such debris surveys (National Marine Debris Database n.d.).

Tangaroa Blue operates the Australian Marine Debris Initiative, which is a network of volunteers that have collected 1,305 tons of debris since 2004 (Tangaroa Blue n.d.). Tangaroa Blue and other partners contribute cleanup data to the Australian Marine Debris Database. Among all debris data entered since January 2005, plastic is the single most common debris type, and 51% of debris comes from the sea (Tangaroa Blue n.d.). There are no records of fouling biota in this database. Other regional cleanup organizations include Eco Barge Clean Seas Inc. in the Whitsunday region (cleaning 203,248 kg of litter since 2009), and South West Marine Debris Cleanup, serving the Tasmania region (Eco Barge Clean Seas Inc. n.d., South West Marine Debris Cleanup n.d.).

#### **4.2.3 New Zealand**

As a country in which 65% of the population live within five kilometers of the coast, the marine environment plays a significant role in the lifestyle of its citizens (Ministry for the Environment 2019a). The relationship between the Māori (indigenous people of New Zealand) and the Moana (coasts and oceans) is central in New Zealand's history (Ministry for the Environment 2019b). Thus, conservation of the marine ecosystem has particular cultural significance to New Zealanders. Between 2010 and 2017, 43% of marine non-indigenous species arriving in New Zealand have established populations in their landing area (Stats NZ 2019). Ship hulls and ballast water are recognized as the most common alien species dispersal vectors (Stats NZ 2019). A snapshot of 44 beach litter surveys in April 2019 shows that plastic accounts for 60.9% of total beach litter (Stats NZ 2019). No connection is made between marine debris and biological invasion risks in official documents.

Litter Intelligence is a database for beach survey reports from citizen scientists. The organization has recruited 4,536 citizen scientists and completed 416 surveys over 128 sites since May 2018 (Litter Intelligence n.d.). At each site, a 100-meter by 20-meter area is surveyed four times each year. Volunteers record debris type, weight, visual assessment for plastic pellets, and findings of large floating or stranded debris (Litter Intelligence n.d.). The data collection sheets do not account for fouling organisms.

Sustainable Coastlines is a nationwide charity organization that hosts regular beach cleanup events under its program Love Your Coast (Sustainable Coastlines n.d.). The program has delivered or enabled the removal of 219,747 kilograms of litter from New Zealand coastlines since 2009, involving 178,991 volunteer cleanup hours (Sustainable Coastlines n.d.). Upon sighting of animals trapped in litter, volunteers are expected to report the location of the animal, but it is unclear whether this scenario applies to fouling biota (Love Your Coast n.d.). In addition to beach cleanup activities, Sea Cleaners coordinates waterway cleanups, and has collected 8.8 million liters of litter from the coast since 2012 (Sea Cleaners n.d.).

#### **4.2.4 Hawaii, U.S.**

The Department of Land and Natural Resources (DLNR) of the State of Hawaii provides a variety of resources regarding marine debris on its website. Marine debris carrying living organisms is a high concern, and citizens are urged to report sightings of such debris to the debris management section of the Hawaii Government (DLNR 2017). JTMD is recognized as a major source of marine debris in Hawaii, and transportation of alien species is acknowledged as a significant hazard of JTMD (DLNR 2017). A response guideline to JTMD is available on the governmental website. The guideline includes detailed steps for removal, preservation, and reporting of debris with living organisms other than gooseneck barnacles (which are not hazardous) (DLNR 2013). Pictures of gooseneck barnacles as well as major alien species of concern are given to assist identification. Other resources available on the website include: an informational flyer about aquatic invasive species and JTMD, a description of the 2011 Japan earthquake and tsunami and the impacts of JTMD, debris handling guidelines, and a hazardous marine debris handbook (DLNR 2017). The Division of Boating and Ocean Recreation of DLNR has a well-developed marine debris removal request form in which citizens are required to answer questions about biofouling levels of the debris (DLNR n.d.). However, no beach survey programs for citizen scientist participation are mentioned on these governmental websites.

The NOAA Marine Debris Program works closely with the removal of marine debris from the Hawaii coastlines. Marine debris as a vector for invasive species is named as a research priority, and there is an identified need for more samples of fouling biota (NOAA 2019). NOAA has removed 848 metric tons of derelict fishing gear from the Northern Hawaiian Islands since 1996 (Pacific Islands Fisheries Science Center 2019). Founded in 2015, 808 Cleanups has organized beach cleanups and derelict fishing gear removal. The weight of litter is recorded in each cleanup event, contributing to the organization's monthly and yearly reports (808 Cleanups n.d.). Sustainable Coastlines Hawaii (SCH), founded in 2011 with inspiration from Sustainable Coastlines New Zealand, has collected 519,286 pounds of coastal litter through its 30,875 volunteers (SCH n.d.). Other beach cleanup efforts include Hawaii Ocean Ambassadors, local chapters of Surfrider Foundation, Friends of Hanauma Bay, etc.

#### **4.2.5 California, U.S.**

The California Coastal Commission (CCC) (2018) recognizes the spread of non-indigenous species by marine debris, along with entanglement and ingestion, as major threats to wildlife. Plastic is the top concern among all types of marine debris. However, the 2018 California Ocean Litter Prevention Strategy does not include response to JTMD and the risk of biological invasion. Although the strategy states the importance of beach cleanups and the removal of derelict fishing gear from the sea, the strategy's emphasis is on source reduction of litter (OPC & NOAA 2018).

California participates in the NOAA Marine Debris Program, which allows volunteers to conduct standing-stock (existing amount of debris, do not



remove debris) or accumulation (deposit rate of debris, remove debris) beach surveys that follow the Marine Debris Program protocol and contribute to the NOAA Marine Debris Monitoring and Assessment Project database (NOAA 2012). An accumulation survey involves a 100-meter shoreline segment, while a standing-stock survey randomly selects four five-meter segments within the 100-meter shoreline. The datasheet for each survey does not ask for information on fouled debris. K-12 schools are highly encouraged to lead student participation in beach surveys, and abundant lecture and activity resources are provided for marine debris education (Nally et al. 2017).

Since 1985, Coastal Cleanup Day takes place each September, attracting a total of nearly 1.6 million volunteers and collecting nearly 25 million pounds of litter from California waterways (CCC n.d.). Following each cleanup day is an extensive survey categorizing each item and accounting for the total amount in each category. Volunteer groups may also choose to “adopt-a-beach,” that is, to commit to cleaning a beach three times per year at minimum (CCC n.d.). Other than the state-wide Coastal Cleanup Day, the majority of beach cleanup organizers are regional, such as regional Coastkeeper branches, chapters of Surfrider Foundation, Save Our Beach (Seal Beach cleanup) etc.

## 5. Discussion

### 5.1 Sampling studies of fouling biota

Sampling studies of fouling biota on floating marine debris are rare, which corresponds to the fact that quantitative surveys of floating marine debris have been scarcely developed, compared to beach surveys (Lippiatt et al. 2013). For studies that sample debris from beachfronts, there is a lack of consistent methodology despite some commonalities in general approaches. This inconsistency in methodology is a source of bias in evaluating the species composition of fouling biota and can make comparing results across studies extremely difficult.

Different studies have chosen different sizes, shapes, and locations of sampling areas in their study sites, producing biased results for debris composition. Studies 1, 5, 6, 7, 8, and 9 collected natural debris in addition to anthropogenic debris. Natural debris is likely to break down over long distances and is therefore less likely to transport invasive species. However, such transportation is possible under significant natural events. While it is debatable whether natural debris should be sampled for the purpose of these studies, it is important that standards should be consistent across different studies. Several studies collected fouled debris only, and Study 9 collected only plastic bottles and Abandoned, Lost or otherwise Discarded Fishing Gear (ALDFG). Different studies have set varying minimum sizes for debris collection. It is almost impossible to compare the proportion of fouled debris among total debris and the abundance of plastic among fouled debris across different studies. Calculating the proportion of debris with rafting biota among total debris in the



sampling area and determining the composition of fouled debris are useful ways to quantify the risk of biological invasion and conduct cross-regional comparisons. However, the heterogeneity in debris types and characteristics of sampling areas across different studies hinders these assessments.

Among the reviewed studies, stranded debris sampling is less opportunistic in location selection than floating debris sampling. However, the sample collected from stranded debris is inherently biased in reflecting the biofouling status of all marine debris in the area. Stranded debris is often only representative of floating debris that is already near the coastline, and therefore may not account for debris in the open ocean. Biota on stranded debris only represent a subset of the taxa attached to the debris when it was floating, as phyla such as hydrozoan tend to degrade soon after being washed ashore (Shabani et al. 2019).

Beach samplings are also likely to result in a mixture of true stranded debris and litter generated from beach activities at the location. Thus, estimating the source of debris is important for evaluating the risk of biological invasion. Existing studies often adopt the rather ambiguous “sea-based” and “land-based” definitions to classify litter. These definitions only reflect the hypothetical location where the debris became litter. For instance, a plastic bottle that was discarded on the beach, washed into the ocean, and transported to another beach would be classified as “land-based,” although the bottle has travelled a significant distance in the sea and has possibly been biofouled. For the purpose of examining the potential of marine debris to transport NIS, it is important to determine whether the debris has been travelling in the water. Study 5 only selected debris suspected to be JTMD, all of which are assumed to have travelled in the water. Studies 1, 3, and 8 collected debris between the high and low tidelines, which increases the chances of collecting true stranded debris. Despite these various methods to increase the proportion of true stranded debris among the entire sample, there is a lack of generally applicable methods to determine whether debris has travelled in the sea.

A sampling study often provides some preliminary assessment of biological invasion risks, and often needs further investigations to validate these proposed risks. Identifying rafting biota to the lowest possible taxonomic levels is a crucial foundation for in-depth studies. For beach samplings, identification of degraded biota is a major challenge, and even genetic barcoding could fail to recognize some species. However, maintaining the commitment to identify as many species as possible is important. Study 9 only identified rafting biota to phyla level, which does not provide any direction for risk management in the region of interest. Studies 1 and 6 employed visual identification methods only, which may result in inaccurate identifications and a large number of unidentifiable taxa.

The goals of different studies also vary tremendously. This may be a reason for the vast differences in sampling and data evaluation methods. For example, Study 4 focuses on the connection between port infrastructure and NIS risks and therefore selected a small number (three) of large size areas around major ports to conduct debris sampling. All studies seem to emphasize on exploring whether the risk for marine debris to transport invasive species exist, and do not sufficiently evaluate the magnitude of the risk. There are still major gaps between the results of sampling studies and the basis for intervening in the status quo.

## **5.2 Coastal plastic management**

Plastic has been recognized as the single most abundant and problematic marine litter by all four regions examined in this component of this study. Although three regions recognize that marine debris (particularly plastic) has the potential to introduce invasive species, only Hawaii elaborates on the issue and provides guidelines for citizens on the governmental website. As an area severely impacted by JTMD, Hawaii's response to biofouled marine debris is only a part of the greater response to JTMD. Other regions that have endured smaller or no impact from JTMD appear to be less motivated in responding to biofouled marine debris. Compared to ingestion and entanglement, NIS transportation has not been presented as a constant risk that requires sustained response. Instead, it is still considered a hazard induced by individual disaster events.

Beach surveys in Australia, New Zealand and California focus mainly on the composition and abundance of litter. This practice evaluates the scale of pollution at the sampling sites, but disregards the specific environmental risks posed by the pollution, such as biological invasion. In Hawaii, step-by-step guidelines and a well-developed reporting system are available for citizens to refer to upon sighting of biofouled debris. Such measures are missing in the three other regions. In all beach cleanup and beach survey programs reviewed in this study, it appears that volunteers are not trained to respond to debris fouled by biota, especially living biota, which may be a significant hazard.

For each of the four regions, there have been several widely successful beach cleanup programs, attracting large numbers of volunteers and collecting large amounts of beach litter. In New Zealand, there is also a prominent floating debris cleanup program (Sea Cleaners) that has collected 8.8 million liters of debris. The majority of cleanups (beach or at-sea) record basic information of the collected debris, often including the weight, type, and count of the debris. Plastic is identified as the single most abundant litter type. None of the cleanups reviewed in this study record information on biofouling on debris. After being collected, the litter is typically disposed of directly or used for recycling, a process in which fouling biota would be lost and could not be retrieved. For decades, the success of a cleanup organization has been measured by the amount of litter collected. Cleanup efforts have been effective in raising public awareness and leveraging public engagement to remove litter, but are largely unaware of the important information the litter may provide.

The ubiquity of marine plastic and its ecological impact has, particularly in economically developed regions, led to high levels of public engagement in the survey and removal of coastal plastic debris. With the discovery of biological invasion threats induced by marine plastic, the practice of simply removing plastic from coastal regions is becoming insufficient, or even counterproductive.

### **5.3 Bridging the gap between samplings studies and coastal plastic management**

Plastic pollution in coastal areas cannot be resolved solely by the efforts of academic institutions. Studies evaluating the risks of biological invasion induced by marine plastic have been sporadic, often focusing on relatively small regions, and are not keeping up with the pace of the massive plastic movements in the oceans. Following the rise of awareness of plastic pollution, public engagement has proven to be successful in removing beach litter and collecting beach litter data, and has gathered data to feed back into scientific studies. While the public may not be able to identify each species and conduct analysis, with proper training, they may be able to gather and report fouling status of coastal litter, in addition to reporting the type and abundance of the litter.

There have been successful examples of incorporating citizen efforts into scientific studies. eBird is a bird observation database that collects millions of observation data submitted by bird watchers across the globe, which contributes to research and conservation efforts (Vriend 2020). The BioBlitz program allows families, teachers, students and other community members to assess biodiversity at a particular area and contribute to the Global Biodiversity Information Facility (National Geographic n.d.). Study 5 in this study also relied heavily on biofouled JTMD samples retrieved by individual beach cleanup organizations and beachcombers.

While citizen science may improve the sporadicity and small-scale nature of sampling studies, it is important that data collection follow certain professional guidelines. For example, collected biota may need to be preserved in designated preservatives, such as ethanol. To maximize the precision of data, sampling methods in beach surveys and beach cleanups need to be standardized to match those of scientific studies. However, this is particularly challenging given that no unified methods are available among scientific studies so far. The development of consistent data collection methods across scientific sampling studies should take place simultaneously with educating the public about the risk of biological invasion via marine plastic.

### **5.4 Limitations**

This study focuses mainly on the response to macroscopic biota on plastic debris and does not offer much discussion on microbiota. The study is limited by language barrier and only examined the coastal plastic management of four English-speaking regions. All selected regions are economically developed and already have high public awareness of plastic pollution, and therefore cannot

reflect the coastal plastic management status of regions with different social structures and economic standings. For each region, all information was drawn from the websites of governmental departments, academic institutions, and non-governmental organizations, which may not represent information available in other forms (e.g. posters, pamphlets, and in-person workshops). Reports on individual cases of NIS rafting on anthropogenic debris are a great resource for studying marine plastic as a biological invasion vector. However, these reports do not involve debris sampling, and are therefore not included in this study.

## 6. Conclusion

Biological invasion is only one of the many threats that marine plastic poses to the environment, and more of these threats may yet be discovered. There are still numerous gaps in the knowledge of the invasive potential of rafting biota, and there is still a long way to go before systematic evaluation of biological invasion risks could be applied on a global scale. Early response to the risks of biological invasion could prevent enormous ecological and economic costs for an area. It is important that a consistent method is developed for future sampling studies so that the risk of biological invasion via marine plastic could be quantified and be comparable across different locations.

More in-depth investigation into coastal plastic management policies and efforts of non-governmental entities may help build more connections between public engagement and response to biological invasion via marine plastic. Regions with different social structures and economic statuses than the four regions in this study may have drastically different coastal plastic management schemes and should be examined separately in future studies. Evaluation of coastal areas' response to marine plastic as a biological invasion vector could also be approached from the biological invasion perspective itself. There is a need for further research on the interactions between rafting biota and its substrates throughout the entire process from attachment to the establishment of a population in a new area.

The success of public engagement in beach cleanups and beach surveys can potentially be applied to gathering data on fouling biota. While managing biological invasions given the enormous scale of marine plastic could be overwhelming, raising public awareness of this new biological invasion vector and developing fouling biota reporting methods for common citizens are important steps to effectively responding to this rising threat induced by marine plastic.

## Bibliography

- Australia Marine Debris Database. "Data Access & Reporting." Accessed July 7, 2020. <http://amdi.tangaroablue.org/dashboard>.
- Barnea, Nir, Liam Antrim, David Lott, George Galasso, Todd Suess, and Steven Fradkin. "Summary Report and Lessons Learned: The Response to the Misawa Dock on the Washington Coast." National Oceanic and

- Atmospheric Administration Marine Debris Program, October 31, 2013. <https://marinedebris.noaa.gov/file/2369/download?token=XaLs5Ghs>.
- Barnes, David K. A. "Invasions by Marine Life on Plastic Debris." *Nature* 416, no. 6883 (2002): 808–809. <https://doi.org/10.1038/416808a>.
- Barnes, D. K. A. and P. Milner. "Drifting Plastic and Its Consequences for Sessile Organism Dispersal in the Atlantic Ocean." *Marine Biology* 146, no. 4 (2004): 815–825. <https://doi.org/10.1007/s00227-004-1474-8>.
- Bravo, Macarena, Juan Carlos Astudillo, Domingo Lancellotti, Guillermo Luna-Jorquera, Nelson Valdivia, and Martin Thiel. "Rafting on Abiotic Substrata: Properties of Floating Items and Their Influence on Community Succession." *Marine Ecology Progress Series* 439 (2011): 1–17. <https://doi.org/10.3354/meps09344>.
- California Coastal Commission. "Adopt-A-Beach®." Accessed July 10, 2020. <https://www.coastal.ca.gov/publiced/aab/aab1.html>.
- California Coastal Commission. "The Problem with Marine Debris." Last updated November 2018. <https://www.coastal.ca.gov/publiced/marinedebris.html>.
- California Coastal Commission. "California Coastal Cleanup Day: Celebrating 35 Years of Cleaning California's Coastal and Inland Waterways!." Accessed July 10, 2020. <https://coastalcomm.maps.arcgis.com/apps/Cascade/index.html?appid=de6a98eb91a24bd3ab51826f5e8eb7d2>.
- California Ocean Protection Council and National Oceanic and Atmospheric Administration Marine Debris Program. "California Ocean Litter Prevention Strategy: Addressing Marine Debris from Source to Sea." June 2018. [https://marinedebris.noaa.gov/sites/default/files/publications-files/2018\\_California\\_Litter\\_Strategy.pdf](https://marinedebris.noaa.gov/sites/default/files/publications-files/2018_California_Litter_Strategy.pdf).
- Campbell, Marnie L., Staci King, Lara D. Heppenstall, Ella van Gool, Ross Martin, and Chad L. Hewitt. "Aquaculture and Urban Marine Structures Facilitate Native and Non-indigenous Species Transfer through Generation and Accumulation of Marine Debris." *Marine Pollution Bulletin* 123, no. 1-2 (2017): 304-312. doi:10.1016/j.marpolbul.2017.08.040.
- Carlton, James T., John W. Chapman, Jonathan B. Geller, Jessica A. Miller, Deborah A. Carlton, Megan I. Mcculler, Nancy C. Treneman, Brian P. Steves, and Gregory M. Ruiz. "Tsunami-driven Rafting: Transoceanic Species Dispersal and Implications for Marine Biogeography." *Science* 357, no. 6358 (2017): 1402-406. doi:10.1126/science.aa01498.
- Commonwealth of Australia. "Threat Abatement Plan: for the Impacts of Marine Debris on the Vertebrate Wildlife of Australia's Coasts and Oceans." 2018. <http://www.environment.gov.au/system/files/resources/e3318495-2389-4ffc-b734-164cdd67fe19/files/tap-marine-debris-2018.pdf>.
- Commonwealth Scientific and Industrial Research Organisation. "Marine Debris: Sources, Distribution and Fate of Plastic and other Refuse – and its Impact on Ocean and Coastal Wildlife." Accessed July 6, 2020. [https://www.sciro.au/~media/OnA/Files/1500156\\_OA\\_MarineDebris4ppFactsheet\\_WEB\\_190405.pdf](https://www.sciro.au/~media/OnA/Files/1500156_OA_MarineDebris4ppFactsheet_WEB_190405.pdf).
- Crooks, Jeffrey A. "Lag Times and Exotic Species: The Ecology and Management of Biological Invasions in Slow-motion." *Écoscience* 12, no. 3 (2005): 316–329. <https://doi.org/10.2980/i1195-6860-12-3-316.1>.

- Department of Agriculture, Water and the Environment, Australian Government. "Marine Debris Including Plastic Pollution." Accessed June 30, 2020. <https://www.environment.gov.au/marine/marine-pollution/marine-debris>.
- Department of Agriculture, Water and the Environment, Australian Government. "Marine Pests." Accessed June 30, 2020. <https://www.environment.gov.au/marine/marine-pollution/marine-pests>.
- Department of Land and Natural Resources. "Japan Tsunami Marine Debris: What to do if you see debris in Hawaii's ocean or beaches." Department of Land and Natural Resources, State of Hawaii. Last updated April 2013. <https://dlnr.hawaii.gov/wp-content/uploads/2013/01/JTMD-Guideline3.pdf>.
- Department of Land and Natural Resources. "Please Kokua: Marine Debris is Everybody's Problem." Department of Land and Natural Resources, State of Hawaii. Last updated June 2017. <https://dlnr.hawaii.gov/marine-debris/#:~:text=To%2oreport%2olarge%2oquantities%2oof,and%2odisasterdebris%40noaa.gov>.
- Department of Land and Natural Resources. "Marine Debris Response and Removal Reporting Form." Division of Boating and Ocean Recreation, Department of Land and Natural Resources, State of Hawaii. Accessed July 3, 2020. <https://dlnr.hawaii.gov/dobor/reportmarinedebrishawaii/>.
- Duhec, Aurélie V., Richard F. Jeanne, Nikolai Maximenko, and Jan Hafner. "Composition and Potential Origin of Marine Debris Stranded in the Western Indian Ocean on Remote Alphonse Island, Seychelles." *Marine Pollution Bulletin* 96, no. 1-2 (2015): 76-86. doi:10.1016/j.marpolbul.2015.05.042.
- Eco Barge Clean Seas Inc. Accessed July 8, 2020. <https://ecobargecleanseas.org.au/>.
- Eriksen, Marcus, Laurent C. M. Lebreton, Henry S. Carson, Martin Thiel, Charles J. Moore, Jose C. Borroero, Francois Galgani, Peter G. Ryan, and Julia Reisser. "Plastic Pollution in the World's Oceans: More than 5 Trillion Plastic Pieces Weighing over 250,000 Tons Afloat at Sea." *PLoS ONE* 9, no. 12 (2014). doi:10.1371/journal.pone.0111913.
- Goldstein, Miriam C., Henry S. Carson, and Marcus Eriksen. "Relationship of Diversity and Habitat Area in North Pacific Plastic-associated Rafting Communities." *Marine Biology* 161, no. 6 (2014): 1441-1453. doi:10.1007/s00227-014-2432-8.
- Gracia, C. A., Nelson Rangel-Buitrago, and Paola Flórez. "Beach Litter and Woody-Debris Colonizers on the Atlantico Department Caribbean Coastline, Colombia." *Marine Pollution Bulletin* 128 (2018): 185-196. <https://doi.org/10.1016/j.marpolbul.2018.01.017>.
- Hardesty, Britta Denise, Tj Lawson, Tonya Van Der Velde, Matt Lansdell, and Chris Wilcox. "Estimating Quantities and Sources of Marine Debris at a Continental Scale." *Frontiers in Ecology and the Environment* 15, no. 1 (2017): 18-25. <https://doi.org/10.1002/fee.1447>.
- Hoeksema, Bert W., Kevin Padoja, and Yohann Poprawski. "Long-distance Transport of a West Atlantic Stony Coral on a Plastic Raft." *Ecology* 99, no. 10 (2018): 2402-2404. doi:10.1002/ecy.2405.

- International Maritime Organization. "Ballast water management." Accessed June 7, 2020. <http://www.imo.org/en/OurWork/Environment/BallastWaterManagement/Pages/Default.aspx>.
- Ivkić, Angelina, Jan Steger, Bella S. Galil, and Paolo G. Albano. "The Potential of Large Rafting Objects to Spread Lessepsian Invaders: the Case of a Detached Buoy." *Biological Invasions* 21, (2019): 1887–1893. <https://doi.org/10.1007/s10530-019-01972-4>.
- Jambeck, Jenna R., Roland Geyer, Chris Wilcox, Theodore R. Sieglar, Miriam Perryman, Anthony Andrady, Ramani Narayan, and Kara Lavender Law. "Plastic Waste Inputs from Land into the Ocean." *Science* 347, no. 6223 (2015): 768–771. doi:10.1126/science.1260352.
- Kiessling, Tim, Lars Gutow, and Martin Thiel. "Marine Litter as Habitat and Dispersal Vector." in *Marine Anthropogenic Litter*, 141–81, (Springer Cham, 2015), doi:10.1007/978-3-319-16510-3\_6.
- Li, W. C., H. F. Tse, and L. Fok. "Plastic Waste in the Marine Environment: A Review of Sources, Occurrence and Effects." *Science of The Total Environment* 566–567 (2016): 333–349. doi:10.1016/j.scitotenv.2016.05.084.
- Litter Intelligence. Accessed July 8, 2020. <https://litterintelligence.org/>.
- Litter Intelligence. "How to Collect Litter Data: Methodology Summary for Citizen Scientists." Accessed July 1, 2020. [https://litterintelligence.org/media/s5thddoy/5-sc\\_litter-intelligence-methodology-v1-4.pdf](https://litterintelligence.org/media/s5thddoy/5-sc_litter-intelligence-methodology-v1-4.pdf).
- Litter Intelligence. "Survey Details: Survey Area & Large Item Information." Accessed July 1, 2020. [https://litterintelligence.org/media/gnyd55py/2-sc\\_litter-intelligence-survey-area-large-items-v1-3.pdf](https://litterintelligence.org/media/gnyd55py/2-sc_litter-intelligence-survey-area-large-items-v1-3.pdf).
- Lippiatt, Sherry, Sarah Opferand, and Courtney Arthur, "Marine Debris Monitoring and Assessment: Recommendations for Monitoring Debris Trends in the Marine Environment." NOAA Technical Memorandum NOS-OR&R-46. November 2013. <https://marinedebris.noaa.gov/sites/default/files/Lippiatt%20et%20al%202013.pdf>.
- Love Your Coast. "Safety Brief." Accessed July 8, 2020. <http://www.loveyourcoast.org.nz/assets/Love-Your-Coast/safety-briefing.pdf>.
- Ministry for the Environment. "Why Our Marine Environment Matters." Last reviewed October 16, 2019. <https://www.mfe.govt.nz/marine/why-our-marine-environment-matters>.
- Ministry for the Environment. "Our Marine Environment 2019." 2019. <https://www.mfe.govt.nz/publications/marine/our-marine-environment-2019>.
- Miralles, Laura, Marta Gomez-Agenjo, Fernando Rayon-Viña, Greta Gyraitė, and Eva Garcia-Vazquez. "Alert Calling in Port Areas: Marine Litter as Possible Secondary Dispersal Vector for Hitchhiking Invasive Species." *Journal for Nature Conservation* 42 (2018): 12–18. <https://doi.org/10.1016/j.jnc.2018.01.005>.
- Molnar, Jennifer L., Rebecca L. Gamboa, Carmen Revenga, and Mark D. Spalding. "Assessing the Global Threat of Invasive Species to Marine Biodiversity." *Frontiers in Ecology and the Environment* 6, no. 9 (2008): 485–92. <https://doi.org/10.1890/070064>.

- Nally, Alyssa, Sherry Lippiatt, Seaberry Nachbar, Naomi Pollack, "Marine Debris Toolkit For Educators." NOAA Marine Debris Program & NOAA Office of National Marine Sanctuaries. 2017. <https://marinedebris.noaa.gov/sites/default/files/publications-files/MarineDebrisMonitoringToolkitForEducators.pdf>.
- National Geographic. "BioBlitz: Counting Species Through Citizen Science." Accessed July 24, 2020. <https://www.nationalgeographic.org/projects/bioblitz/>
- National Marine Debris Database. "Beach Litter Survey Methodology." The Commonwealth Scientific and Industrial Research Organisation. Accessed June 30, 2020. [http://www.cmar.csiro.au/datacentre/cmar\\_public/docs/Debris\\_Transect\\_Survey\\_Methodology.pdf](http://www.cmar.csiro.au/datacentre/cmar_public/docs/Debris_Transect_Survey_Methodology.pdf)
- National Marine Debris Database. "EMU PARADE Guidelines." The Commonwealth Scientific and Industrial Research Organisation. Accessed June 30, 2020. [http://www.cmar.csiro.au/datacentre/cmar\\_public/docs/Debris\\_Emu\\_Parade\\_Survey\\_Methodology.pdf](http://www.cmar.csiro.au/datacentre/cmar_public/docs/Debris_Emu_Parade_Survey_Methodology.pdf).
- National Oceanic and Atmospheric Administration Marine Debris Program. "NOAA Marine Debris Shoreline Survey Field Guide." January 2012. <https://marinedebris.noaa.gov/sites/default/files/publications-files/ShorelineSurveyFieldGuide.pdf>.
- National Oceanic and Atmospheric Administration Marine Debris Program. "Invasive Species: Marine Debris as a Potential Pathway for Invasive Species" Silver Spring, MD: National Oceanic and Atmospheric Administration Marine Debris Program. March 2017. [https://marinedebris.noaa.gov/sites/default/files/publications-files/2017\\_Invasive\\_Species\\_Topic\\_Paper.pdf](https://marinedebris.noaa.gov/sites/default/files/publications-files/2017_Invasive_Species_Topic_Paper.pdf).
- National Oceanic and Atmospheric Administration Marine Debris Program. "2019 Hawai'i Marine Debris Action Plan Research Workshop." Silver Spring, MD: National Oceanic and Atmospheric Administration Marine Debris Program. October 2019. [https://marinedebris.noaa.gov/sites/default/files/publications-files/2019%20HI-MDAP%20Research%20Workshop%20Proceedings\\_FINAL\\_508\\_0.pdf](https://marinedebris.noaa.gov/sites/default/files/publications-files/2019%20HI-MDAP%20Research%20Workshop%20Proceedings_FINAL_508_0.pdf).
- Pacific Islands Fisheries Science Center. "Marine Debris Research and Removal in the Northwestern Hawaiian Islands." National Oceanic and Atmospheric Administration Fisheries. Last updated December 18, 2019. <https://www.fisheries.noaa.gov/pacific-islands/marine-debris-research-and-removal-northwestern-hawaiian-islands>.
- Pimentel, David. *Biological Invasions: Economic and Environmental Costs of Alien Plant, Animal, and Microbe Species*. Boca Raton, Fl: CRC, 2011.
- Rech, Sabine, Yaisel J. Borrell Pichs, and Eva García-Vazquez. "Anthropogenic Marine Litter Composition in Coastal Areas May Be a Predictor of Potentially Invasive Rafting Fauna." *Plos One* 13, no. 1 (2018a). doi:10.1371/journal.pone.0191859.
- Rech, Sabine, Simone Salmina, Yaisel J. Borrell Pichs, and Eva García-Vazquez. "Dispersal of Alien Invasive Species on Anthropogenic Litter from European Mariculture Areas." *Marine Pollution Bulletin* 131 (2018b): 10–16. <https://doi.org/10.1016/j.marpolbul.2018.03.038>.
- Sea Cleaners. "We Are Results Driven." Accessed July 8, 2020. <https://seacleaners.com/>.



- Shabani, Fatemeh, Ali Nasrolahi, and Martin Thiel. "Assemblage of Encrusting Organisms on Floating Anthropogenic Debris along the Northern Coast of the Persian Gulf." *Environmental Pollution* 254 (2019): 112979. doi:10.1016/j.envpol.2019.112979.
- Simberloff, Daniel. *Invasive Species: What Everyone Needs to Know*. Oxford: Oxford University Press, 2013. [https://www.google.com/books/edition/Invasive\\_Species/DKLEvpmooUC?hl=en&gbpv=1](https://www.google.com/books/edition/Invasive_Species/DKLEvpmooUC?hl=en&gbpv=1).
- Simberloff, Daniel, Jean-Louis Martin, Piero Genovesi, Virginie Maris, David A. Wardle, James Aronson, Franck Courchamp, Bella Galil, Emili García-Berthou, Michel Pascal, Petr Pyšek, Ronaldo Sousa, Eric Tabacchi and Montserrat Vilà. "Impacts of Biological Invasions: What's What and the Way Forward." *Trends in Ecology & Evolution* 28, no. 1 (2013): 58–66. <https://doi.org/10.1016/j.tree.2012.07.013>.
- Simkanin, Christina, James T. Carlton, Brian Steves, Paul Fofonoff, Jocelyn C. Nelson, Cathryn Clarke Murray, and Gregory M. Ruiz. "Exploring Potential Establishment of Marine Rafting Species after Transoceanic Long-distance Dispersal." *Global Ecology and Biogeography* 28, no. 5 (2019): 588–600. doi:10.1111/geb.12878.
- Smithsonian Institution. "Written Testimony of Linsey E. Haram, Research Ecologist, Smithsonian Environmental Research Center: Plastic Marine Debris Research." U.S. House of Representatives Document Repository. September 19, 2019. <https://docs.house.gov/meetings/AP/AP06/20190919/109934/HHRG-116-AP06-Wstate-HaramL-20190919.pdf>.
- South West Marine Debris Cleanup. Accessed July 8, 2020. <http://whamarinedebris.blogspot.com/p/history-and-statistics.html>.
- Stats NZ. "New Zealand's Environmental Reporting Series: Environment Aotearoa 2019: Stats NZ." Accessed July 23, 2020. <https://www.stats.govt.nz/information-releases/new-zealands-environmental-reporting-series-environment-aotearoa-2019>.
- Sustainable Coastlines. "Our Programs." Accessed July 8, 2020. <https://sustainablecoastlines.org/about/our-programmes/>.
- Sustainable Coastlines. "Our Impacts." Accessed July 8, 2020. <https://sustainablecoastlines.org/about/our-impacts/>.
- Sustainable Coastlines Hawaii. Accessed July 9, 2020. <https://www.sustainablecoastlineshawaii.org/>.
- Tangaroa Blue. Accessed July 6, 2020. <https://www.tangaroablue.org/>.
- Thiel, Martin, and Lars Gutow. "The Ecology of Rafting in the Marine Environment. II. The Rafting Organisms and Community." *Oceanography and Marine Biology : An Annual Review* 43, (2005): 279–418. doi: 10.1201/9781420037449.ch7.
- Thresher, Ronald E., and Armand M. Kuris. "Options for Managing Invasive Marine Species." *Biological Invasions* 6, no. 3 (2004): 295–300. <https://doi.org/10.1023/b:binv.0000034598.28718.2e>.
- Vriend, Stefan. "The Bird Watching Community: Citizen Science at its Finest." *Ecology for the Masses*. June 11, 2020. <https://ecologyforthemasses.com/2020/06/11/the-bird-watching-community-citizen-science-at-its-finest/>.

- Whitehead, T. Otto, Aiden Biccard, and Charles L. Griffiths. "South African Pelagic Goose Barnacles (Cirripedia, Thoracica): Substratum Preferences and Influence of Plastic Debris on Abundance and Distribution." *Crustaceana* 84, no. 5 (2011): 635–649. <https://doi.org/10.1163/001121611x574290>.
- Wu, Hsing-Hao. "A Study on Transnational Regulatory Governance for Marine Plastic Debris: Trends, Challenges, and Prospect." *Marine Policy* (2020): 103988. <https://doi.org/10.1016/j.marpol.2020.103988>.
- 808 Cleanups. "Create a Team: Adopt a Site." Accessed July 9, 2020. <https://808cleanups.org/adopt-a-site/>.



# Semi-Supervised Classification with Generative Adversarial Networks on Medical Datasets with Limited Size and Label

---

*Lutfi Eren Erdogan*

**Author Background:** *Lutfi Eren Erdogan grew up in Turkey and currently attends Uskudar American Academy in Istanbul, Turkey. His Pioneer research concentration was in the field of computer science/biology and titled “Biology in Computing: Why Are Networks Neural?”*

## **Abstract**

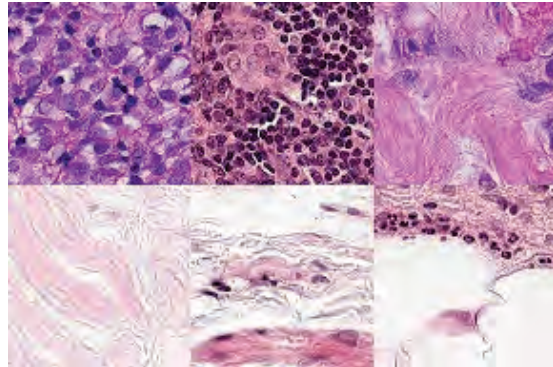
One of the challenges faced within medical image classification studies is the fact that they usually suffer from limited sized datasets with few annotations as collecting large amounts of annotated medical data requires extensive financial resources and time. In particular, too little data can cause overfitting, and the images with no annotations cannot be utilized for supervised training. In this paper, we propose a semi-supervised approach that is built upon the recent advances in generative models and that can train a powerful classifier on medical datasets with limited data and few annotations. Our experiments demonstrate the consistent gains of our method over state-of-the-art classification networks. With only 20% of the annotations used for training, our model is able not only to match the performance of a state-of-the-art classification network trained with all the annotations, but even to outperform it. Our code is available at: <https://github.com/xTRam1/Semi-Supervised-Classification-with-GANs-on-Medical-Datasets-with-Limited-Size-and-Label>

## **1. Introduction**

Deep learning, and in particular convolutional neural networks (CNNs), is driving significant advances in the analysis and interpretation of medical imaging applications including reconstruction, segmentation, diagnosis, and treatment response assessment in radiology, dermatology, and pathology (Greenspan et al., 2016; Esteva et al., 2017; Gulshan et al., 2016). For classification, medical imaging studies typically employ supervised convolutional neural networks that require large datasets annotated by experts to obtain high predictive performance. However, obtaining large volumes of labelled training data is time-consuming and requires extensive expertise. Thus, most medical datasets are still limited in size, and in many applications there is a need for multiple annotations to reduce

labelling noise (Gulshan et al., 2016; Krause et al., 2018). Because of this, two problems arise for medical image classification: 1. The limited size of the dataset; 2. The limited number of annotated images. Unfortunately, no past study has tried to tackle both problems simultaneously; all have instead focused on just one of them. While a solution to the second problem was attempted by several recent studies, they failed to give enough attention to the first problem, and similarly, while others attempted to solve the first problem, their studies didn't put enough emphasis on the second one.

The increasingly impressive results of generative adversarial networks (GANs) (Goodfellow et al., 2014; Miyato and Koyama, 2018; Miyato et al., 2018; Brockett et al., 2018; Karras et al., 2017, 2018, 2019) can help tackle the annotation burden. GANs are a type of a generative neural network model composed of two networks, one centered on generating images and the other on discrimination. These two networks are trained simultaneously to beat each other, generator trying to fool the discriminator by generating more realistic images and the discriminator trying to detect these fake generated images. Recent works successfully implemented GANs to image synthesis and semi-supervised learning where given a small labelled dataset and a large unlabelled dataset, the model can train a powerful classifier (Yi et al., 2018; Madani et al., 2018; Hu et al., 2019; Lecouatet et al., 2018; Diaz-Pinto et al., 2019). Particularly, some studies were able to show the applicability of Semi-Supervised GANs (SS-GAN) (Salimans et al., 2016) in classifying skin and heart diseases (Yi et al., 2018; Madani et al., 2018; Madani et al., 2018).



**Figure 1.** Our GAN is attempting to classify images taken from scans of lymph nodes that may or may not be cancerous. On the top are three images of cancerous lymph nodes scans, and on the bottom are three of the non-cancerous ones.

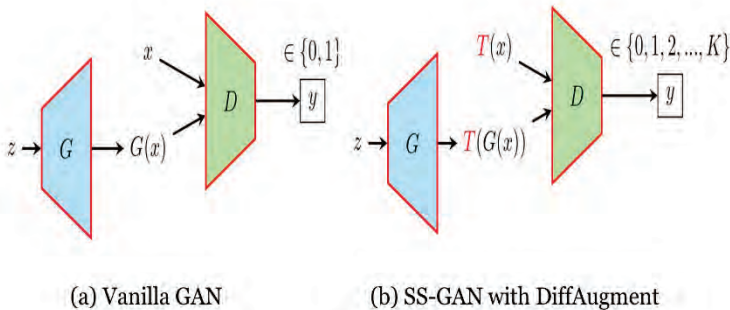
On the other hand, in order to tackle the limited data size problem, several studies used GANs for synthesizing images as a means of data augmentation for supervised classification tasks on small fully annotated datasets, training their GANs on these small datasets (Salehinejad et al., 2017; Frid-Adar et al., 2018). However, the limited amount of training data for GAN training results in drastic degradation and overfitting in GAN performance, causing the discriminator to memorize the exact training set and causing training to diverge (Karras et al., 2020; Zhao et al., 2020). This severe overfitting then disrupts the training dynamics and leads to degraded image quality, which is highly undesirable in medical imaging. One study, in an attempt to increase its GAN training dataset size, tried to use data augmentation for skin lesion classification on their small training dataset (Frid-Adar et al., 2018). This unfortunately led to outputs suffering from distribution shifts and artifacts, since applying data augmentation to GANs is fundamentally different to the normal neural network training.

While for reducing overfitting in image classification, "data augmentation" is a widely used strategy that increases the diversity of training data without collecting new samples (Krizhevsky et al., 2012; Simard et al., 2003; Wan et al., 2013), for GAN training, a GAN trained under dataset augmentations learns to generate the augmented distribution (Karras et al., 2020). For example, a noise augmentation leads to noisy results, even if there are none in the dataset. In general, such "leaking" of augmentations to the generated samples is highly undesirable. As a result, these studies were unable to fully tackle the first problem: the limited size of the dataset.

Guided by recent studies aiming to combat the limited data sizes in GANs (Karras et al., 2020; Zhao et al., 2020), we propose a semi-supervised approach that can tackle both the limited data size and the limited annotations problems in medical image datasets. In particular, by using SS-GAN and Differentiable Augmentation (DiffAugment) (Zhao et al., 2020), a recent augmentation method that imposes various differentiable transformations on both real and fake images during GAN training, we hope to effectively train a classifier that is both stable to train and likely to converge better than current classifiers. In addition to the semi-supervised augmented strategy, we also incorporate recent insights related to GAN training stabilization by enforcing spectral normalization technique (Miyato et al., 2018) to both the discriminator and the generator.

For clarity, we emphasize that our aim is not to synthesize real looking high-quality images but to *train a powerful classifier*. In fact, Dai et al. (2017) theoretically shows that "given the discriminator objective, good semi-supervised learning indeed requires a bad generator."

## 2. Related Works



**Figure 2.** GAN Architectures. In both figures,  $z$  is the random noise,  $x$  is the real data, and  $y$  is the output of  $D$ . For 2a,  $x$  is always an unlabeled sample, while in 2b  $x$  can be both an unlabeled sample or a labeled sample.  $D$  of 2a can be thought of as a binary classifier trying to discriminate between fake and real images while  $D$  for 2b is a  $K + 1$  multi-class classifier trying to assign one of  $K$  real labels to real sample and the  $(K + 1)$ th label to the fakesamples.  $T$  represents the DiffAugment transformation applied to both  $x$  and  $G(z)$ .

The effectiveness of GANs on medical datasets has been shown by several studies for several tasks, including unsupervised image synthesis, unsupervised object detection, image-to-image modality translation, CT de-noising, and semi-supervised classification (Wolterink et al., 2017; Frid-Adar et al., 2018; Hiasa et al., 2018; Chen and Konukoglu, 2018; Yi et al., 2018; Madani et al., 2018; Madani et al., 2018).

For classification, semi-supervised and unsupervised methods using GANs have drawn recent attention within the medical scene because expert annotations for medical datasets are so time-consuming. Meanwhile, data augmentation using GANs and synthetically generated images on small-sized datasets was also studied by several papers. Madani et al. (2018) demonstrated that augmenting the original imbalanced dataset with GAN generated images improves the chest pathology classification performance. They used synthetic images for balancing the dataset where classes that are lacking in examples are preferentially augmented. Using GANs for augmentation of the CNN training set, Frid-Adar et al. (2018) synthesized high-quality focal liver lesions from CT images for improved classification results. Hu et al. (2019) used a unified WGAN-GP and InfoGAN architecture with an auxiliary network on top of the discriminator to classify cells for unsupervised classification settings. Their model achieved promising results in the unsupervised classification of bone marrow cellular components and was capable of providing interpretable visualizations. Lecouat et al. (2018) proposed a semi-supervised classification framework where before being used for training or inference with a semi-supervised GAN, high-resolution medical images were divided into equal-sized patches. They demonstrated their method's effectiveness on retinal image classification. Yi et al. (2018) used a method that combines categorical generative adversarial network (catGAN) (Springenberg, 2015) and Wasserstein distance (Arjovsky et al., 2017) (also known as earthmover's distance) for unsupervised and semi-supervised classification on dermoscopy images. Diaz-Pinto et al. (2019) trained a semi-supervised GAN for retinal image synthesis and glaucoma assessment on a small glaucoma-labelled and large unlabeled database. They used 86926 images from fourteen public databases, a very large-sized training dataset.

The core of the problem lies at this point, we think. Most of the preceding studies showed the effectiveness of semi-supervised classification tasks with datasets of large sizes, which is highly unrealistic for real-world applications, as finding data is challenging due either to financial costs or privacy issues. We therefore propose a GAN method that is inspired by recent advances in GAN architectures and augmentation strategies for semi-supervised classification on small-sized medical datasets.

### 3. Methodology

#### Vanilla GAN

The vanilla GAN<sup>1</sup> (Goodfellow et al., 2014) is a generative model that is designed to directly draw samples from a desired data distribution without explicitly modelling its probability density function. It is composed of two networks: a generator,  $G$ , and a discriminator,  $D$ .  $z$ , a pure random noise sampled from a prior distribution  $p(z)$ , is input to  $G$ . For simplicity,  $p(z)$  is commonly chosen as a Gaussian or a uniform distribution.  $x_r$  is a real sample that is drawn from a real data distribution,  $p_r(x)$ , and  $x_g$ , the output of  $G$  (fake image), is expected to have visual similarity with  $x_r$ .  $G$  learns a non-linear mapping function, which is parameterized by  $g$  and is denoted as  $x_g = G(z; \theta_g)$ .

The input of  $D$  is either a real,  $x_r$ , or a generated,  $x_g$ , sample.  $y$ , the output of  $D$ , is a binary value indicating the probability of the input being a real or a generated sample. The mapping learned by  $D$  is denoted as  $y = D(x; d)$ , which is parameterized by  $d$ . After successful training,  $p_g(x)$ , the distribution formed by generated samples, is desired to be an approximation of  $p_r(x)$ . Figure 2a shows an illustration of a vanilla GAN's configuration.  $D$ 's objective is to distinguish between real and fake (generated) samples while  $G$ 's objective is to confuse  $D$  as much as possible. During training, the gradient information is propagated from  $D$  back to  $G$ , so  $G$  adapts its parameters in order to produce an output sample that can fool  $D$ . Mathematically, the training objectives of both networks become:

$$\begin{aligned} L_D &= \max_D E_{x_r \sim p_r(x)} [\log D(x_r)] + E_{x_g \sim p_g(x)} [\log(1 - D(x_g))] \\ L_G &= \min_G E_{x_g \sim p_g(x)} [\log(1 - D(x_g))] \end{aligned} \quad (1)$$

As seen above,  $D$  is simply a binary classifier with a maximum log-likelihood objective. Furthermore, when  $D$  is trained to the optimal point before the next update of  $G$ , it is proven that  $L_G$  is equivalent to minimizing the Jensen-Shannon (JS) divergence between the real and fake distributions (Goodfellow et al., 2014). After training,  $x_g$  samples should approximate samples from  $p_r(x)$ .

#### Semi-Supervised GAN (SS-GAN)

We build upon the semi-supervised feature-matching GAN framework (Salimans et al., 2016). In addition to determining whether the image is real or generated,

---

<sup>1</sup> Vanilla GAN explanation was adapted from Karras et al. (2017) and Goodfellow et al. (2014).

with this method, the discriminator  $D$  is expected to predict the specific class of the images  $y \in \{1 \text{ fake class, } K \text{ real classes}\}$ , as seen in Figure 2b. For example, suppose we are given a dataset of images  $S = S_L \cup S_U$  consisting of a labeled set  $S_L = f(x_i; y_i)$  and an unlabeled set  $S_U = x_i$  of images; here  $x_i$  and  $y_i$  are the images and labels ( $K$  classes) respectively. Then, during training, we simultaneously optimize  $D$  and  $G$  using stochastic gradient descent, minimizing loss functions  $L_D$  for the discriminator and  $L_G$  for the generator.  $L_D$  is the sum of the loss term on the labelled images  $L_{supervised}$  and the vanilla GAN loss (on unlabeled images)  $L_{unsupervised}$  so that:

$$\begin{aligned} L_D &= L_{unsupervised} + L_{supervised} \\ L_{supervised} &= -E_{(x,y) \sim \mathbb{S}_L} [\log p_D(y | x, y < K + 1)] \\ L_{unsupervised} &= -E_{x \sim \mathbb{S}} [\log[1 - p_D(y = K + 1 | x)]] - E_{x \sim G} [\log[p_D(y = K + 1 | x)]] \end{aligned} \quad (2)$$

$L_G$ , the feature-matching loss, is designed to encourage generated images to have similar features to the real images (Salimans et al., 2016).

$$L_G = |E_{x \sim \mathbb{S}}[h(x)] - E_{z \sim p_z(z)}[h(g(z))]|_1 \quad (3)$$

$h(x)$  represents activations on an intermediate layer of the discriminator. In our experiments, the last activation layer after max-pooling was chosen as the intermediate  $h(x)$ .

### DiffAugment

With a limited amount of training data, the discriminator memorizes the exact training set so the performance of GANs heavily deteriorates (Arjovsky and Bottou, 2017; Zhang and Khoreva, 2019). To tackle this problem, Zhao et al. (2020) propose Differentiable Augmentation (DiffAugment), a recent augmentation method that imposes various differentiable transformations on both real and fake images during GAN training. This is depicted in Figure 2b. DiffAugment enables the gradients to be propagated through the augmentation back to the generator and regularizes the discriminator without manipulating the target distribution, all the while maintaining the balance of training dynamics. Directly augmenting the training data manipulates the distribution of real images, which is undesirable; DiffAugment is able to tackle this problem, leading to better convergence. Given medical datasets' limited size, DiffAugment enables us to stably train our semi-supervised GAN while our discriminator is not able to overfit to our dataset easily. With DiffAugment, the training objectives of semi-supervised  $D$  and  $G$  becomes:

$$\begin{aligned} L_D &= E_{x \sim p_{\text{data}}}(x) \left[ f_D \left( -D(T(x)) \right) \right] + E_{z \sim p(z)} \left[ f_D \left( D \left( T(G(z)) \right) \right) \right] \\ L_G &= E_{z \sim p(z)} \left[ f_G \left( -D \left( T(G(z)) \right) \right) \right] \end{aligned} \quad (4)$$



where  $T$  is a transformation function applied to the input of the discriminator. Note that  $T$  can be a random function. It does not necessarily produce the same transformations across the three positions in Equation 4. Following authors' suggestions for severe data size problems, throughout the paper, we use the strongest transformation policy: Translation (within  $[-1/8, 1/8]$  of the image size, padded with zeros), Cutout (DeVries and Taylor, 2017) (masking with a random square of half image size), and Color (including random brightness within  $[-0.5, 0.5]$ , contrast within  $[0.5, 1.5]$ , and saturation within  $[0, 2]$ ).

### Spectral Normalization

One challenge in GAN training is the control of discriminator training. The discriminator is often inaccurate and unstable to train in high dimensional spaces, causing the generator network to fail to learn the target distribution. Even worse, when the support of the model distributions and that of the real distribution are disjoint, there exists a discriminator that can perfectly distinguish the model distribution from the real one. This causes the generator training to stop as the back-propagated gradient of the discriminator becomes zero, indicating that the discriminator does not send any feedback to the generator for it to improve its image quality (Arjovsky and Bottou, 2017). To eliminate this possibility, Miyato et al. (2018) proposed a weight normalization method called spectral normalization that can stabilize the training of discriminator networks and is simple to implement while the additional costs are small. We urge readers to take a look at their work for more details.

Since GAN training with limited data size tends to be unstable, we incorporated spectral normalization to both our discriminator and our generator, following Zhang et al. (2018). We are also following the ResNet (He et al., 2015) architecture of Miyato et al. (2018) for the SVHN dataset presented in the paper, as residual blocks are more likely to result in better performance as opposed to the simple convolutional layers presented by Radford et al. (2015). The technical details of the architecture of the generator and the discriminator, along with the residual blocks, are included in Appendix A.

### Baseline

To demonstrate the effectiveness of our semi-supervised approach in both its elimination of the need for annotations and data, we compared it with a state-of-the-art supervised classification model. Since we used the residual architecture of Miyato et al. (2018), for it to be a fair comparison, our baseline is ResNet50 architecture (He et al., 2015) which has proven to be effective on large, annotated datasets, such as ImageNet. For every comparison, we used the same dataset to train our baseline and our model.

## 4. Experiments

In this section, we describe the dataset we use. Then we elaborate on the implementational details of our Semi-Supervised GAN. Finally, we present our

training and inference results on our chosen dataset both for our GAN and our baseline.

## Dataset

The dataset we used to train both our GAN and the baseline is a Kaggle Competition Dataset called "Histopathologic Cancer Detection"<sup>2</sup> which is a slightly modified version of the PatchCamelyon (PCam) dataset (Veeling et al., 2018). The PatchCamelyon dataset is a challenging binary image classification dataset. It consists of 327,680 color images (96 x 96px) extracted from histopathologic scans of lymph node sections. Each image is annotated with a binary label indicating the presence of metastatic tissue. To test our GAN approach, out of 327,680 images, while preserving the "positive to negative" sample ratio, we randomly chose 500 images for training, 100 images for validation, and 100 images for testing. For computational speed, we resized the images to (48 x 48px). Both validation and training accuracy were tracked during training. Finally, test accuracy was reported.

## Experimental Approach

To test the effectiveness of our proposed GAN architecture (SS-DiffAugment-GAN), we trained it with only 10% and 20% of the training images annotated (500 images, [50 labels, 100 labels respectively]). To compare it with our baseline, we trained our baseline with 10%, 20%, and 100% of the training images ([50, 100, 500 images], [50, 100, 500 labels respectively]). The next section describes how we chose our hyperparameters for the final training.

## Hyperparameter Selection

As GAN training is known to be extremely sensitive to hyperparameters, we needed to select the best hyperparameters. One of the most important hyperparameters is the learning rate. To select the best performing learning rate, we chose three of the most widely used GAN training learning rates for our experiments<sup>3</sup>: 0.0001-0.0004 (Zhang et al., 2018), 0.0002-0.0002 (Miyato et al., 2018), 0.001-0.003 (Dai et al., 2017). In order to compare the learning rates equally, we set the batch size for all experiments to be 16, and the maximum iteration (generator steps) to be 5000. We conducted our experiments on 20% regime and calculated validation accuracy, test accuracy, unsupervised loss, and supervised loss. Since, for each case, the generator loss was increasing just as we had expected and nearly at the same steady rate, we decided not to base our learning rate selection on the generator loss.

Looking at the progression of the supervised loss from Figure 3a, it is evident that while for 0.0001-0.0004 and 0.0002-0.0002, the loss steadily decreases with similar rates, the loss for 0.001-0.003 does not seem to improve

---

<sup>2</sup> The dataset is available at <https://www.kaggle.com/c/histopathologic-cancer-detection>

<sup>3</sup> The first ones are for the generator and the second ones are for the discriminator

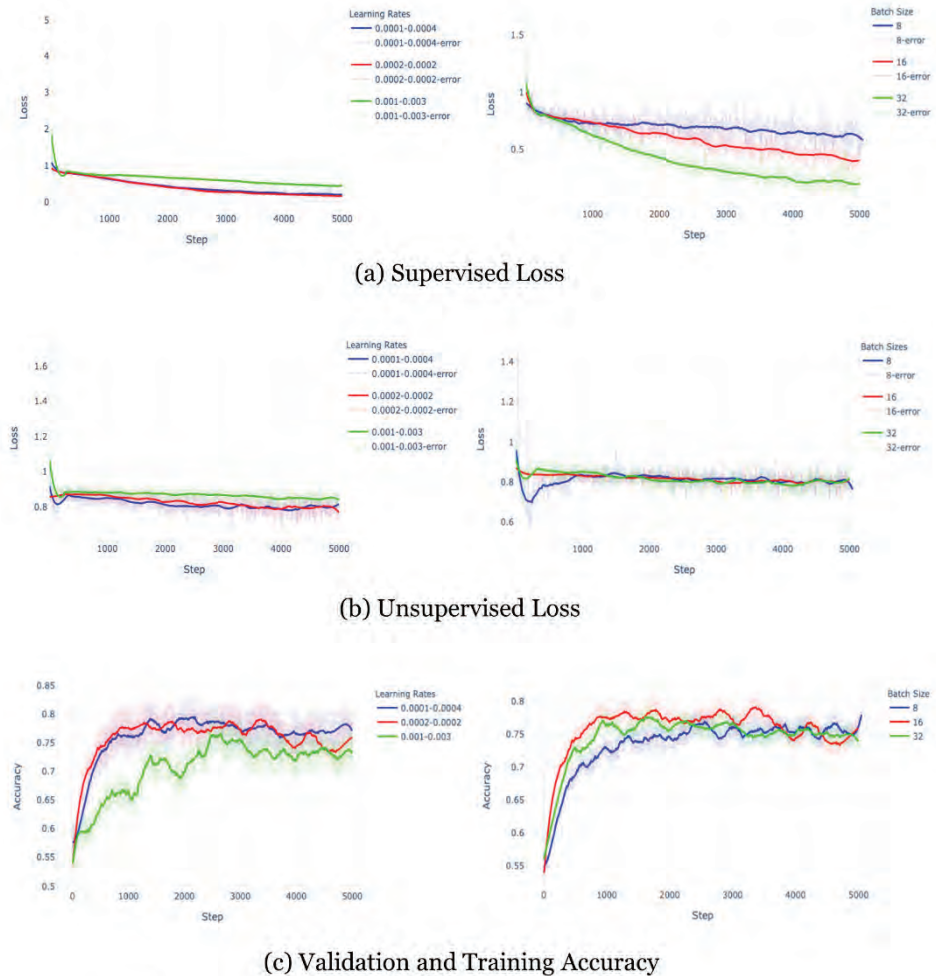
as much as the other two. The same case is also applicable for unsupervised loss, as seen in Figure 3b. While the unsupervised loss for the first two decreases at a stable and steady rate, for the 0.001-0.003 model, during the early iterations, it just does not improve (indicated by the almost horizontal line) and during the later iterations, it decreases at a very undesirably slow rate. As for the validation accuracy, the 0.0001-0.0004 and 0.0002-0.0002 models' validation accuracies are able to increase more stably than that of the 0.001-0.003 model, as seen in Figure 3c. The irregular changes in 0.001-0.003 training possibly indicate the model's inability to converge to the optimal point due to the learning rate being too high. A similar "shooting off" is observable at the later iterations of 0.0002-0.0002 training, possibly indicating that, if we are to choose it for final training, a learning rate scheduler should be used. Although this "shooting off" is not observable in 0.0001-0.0004 training, this training is slower to reach the accuracy that is achieved by the 0.0002-0.0002 model and seems to not improve the later iterations. Thus, we chose to utilize the 0.0002-0.0002 training regime paired with a linear learning rate scheduler starting from a specific iteration until the end, following the work of Miyato et al. (2018), so that the learning rate at the end of the training would equal 0. We speculate that this could overcome the "shooting off" problem of 0.0002-0.0002 training and utilize its faster convergence.

We based our experiments on three batch sizes: 8, 16, and 32. Following the learning rate experiments, we chose the maximum iteration to be 5000 and the learning rates 0.002-0.0002. We did not implement the linear learning rate scheduler for now, as we wanted to see how the batch size selection as it is could provide useful information for the training dynamics.

Looking at the validation accuracy curves in Figure 3c, it is evident that 32 and 16 batch size models are able to converge faster compared to 8 batch size training. This is as expected, since 16 and 32 batch size models see more real images than the 8 batch size model sees in the same iteration interval. However, in later iterations, the validation accuracies of 32 and 16 batch size models decrease steadily (even worse for the 32 batch size regime) to the point that both models perform worse than the 8 batch size model does. We speculate that using a large batch size hurts the performance during the initial training. As the discriminator receives a lot of examples to train on, it might overpower the generator, which then has a negative effect on the overall training. This might cause the generator to stop improving or to generate the same images since the discriminator no longer sends any meaningful feedback. This could, in turn, stop the discriminator improvement, as it is now easier for the discriminator to only memorize the training images and not generalize to new samples. Conversely, 8 batch size training improves steadily across the whole training process, although it is slower than the other two in the beginning. This tells us that in the 8 batch size regime, the training does not suddenly become a "one-sided" match and the discriminator can learn more generalizable and meaningful parameters.

Looking at the discriminator loss curves in Figure 3a and 3b, the overfitting of the 32 batch size training is evident as the extreme decrease of supervised loss does not correspond to the mild unsupervised loss decrease. This indicates that the discriminator is trying to just memorize the labelled images. Although this overfitting case is not as easily observable in 16 batch size training, we think that the decrease of the supervised loss in 16 batch size training is still

“fast-paced” and can cause training instabilities. Thus, we think that with both of its discriminator losses decreasing in a steady but stable way and with its stable validation accuracy curve, 8 batch size is the safer and the more performant choice.

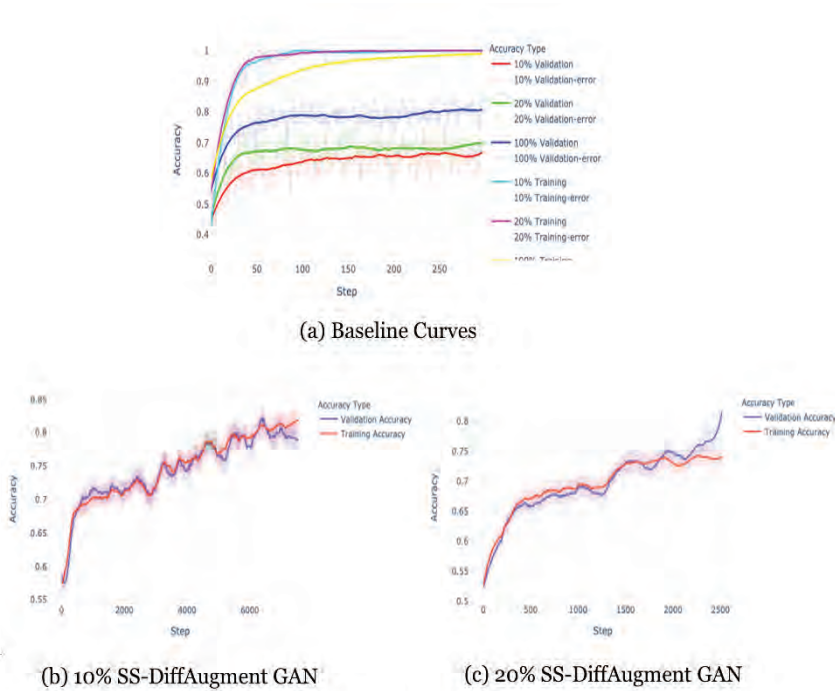


**Figure 3.** Training Curves for Different Learning Rates and for Different Batch Sizes. On the left are the curves for the different learning rates, and on the right are the curves for the different batch sizes. Continuous Error bars are represented by the lightly shaded areas around the curves.

## Training Approach

For our GAN training, we used a batch size of 8 and an Adam optimizer with a learning rate of  $2 \cdot 10^{-4}$  for both the generator and the discriminator ( $\beta_1 =$

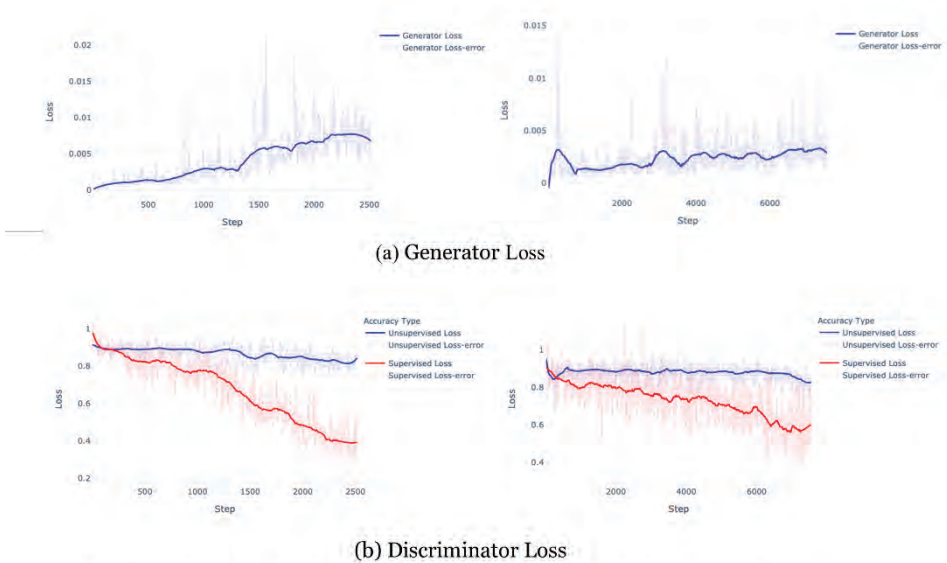
0,  $\beta_1 = 0.9$ ). We used a latent code  $z$  of 128 dimensions. We trained our 20% GAN model for 7500 iterations (generator steps) and employed linear learning rate decay starting from the 5000th iteration. For 10% training, we found that maximum iterations of 7500 cause slight overfitting, which was easily tackled without sacrificing any major performance decreases by setting the maximum iterations to 2500, and by employing the decay of the linear learning rate starting from the 1200<sup>th</sup> iteration. The starting iterations of the linear learning rate decay were set to when the "un-decayed" training's validation curve has its first major plateau.



**Figure 4.** Accuracy Curves for ResNet50 baseline and SS-DiffAugment-GAN. Continuous Error bars are represented by the lightly shaded areas around the curves.

**Table 1.** Percentage accuracies of our GAN approach and the baseline on the test set. For the GAN approach, the complete training dataset was used with only the given percentage of the images with labels, while for the baseline, only the given percentage of the images in the fully-annotated training dataset was used for training. It is evident from the experimental results that our GAN approach is able to outperform the baseline 100% classification modeven when only 10% and 20% of the images are annotated (by 0.61% and 3.71% respectively).

Method	10%	20%	100%
Resnet50 (Baseline)	62.50 ± 0.52	70.83 ± 0.72	73.34 ± 0.37
Semi-Supervised DiffAugment GAN	<b>73.95 ± 0.46</b>	<b>77.05 ± 0.83</b>	-



**Figure 5.** Loss Curves for Different Learning Rates and for Different Batch Sizes. On the left are the curves for 10% GAN training, and on the right are the curves for 20% GAN training. Continuous error bars are represented with the lightly shaded areas around the curves.

For our baseline training, we followed the common transfer learning training schedule on a pre-trained ImageNet model. We used a batch size of 32 and Stochastic Gradient Descent (SGD) with a learning rate of  $1 \cdot 10^{-3}$ . We trained every baseline model (10%, 20%, and 100%) for a maximum epoch of 100. We did not employ any learning rate decay.

## Results and Discussion

Looking at the accuracy curves of every baseline training in Figure 4a, it is evident from the large gap between the validation accuracy and the training accuracy that the model overfitted to the training set. In contrast, looking at the accuracy curves of the SS-Diff-Augment for both 10% and 20% regimes in Figure 4b and 4c, from the similar and close increase in validation and training accuracy curves, we can deduce that the model did not overfit to the training data.

Our SS-DiffAugment-GAN model was able to outperform every baseline model with only 10% or 20% of the labels present in the training data. Looking at the final test accuracies from Table 1, we can see that both the 10% regime and the 20% regime were able to outperform their corresponding baseline models by large differences and the 100% model by 0.61% and 3.71% respectively. The generator losses for both 20% and 10% are increasing just as expected, as seen in Figure 5a. For the discriminator loss, both losses seemed to increase quite steadily with no sudden jumps, as seen in Figure 5b. However, we also observed that the supervised loss decreases faster, possibly due to the fact that the number of

labelled examples (50 and 100 images respectively) is less than the number of unlabelled examples (450 and 500 images respectively). Looking at the slow decrease of unsupervised loss and the steady increases in the accuracy curves, we speculate that training the model for more iterations with the stable training dynamics of the SS-DiffAugment-GAN and a more aggressive learning rate scheduling approach, such as decreasing the learning rate by 0.1 at specific steps, would result in an even better performance.

## 5. Conclusion

We present the SS-DiffAugment-GAN, a semi-supervised classification approach to medical datasets with limited images and few labels, utilizing the recent deep generative model works such as Semi-Supervised Generative Adversarial Networks (Salimans et al., 2016), Spectral Normalization (Miyato et al., 2018), and DiffAugment (Zhao et al., 2020). Our model is able to outperform one of the state-of-the-art classification networks and tackle the data size and annotation problems within medical datasets simultaneously. However, although it is able to outperform or match the existing classification approaches, we do think that there is room for improvement within our work. Since existing semi-supervised approaches were developed on standard computer vision datasets wherein image composition and structure differ with the target annotations and since medical datasets have more structural similarity and redundancy amongst samples, new design methods must be developed for medical datasets' special requirements. In the future work, we would like to further investigate where our model stands amongst other semi-supervised classification models and experiment our model with more complex medical datasets or natural image datasets.

## 6. References

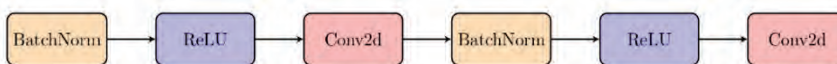
- Arjovsky, Martin, and Léon Bottou. "Towards Principled Methods for Training Generative Adversarial Networks." *International Conference on Learning Representations (ICLR)*, 2017.
- Arjovsky, Martin, et al. "Wasserstein Gan." *Proceedings of the 34th International Conference on Machine Learning, PMLR 70:214-223*, 2017.
- Brock, Andrew, et al. "Large Scale Gan Training for High Fidelity Natural Image Synthesis." *International Conference on Learning Representations*, 2018.
- Chen, Xiaoran, and Ender Konukoglu. "Unsupervised Detection of Lesions in Brain MRI Using Constrained Adversarial Auto-Encoders." *MIDL*, 2018.
- Dai, Zihang, et al. "Good Semi-Supervised Learning That Requires a Bad GAN." *In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*. Curran Associates Inc., Red Hook, NY, USA, pages 6513–6523, 2017.
- DeVries, Terrance, and Graham W. Taylor. "Improved Regularization of Convolutional Neural Networks With Cutout." *ArXiv*, 2017.
- Esteva, Andre, et al. "Dermatologist-Level Classification of Skin Cancer With Deep Neural Networks." *Nature* 542, 2017.

- Goodfellow, Ian J., et al. "Generative Adversarial Networks." *Advances in Neural Information Processing Systems*, 2014.
- Greenspan, Hayit, et al. "Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique." *IEEE Transactions on Medical Imaging*, 35:1153–1159, 05 2016. doi: 10.1109/TMI.2016.2553401.
- Gulshan, Varun, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. 8 *JAMA*, 316(22):2402–2410, 12 2016. ISSN 0098-7484. doi: 10.1001/jama.2016.17216. URL: <https://doi.org/10.1001/jama.2016.17216>.
- He, Kaiming, et al. "Deep Residual Learning for Image Recognition." *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- Hiasa, Yuta, et al. "Cross-Modality Image Synthesis From Unpaired Data Using CycleGAN: Effects of Gradient Consistency Loss and Training Data Size." *Simulation and Synthesis in Medical Imaging. SASHIMI*, 2018.
- Hu, Bo, et al. "Unsupervised Learning for Cell-Level Visual Representation in Histopathology Images With Generative Adversarial Networks." *IEEE Journal of Biomedical and Health Informatics*, 23(3):1316–1328, May 2019. ISSN 2168-2208. doi:10.1109/jbhi.2018.2852639. URL: <http://dx.doi.org/10.1109/jbhi.2018.2852639>.
- Karras, Tero, et al. "A Style-Based Generator Architecture for Generative Adversarial Networks." *CVPR*, 2019.
- Karras, Tero, et al. "Analyzing and Improving the Image Quality of StyleGAN." *CVPR*, 2020.
- Karras, Tero, et al. "Progressive Growing of Gans for Improved Quality, Stability, and Variation." *ICLR*, 2018.
- Karras, Tero, et al. "Training Generative Adversarial Networks With Limited Data." *NeurIPS*, 2020.
- Krause, Jonathan, et al. "Grader Variability and the Importance of Reference Standards for Evaluating Machine Learning Models for Diabetic Retinopathy." *Ophthalmology*, 125(8):1264–1272, Aug 2018. ISSN 0161-6420. doi: 10.1016/j.ophtha.2018.01.034. URL: <http://dx.doi.org/10.1016/j.ophtha.2018.01.034>.
- Krizhevsky, Alex, et al. "Imagenet Classification with Deep Convolutional Neural Networks." In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012. URL: <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>.
- Lecouat, Bruno, et al. "Semi-Supervised Deep Learning for Abnormality Classification in Retina Images." *ArXiv*, 2018.
- Maayan Frid-Adar, Idit Diamant, Eyal Klang, Michal Amitai, Jacob Goldberger, and Hayit Greenspan. "Gan-based Synthetic Medical Image Augmentation for Increased CNN Performance in Liver Lesion Classification." *Neurocomputing*, 321:321–331, Dec 2018. ISSN 0925-2312. doi:10.1016/j.neucom.2018.09.013.
- Madani, Ali, et al. "Semi-Supervised Learning With Generative Adversarial Networks for Chest X-Ray Classification With Ability of Data Domain

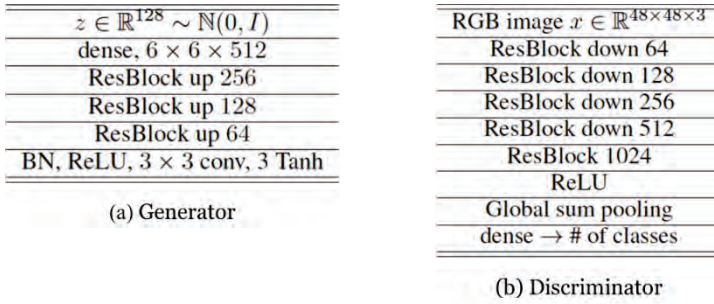


- Adaptation.” In 2018 *IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 1038–1042, 2018.
- Madani, Ali. et al. “Deep Echocardiography: Data-efficient Supervised and Semi-supervised Deep Learning Towards Automated Diagnosis of Cardiac Disease.” *DigitalMedicine*, 1, 12 2018. doi: 10.1038/s41746-018-0065-x.
- Miyato, Takeru, and Masanori Koyama. “cGANs With Projection Discriminator.” *ArXiv*, 2018.
- Miyato, Takeru, et al. “Spectral Normalization for Generative Adversarial Networks.” *ICLR*, 2018.
- Pinto, Diaz, et al. “Retinal Image Synthesis and Semi-Supervised Learning for Glaucoma Assessment.” *IEEE Transactions on Medical Imaging*, 38(9):2211–2218, 2019.
- Radford, Alec, et al. “Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks.” *CoRR*, 2016.
- Salehinejad, Hojjat, et al. “Generalization of Deep Neural Networks for Chest Pathology Classification in X-Rays Using Generative Adversarial Networks.” *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2017.
- Salimans, Tim, et al. “Improved Techniques for Training GANs.” *Advances in Neural Information Processing Systems*, 2016.
- Simard, Patrice Y., et al. “Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis.” In *Seventh International Conference on Document Analysis and Recognition, 2003. Proceedings.*, pages 958–963, 2003.
- Springenberg, Jost Tobias. “Unsupervised and Semi-Supervised Learning With Categorical Generative Adversarial Networks.” *CoRR*, 2015.
- Veeling, Bastiaan S., et al. “Rotation Equivariant CNNs for Digital Pathology.” *Medical Image Computing and Computer Assisted Intervention*, 2018.
- Wolterink, Jelmer M., et al. “Generative Adversarial Networks for Noise Reduction in Low-Dose CT.” *IEEE Transactions on Medical Imaging*, 36(12):2536–2545, 2017.
- Yi, Xin, et al. “Unsupervised and Semi-Supervised Learning With Categorical Generative Adversarial Networks Assisted by Wasserstein Distance for Dermoscopy Image Classification.” *ArXiv*, 2018.
- Zhang, Dan, and Anna Khoreva. “Progressive Augmentation of GANs.” *NeurIPS*, 2019.
- Zhang, Han, et al. “Self-Attention Generative Adversarial Networks.” *Proceedings of the 36th International Conference on Machine Learning*, 2018.
- Zhao, Shengyu, et al. “Differentiable Augmentation for Data-Efficient Gan Training.” *NeurIPS*, 2020.

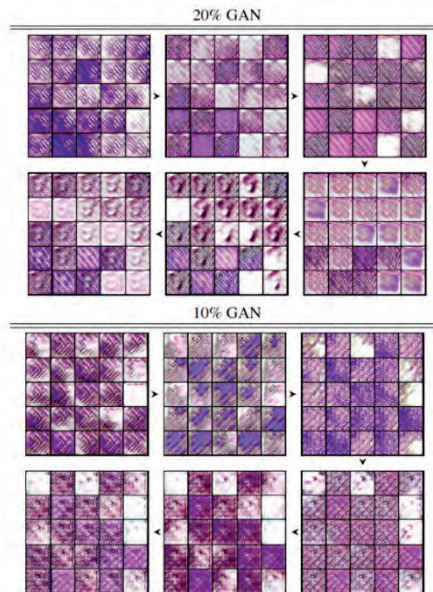
## Appendix: Supplemental Materials



**Figure 6.** *ResBlock Architecture. For Discriminator, however, BatchNorm layers are removed.*



**Figure 7.** ResNet GAN Architectures (Miyato et al., 2018).



**Table 2.** Images generated by the generator from a fixed  $z$  every 20th epoch. Viewed better when zoomed in.



# Search Asymmetry Manipulation with a Common Feature

---

*Maanasi Nair*

**Author Background:** *Maanasi Nair grew up in San Jose, California and then Bangalore, Karnataka, India. She currently attends The International School Bangalore (TISB) in Bangalore, India. Her Pioneer research concentration was in the field of neuroscience and titled "Vision Impossible: An Introduction to the Psychology and Neuroscience of Seeing."*

## Abstract

Search asymmetry has been used widely in vision research in order to understand the way the brain processes visual stimuli. Search asymmetry arises when a simple search, i.e. finding a target A among distractors B, takes longer than its inverse search, i.e. finding target B among distractors A. The mechanism behind the existence of search asymmetries is not fully understood. Existing research proposes theories that remain imprecise and subjective. In order to improve the understanding of search asymmetry, the properties of search asymmetry must be fully known. This study aims to expand on the current knowledge of search asymmetry properties by exploring how a common search asymmetry behaves when a common feature is added equally to both the target and distractor. Participants were asked to perform a series of searches and the time taken to locate the target in each search was recorded. Search asymmetry was quantified by the difference in search time between inverse searches. In the presence of a common feature, not only did participants take longer but also the previously established search asymmetry in the absence of the feature disappeared. This calls into question the conditions currently accepted as the reason behind the existence of search asymmetries.

## Introduction

Much of how vision happens is subconscious and taken for granted. This means that we are often unaware of how we perform basic tasks, even ones that we do all the time. Searching for things is a clear example of this: how do you find that misplaced notebook, or your keys on a cluttered table, or your phone when you realize you don't

have it, just before leaving the house? Answers to these questions can be found under the topic of visual search, a research paradigm that often involves very simple stimuli, like finding a T among a field of Ls, in order to clarify the mechanisms that support search in the visual system.<sup>1</sup>

Research shows that often searching does not work the way we would expect. For example, it makes sense that finding your keys among a bunch of papers scattered on a table is just as easy as finding a paper among a bunch of keys scattered on the table. However, in certain cases this is not so. The search performance can be asymmetric. Finding a key among papers might take more time than finding a paper among keys. Although these two searches both use keys and papers, search asymmetry shows that finding one of those objects can turn out to be harder, meaning it takes more time to locate the target.

In the research literature, the object that is being looked for is called the target. For example, finding the keys means that the key is the target. The objects that surround the target as a form of clutter from which you have to distinguish the target are called distractors, i.e. the papers that fill the desk. The time taken to locate your target from the set of distractors is known as search time.<sup>2</sup>

A visual search asymmetry is said to occur when searching for object A among objects B returns different results as opposed to searching for object B among objects A. In essence, one target appears to be easier to locate.<sup>3</sup> Another example of search asymmetry is locating a magenta spot among blue distractor stimuli. The reverse, finding blue among magenta distractors, takes a longer mean time.<sup>4</sup>

Previous research has attempted to understand the mechanisms behind search asymmetries. However, no concrete answer has been suggested, which means that the cause of asymmetries remains uncertain. Treisman and Gormican suggested some potential explanations:

- a) Searching for the absence of a feature is slower than searching for the presence.
- b) Searching for less of a feature is slower than searching for more of that feature.
- c) Searching for prototypical features is slower than searching for non-prototypical features<sup>5</sup>.

These statements are not satisfying because they denote a level of subjectivity in the terms used to define the condition for a search asymmetry. Primarily, these statements reference a feature. While there have been attempts to create a list of features that cause a search asymmetry, the cause still remains unclear. It is still not known what trait of a target leads to the formation of a distinguishable feature. Since

---

<sup>1</sup>"Learning in Visual Search". *APA*, 2020, <https://www.apa.org/pubs/highlights/peeps/issue-27>.

<sup>2</sup> Bruce, Neil D. B., and John K. Tsotsos. "Visual Representation Determines Search Difficulty: Explaining Visual Search Asymmetries". *Frontiers In Computational Neuroscience*, vol 5, 2011. *Frontiers Media SA*, doi:10.3389/fncom.2011.00033.

<sup>3</sup>*Search.Bwh.Harvard.Edu*, 2020, <https://search.bwh.harvard.edu/new/pubs/IntrotoSearchAsym.pdf>.

<sup>4</sup> Nagy, Allen L, and Scott M Cone. "Asymmetries in Simple Feature Searches for Color". *Vision Research*, vol 36, no. 18, 1996, pp. 2837-2847. *Elsevier BV*, doi:10.1016/0042-6989(96)00046-6.

<sup>5</sup>Treisman, Anne, and Stephen Gormican. "Feature Analysis in Early Vision: Evidence from Search Asymmetries". *Psychological Review*, vol 95, no. 1, 1988, pp. 15-48. *American Psychological Association (APA)*, doi:10.1037/0033-295x.95.1.15.

this definition is ambiguous, the absence or presence of a feature can be interpreted in multiple ways. For example, in a search asymmetry paradigm containing Os and Cs, the gap in the circular shape could be seen as the defining feature. This would indicate that the O has the absence of the feature. However, it could also be said that O has a complete boundary to form a closed shape, hence it has the presence of that feature. In this case, C would be seen to have the absence of the closed boundary feature. Consequently, the direction of the asymmetry (ascertaining which target is actually harder to find) is defined differently, based on the way the presence or absence of a feature is evaluated. This leads to subjectivity and variability.

In order to narrow down the number of possible explanations for search asymmetries, it is important to understand the limitations and the circumstances that cause search asymmetry to manifest itself differently. Specifically, this study looks at the effect that the addition of a common feature to both the target and the distractor has on the search asymmetry between the two reverse searches.

The conclusions of the existing theories emphasize the presence of a feature to define the direction of an asymmetry. Hence, as long as the target has the “pop up” feature, the paired searches should result in a search asymmetry. Adding a common feature equally to both the target and the distractor does limit the effect of this differentiating feature in the target. Therefore, regardless of whether the common feature is present or not, the search asymmetry would be expected to remain the same.

However, the hypothesis of this study is that the equal addition of a common feature to both the target and the distractor would actually diminish the search asymmetry.

Search asymmetry can be quantified as the difference in search time for each search. A search asymmetry would be present if there was a difference in search time between finding keys among papers and finding a paper among keys. In order to investigate the effect of the addition of a common feature, a classic, simple asymmetry paradigm would be replicated to achieve the original result of establishing the search asymmetry. Then, a simple additional feature would be added equally to both objects, and the search paradigm would be conducted similarly, allowing for a comparison of the difference in search time between both pairs of searches.

Understanding search asymmetry is integral to the study of visual attention and other fundamental mechanisms of vision. For example, studying asymmetries can shed light on whether the visual system is coded on the basis of predetermined features or how the visual system processes dimensions<sup>6</sup>. Therefore, it is important to uncover the reality of the mechanism behind this phenomenon in order to further our understanding of vision accurately.

## Methods

Participants were asked to conduct a set of four searches continuously, repeated twice. For each search, there was only one target and 30 distractors. Both target and distractors were randomly oriented, randomly aligned and randomly positioned. The time taken for the participant to locate the target once the pattern was displayed was recorded in seconds. Consequently, one run for one participant would result in the

---

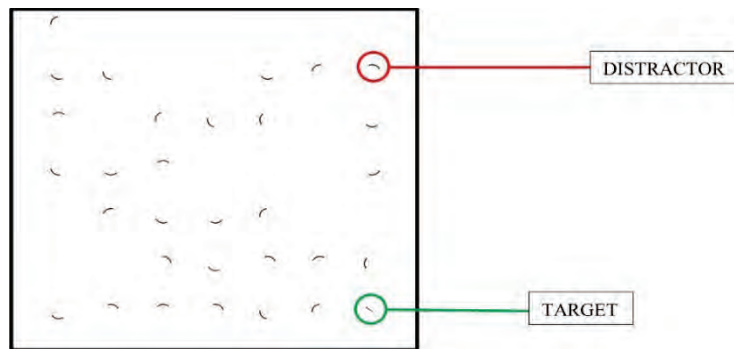
<sup>6</sup> Rosenholtz, Ruth. “Search Asymmetries? What Search Asymmetries?” *Perception & Psychophysics*, vol. 63, no. 3, 2001, pp. 476–489., doi:10.3758/bf03194414.

return of four time intervals. In this study, the set size remained constant as efficiency was not measured. 100% of the assigned scenarios contained the target. The total number of participants for this data collection was 24.

The first two searches were established to produce a search asymmetry; hence they were simply replicated for the participant.<sup>7</sup> The curve-line paradigm of search asymmetry was used as the base of this experimental model.

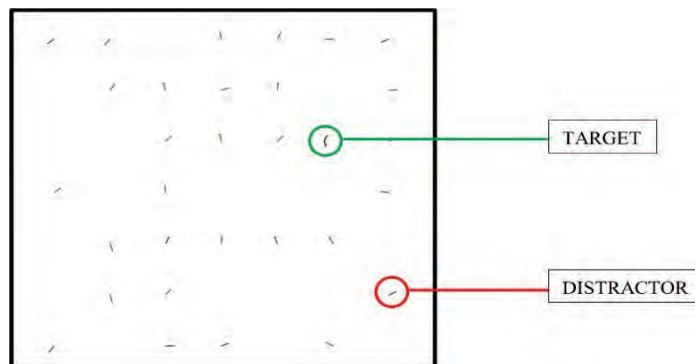
Note: The border seen in the following examples was not included in the actual implementation, as it was projected in an application window on a laptop.

Search 1: To find a single straight line in any orientation or direction among many curved lines in any orientation or direction.



**Figure 1.** Search 1: line among curves

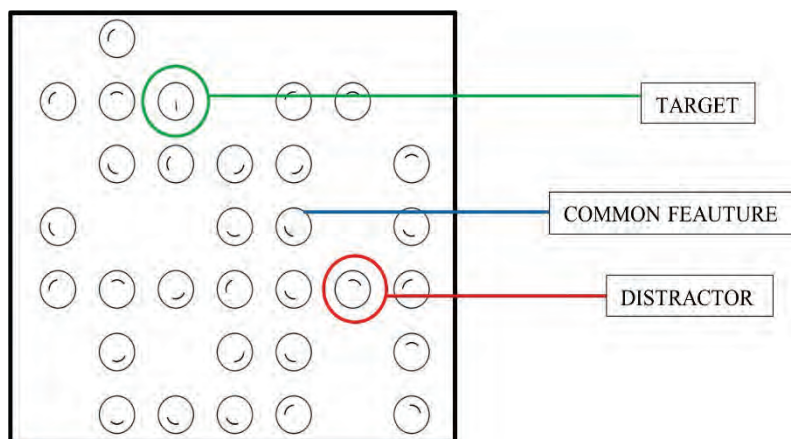
Search 2: To find a single curve in any orientation or direction among many straight lines in any orientation or direction.



**Figure 2.** Search 2: curve among lines

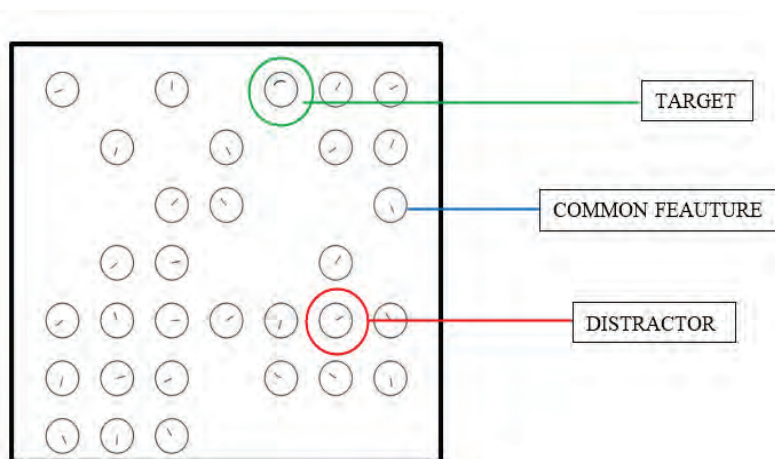
<sup>7</sup>Li, Z. "Contextual Influences in V1 As A Basis For Pop Out and Asymmetry in Visual Search". *Proceedings Of The National Academy Of Sciences*, vol 96, no. 18, 1999, pp. 10530-10535. *Proceedings Of The National Academy Of Sciences*, doi:10.1073/pnas.96.18.10530.

Search 3: To find a straight line enclosed in a circle in any orientation or direction among many curves each enclosed in circles in any orientation or direction.



**Figure 3.** Search 3: line among curves with common feature

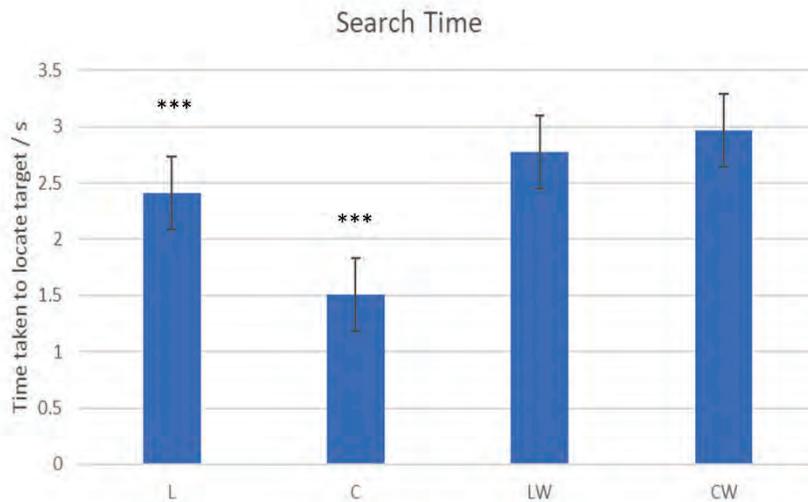
Search 4: To find a curve enclosed in a circle in any orientation or direction among many straight lines each enclosed in circles in any orientation or direction.



**Figure 4.** Search 4: curve among lines with common feature

## Results

A total of 12 people were asked to participate in this study, completing all four searches twice. Hence, for each search, 24 search times were obtained. These were averaged to produce the results shown in the graph below.



**Figure 5.** Bar graph showing search time in seconds for all 4 searches. Significant difference between bars marked with \*\*\* is  $p < 0.05$ ). Error bars are standard errors of the mean ( $n=24$ ). Search conditions are as follows: L- target line among curves, C- target curve among lines, LW- target line with circles among curves with circles, CW- target curve with circles among lines with circles.

## Analysis

Search asymmetry can be quantified as the difference in the time taken to locate the target in two seemingly equivalent tasks, i.e. finding a line among curves vs. finding a curve among lines. A large search asymmetry would indicate a larger difference between the two calculated search times. A reduced search asymmetry would indicate a reduced difference between the two, regardless of the absolute value of the search time for each condition.

These four searches consist of two pairs of opposite searches. The first two, L and C, refer to first finding a line among curves and then vice versa, respectively. As is shown on the graph, finding the curve took significantly less time than finding the line, establishing the presence of a search asymmetry between these two conditions.

Similarly, the second pair of situations was also designed as equivalent opposites that produced a search asymmetry. The only difference the second pair held, when compared to the first, is the presence of a common feature, an added equal component to each individual distractor or target. In this particular case, it was the presence of a circle around the line or curve. Fundamentally, it was the same search, and hence, should have produced a similar search asymmetry, quantified by the difference in search time.

Three main observations can be drawn in relation to the second pair of searches, LW and CW, which involved finding a line with circles and finding a curve with circles respectively.



1. The search time for both LW and CW was much greater than their counterparts L and C. The introduction of a common feature made the search for both cases harder, leading to the necessity of more time to find both targets. It is important to note that this occurred for both searches that included the common circle and was not partial to a certain target. Therefore, while it does affect the search time greatly, this alone does not alter the possibility of a search asymmetry.
2. The difference between the time taken to locate the target in LW and CW was considerably less, as compared to the difference between the times for L and C. As the difference between the search times is a measure of the search asymmetry, it can be noted that the asymmetry was reduced in the second pair of searches. As the only difference between the two pairs was the addition of the common feature, i.e. the circle, the reduction in search asymmetry can be correlated to the equal addition of a common feature to both the target and the distractor.
3. It appears that the search asymmetry was not only reduced, but also flipped in direction. In the regular pair, L and C, as noted in the literature, finding a curve among lines was easier and therefore quicker than finding a line among curves. This was shown in the results above, as the search time for C was less than it was for L. However, in the case of LW and CW, this seemed to reverse. The average search time taken for CW was in fact slightly greater than the time for LW, the opposite of what had been established and shown without the common feature.

In order to understand whether the differences depicted between the search times for each search scenario is significant, a statistical analysis using the t test was used to determine the data's reliability. The difference between L and C is significant,  $t(22) = 2.64$ ,  $p < 0.05$ . The difference between LW and CW is not significant,  $t(22) = -0.38$ ,  $p < 0.05$ .

## Discussion

When asked to search for a line among a field of curves, participants were much slower than when they were asked to search for the opposite, a curve among lines. As expected, this visual search produced a search asymmetry, due to the difference in speed when locating the target. In the second pair of searches, significantly, the participants took longer for both searches that included the circles than for their counterparts without the circles, and also did not take longer for either the line or the curve as the target. The search asymmetry was no longer present.

This study sought to test the commonly accepted theories about asymmetries by applying them to an unconventional situation: the addition of a common feature. However, the results differed from the predicted result.

Suppose that searching for the absence of a feature is slower than searching for its presence. Adding the common feature does not detract from the presence of the previously mentioned feature, whatever it may be, from the target that creates an

easier search. In this case, adding the circle around both the curve and line does not remove any differences between the curve and line itself. Therefore, by this reckoning, the search asymmetry should have been preserved.

Suppose that searching for less of a feature is slower than searching for more of that feature. The addition of a common feature does not detract from the nature of the distinguishing feature. It simply acts as additional and irrelevant information that should not affect the search, other than perhaps prolonging the time taken to find the target for both searches equally. Hence, even with this condition, the search asymmetry should have been preserved.

Suppose that searching for prototypical features is slower than searching for non-prototypical features. Similar to the first condition, the addition of a common feature does not tamper with the effect that the prototypical and non-prototypical features may have on the search asymmetry. The common feature does not change the conformation of the existing objects in the search, but simply adds to it.

The feature analysis model suggests that a larger pop-out created by greater differences between target and distractor should result in smaller or absent search asymmetries. However, this experiment shows the opposite effect. The added common feature reduced the 'pop-out', and also had a minimalized asymmetry in search times. This is also seen in asymmetry in searches between color pairs (Nagy and Conn)<sup>8</sup>. These results also suggest the need for an alternative theory.

The results of this experiment therefore pose a challenge to the current understanding of search asymmetries. Following logical reasoning, if  $A > B$ , then  $A+C > B+C$ . However, the results show that even though  $A > B$ ,  $A+C$  is in fact essentially equal to  $B+C$ . Therefore, there is a fallacy in the theory that predicts the creation of a search asymmetry and it needs to be revised to incorporate this behavior of asymmetries with common features.

This investigation indicates that search asymmetries are a complex phenomenon that arises due to more than just one distinguishing feature. Instead, the results suggest that the nature of the search is a culmination of the interactions of many features that determine how similar the target and the distractor are.

## References

- Bruce, Neil D. B., and John K. Tsotsos. "Visual Representation Determines Search Difficulty: Explaining Visual Search Asymmetries." *Frontiers in Computational Neuroscience*, vol. 5, 2011. Frontiers Media SA, doi:10.3389/fncom.2011.00033.
- Dosher, Barbara Anne et al. "Parallel Processing in Visual Search Asymmetry." *Journal of Experimental Psychology: Human Perception And*

---

<sup>8</sup> Nagy, Allen L, and Scott M Cone. "Asymmetries in Simple Feature Searches for Color." *Vision Research*, vol. 36, no. 18, 1996, pp. 2837–2847., doi:10.1016/0042-6989(96)00046-6.

- Performance*, vol 30, no. 1, 2004, pp. 3-27. American Psychological Association (APA), doi:10.1037/0096-1523.30.1.3. Accessed 29 June 2020.
- "Learning in Visual Search." American Psychological Association. <https://www.apa.org/pubs/highlights/peeps/issue-27>.
- Li, Z. "Contextual Influences in V1 As A Basis for Pop Out and Asymmetry in Visual Search." *Proceedings of the National Academy Of Sciences*, vol. 96, no. 18, 1999, pp. 10530-10535. doi:10.1073/pnas.96.18.10530.
- Nagy, Allen L, and Scott M Cone. "Asymmetries in Simple Feature Searches for Color". *Vision Research*, vol. 36, no. 18, 1996, pp. 2837-2847. Elsevier BV, doi:10.1016/0042-6989(96)00046-6.
- Rosenholtz, Ruth. "Search Asymmetries? What Search Asymmetries?" *Perception & Psychophysics*, vol. 63, no. 3, 2001, pp. 476-489., doi:10.3758/bf03194414.
- Saiki, J. "Features Underlying Visual Search Asymmetry Revealed by Classification Images." *Journal of Vision*, vol. 5, no. 8, 2005, pp. 780-780. Association for Research In Vision And Ophthalmology (ARVO), doi:10.1167/5.8.780. Accessed 29 June 2020.
- Treisman, Anne, and Janet Souther. "Search Asymmetry: A Diagnostic for Preattentive Processing of Separable Features." *Journal of Experimental Psychology: General*, vol 114, no. 3, 1985, pp. 285-310. American Psychological Association (APA), doi:10.1037/0096-3445.114.3.285.
- Treisman, Anne, and Stephen Gormican. "Feature Analysis in Early Vision: Evidence from Search Asymmetries." *Psychological Review*, vol. 95, no. 1, 1988, pp. 15-48. American Psychological Association (APA), doi:10.1037/0033-295x.95.1.15.
- Wang, L. et al. "Asymmetry in Visual Search for Local Biological Motion Signals". *Journal of Vision*, vol. 10, no. 7, 2010, pp. 782-782. Association for Research In Vision And Ophthalmology (ARVO), doi:10.1167/10.7.782.
- Wolfe, Jeremy. Asymmetries in Visual Search: An Introduction. *Perception & Psychophysics*, vol. 63, no. 3, pp. 381-389. <https://search.bwh.harvard.edu/new/pubs/IntrotoSearchAsym.pdf>.



# **Complicated Implications of Simple Images: Schizophrenic Patients' and Their Families' Views on MRI and fMRI in the United States from 2000-2020**

---

*Mengyao (Angel) Yang*

**Author Background:** *Mengyao (Angel) Yang grew up in China and currently attends San Domenico High School in San Anselmo, California in the United States. Her Pioneer research concentration was in the field of history/science, technology, and society and titled "History of Medical Technologies."*

## **Introduction**

Affecting about 1% of the US population, or roughly 3 million people in the US, schizophrenia is a chronic mental disorder usually first detected in early adulthood. According to the World Health Organization, schizophrenia is characterized by distortions in perceptions, emotions and the self-image. Typical symptoms of schizophrenia include hallucination, delusion, abnormal behavior, disorganized speech, and disturbance of emotions.<sup>1</sup> Because of the severe symptoms schizophrenic patients display, their sudden appearance in early adulthood, and a variety of stigmas about this disorder, families of schizophrenic patients are highly involved in activism and advocacy groups. Historically, schizophrenia was thought to be a degenerative disease, as influenced by German psychiatrist Emil Kraepelin, who categorized two types of psychosis: 'manic depression' and 'dementia praecox' (later replaced by schizophrenia).<sup>2</sup> According to Psychiatric Professor Peter Falkai and his colleagues, due to the historical perception of schizophrenia, "the interest in unravelling the neurobiological basis of schizophrenia has a long and distinctive history."<sup>3</sup>

Neuroimaging technologies such as MRI and fMRI were introduced to psychiatry in the late 20th century, enabling more advanced and safer research for schiz-

---

<sup>1</sup> "Schizophrenia," WHO, accessed Aug 24, 2020, <https://www.who.int/news-room/factsheets/detail/schizophrenia>.

<sup>2</sup> Martha E. Shenton, Thomas J. Whitford, and Marek Kubicki, "Structural Neuroimaging in Schizophrenia from Methods to Insights to Treatments," *Dialogues in Clinical Neuroscience*, vol. 12, 3 (September 2010): 317, <https://doi.org/10.31887/DCNS.2010.12.3/mshenton>.

<sup>3</sup> Peter Falkai, Andrea Schmitt, and Nancy Andreasen, "Forty Years of Structural Brain Imaging in Mental Disorders: Is it Clinically Useful or Not?" *Dialogues in Clinical Neuroscience*, vol. 20, 3 (September 2018): 182, <https://doi.org/10.31887/DCNS.2018.20.3/pfalkai>.

ophrenia. Psychiatrist and radiologist Dr. Martha E. Shenton and her colleagues indicate that “prior to the advent of MRI, brain abnormalities in schizophrenia were based on crude measurements,” some of which are invasive.<sup>4</sup> As neuroimaging technologies visually portray the brain structure and activity, they provide a seemingly solid demonstration of mental disorders. According to bioethics scholar Emily L. Borgelt and her colleagues, neuroimaging emphasizes the biomedical model of mental illness, replacing “the maladies of the soul” with brain abnormalities.<sup>5</sup> Historian Joseph Dumit argues that the biomedical model categorizes people as normal and abnormal as scientists and media reduce the description of brain images to often one-word labels such as *normal*, *depressed*, *schizo*. Dumit elucidates that by employing these one-word labels, authorities such as researchers and magazines convey to the public that different types of brain parallel “different kinds of human.”<sup>6</sup> Despite the danger of these simple categorizations of individuals, many people, particularly families of mentally ill patients, believe that neuroimaging technologies are able to alleviate stigmas associated with mental disorders by offering a biological basis to these disorders. For instance, the National Alliance on Mental Health (NAMI), one of the best-known advocacy organizations in the mental health community in the US, claims on its website that “Identifying brain activity associated with depression ... will help to destigmatize the illness, a disease of the brain.”<sup>7</sup> The paradox of the power of neuroimaging technologies in creating simplistic labels while attacking existing stigmas prompts the question: how do mentally ill patients and their families perceive the capabilities and limitations of neuroimaging technologies in the modern US?

Focusing on the schizophrenic community, this paper will center on the way Magnetic Resonance Imaging (MRI) technology and its counterpart functional MRI (fMRI) communicate the biomedical model to the public and their roles in affecting the public’s understanding of schizophrenia. By examining surveys, personal accounts, and messages delivered by American advocacy groups from 2000 to 2020, this study will show that schizophrenic patients and their families enthusiastically embrace the biomedical model as they perceive neuroimaging’s potential to redefine schizophrenia as a brain disease.

## The Power of Brain Imaging

Although among the most recent medical technology tools, MRI and fMRI have rapidly become influential in both medicine and the general public by providing a new approach to the body and promoting the biomedical model of mental illness. The technological innovation of MRI and fMRI, to a great extent, facilitates scientific research of mental disorders despite their limited application in clinical settings for

---

<sup>4</sup> Shenton, Whitford, and Kubicki, “Structural Neuroimaging in Schizophrenia,” 321.

<sup>5</sup> Emily L. Borgelt, Daniel Z. Buchman and Judy Illes, “Neuroimaging in mental health care: voices in translation,” *Frontiers in Human Neuroscience*, vol. 6, no. 293 (October 2012): 1, <https://doi.org/10.3389/fnhum.2012.00293>.

<sup>6</sup> Joseph Dumit, “Traveling Images, Popularizing Brains,” in *Picturing Personhood: Brain Scans and Biomedical Identities* (Princeton: Princeton University Press, 2004), 141-142.

<sup>7</sup> Dumit, “Traveling Images,” 155.

schizophrenia. Magnetic Resonance Imaging (MRI), previously called Nuclear Magnetic Resonance (NMR), originated in the work of Austrian physicist Wolfgang Pauli, who in 1924 suggested that some conditions can trigger an atom's property called 'spin,' which allows atoms to react to a magnetic field. In the 1970s, Paul C. Lauterbur and Peter Mansfield accomplished groundbreaking work that enabled the production of MRI images. By measuring the time for which hydrogen atoms release energy in a specific part of the human body, MRI captures the information of the soft tissues. Figure 1 shows an MRI image that depicts the soft tissues of the brain. The ability to picture soft tissues is innovative as it addresses the parts that the X-ray and the CT scanner cannot demonstrate. The image of soft tissues can portray the structure of the brain, thus rendering MRI particularly suitable for brain disease diagnosis. Moreover, since MRI employs strong magnetic fields, it does not have the kinds of harmful side effects that X-ray and CT scans have as a result of their use of ionizing radiation.<sup>8</sup> MRI's functional counterpart even further advanced the progress of medical imaging. Dr. Seigi Ogawa, a physicist at the AT&T Bell Laboratory in New Jersey, invented functional MRI in the early 1990s. Unlike MRI, which provides a view of the anatomical structure of the brain, fMRI captures the brain 'in action.' Taking advantage of the fact that neuron activities consume oxygen, fMRI tracks the oxygen level in particular parts of the brain and creates an image of the brain with lighted regions to show brain activity in those regions (see Figure 2).<sup>9</sup> With its ability to display brain activity in different regions, fMRI helps researchers study the functions of different parts of the brain. Researchers frequently use MRI and fMRI to explore the pathology of mental disorders, since they allow researchers to make connections between symptoms of mental illness and brain structure or function in living subjects.<sup>10</sup>

In clinical settings, MRI and fMRI are useful assistance for the classification, diagnosis, and prognosis of psychiatric disorders. However, the neuropathology for many mental disorders still remains unknown. STS professor Erik Parens and bioethics research scholar Josephine Johnston reveal that "different individuals can exhibit 'the same behavior' by recruiting different brain areas or different brain circuits or recruiting the same areas or circuits but to different degrees ... [or] to initiate different behaviors in different individuals."<sup>11</sup> This complicated relationship between different parts of the brain and the behaviors people manifest renders it challenging to interpret MRI and fMRI scans for diagnosis. Psychiatric Professor Peter Falkai and his colleagues refer to a meta-analysis, concluding that after 40-years of research of structural brain, researchers still cannot identify a diagnostic biomarker for schizophrenia. Though brain scans cannot diagnose schizophrenia by themselves, they help physicians rule out other brain diseases or psychotic illnesses in the diagnostic process.<sup>12</sup>

---

<sup>8</sup> Bettyann Holtzmann Kevles, *Naked to the Bone: Medical Imaging in the Twentieth Century* (New Jersey: Rutgers University Press, 1997), 173-198.

<sup>9</sup> Erik Parens and Josephine Johnston, "Neuroimaging: Beginning to Appreciate Its Complexities," *The Hastings Center Report*, Vol. 44, No. 2 (March-April 2014): S3, [https://doi:10.1002/hast.293](https://doi.org/10.1002/hast.293).

<sup>10</sup> "Scanning the Brain," American Psychological Association, accessed August 20, 2020, <https://www.apa.org/action/resources/research-in-action/scan>.

<sup>11</sup> Parens and Johnston, "Neuroimaging," S4.

<sup>12</sup> Falkai, Schmitt, and Andreasen, "Forty Years," 183.



**Figure 1.** (Left) An MRI brain scan showing the anatomical structure of the brain. It is included in an article from November 5, 2013 on the Neuroscience News website for illustrative purposes.<sup>13</sup>

**Figure 2.** (Right) This is a fMRI brain scan showing brain activation. It is from an online article published on July 20, 2017 on *The Harvard Gazette*.<sup>14</sup>

MRI and fMRI not only advance the medical imaging industry with their innovative approaches to create images, but also concentrate the public's attention on brain science by promoting the biomedical model, which explains mental illnesses as biological problems. This explanation is readily accepted by the public, which has a long history of fascination with self-explanatory images of the human body. Wilhelm Conrad Roentgen's discovery of X-rays in 1895 opened an era of medical imaging. According to STS professor Lisa Cartwright, the introduction of X-ray images made them instantly popular among the public. Women, fascinated with the beauty of the X-ray of Roentgen's wife's hand, asked to be photographed with the X-ray, despite harmful side effects of radiation.<sup>15</sup> The public's obsession with images of the interior of human bodies, to some degree, explains the current interest in brain images.

In the case of neuroimaging, the public's fascination with inner body images, along with the way media and researchers present brain imaging to the public, grants brain images agency to convey information unaided by elaborate explanation. Semiotologist Roland Barthes proposes that, historically, the text had dominated the message of visual forms, but this relationship could be reversed in some conditions, when "images no longer illustrates the words ... [and] the words ... are parasitics on the image."<sup>16</sup> Speaking of PET<sup>17</sup> images, Dumit proposes that "scientific visualizations, such as PET images, participate in this reversal of veridictory authority."<sup>18</sup> Dumit identifies the self-explanatory power of brain images and suggests that this power may have arisen in the process during which brain images were transferred from the scientific province to the popular arena. An article on the website of the American

<sup>13</sup> Paul Gabrielsen, "Quantity, Not Just Quality, in New Brain Scan Method," *Neuroscience News*, Nov. 5, 2013, accessed Sep. 3, 2020, <https://neurosciencenews.com/neuroimaging-mri-quantify-brain-volume-571/>.

<sup>14</sup> Sue McGreevey, "Finding signs of life when it matters most," *The Harvard Gazette*, Jan. 20, 2017, accessed Sep. 3, 2020, <https://neurosciencenews.com/neuroimaging-mri-quantify-brain-volume-571/>.

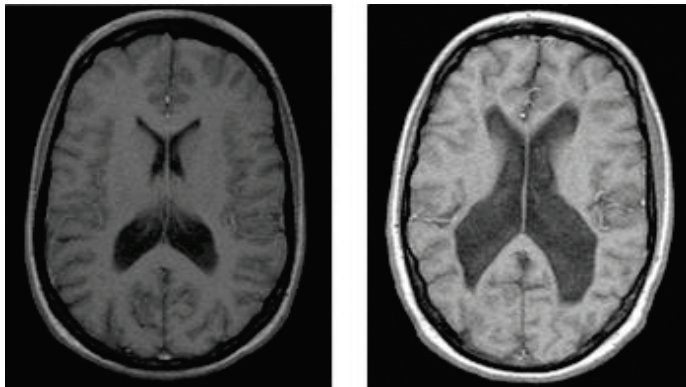
<sup>15</sup> Lisa Cartwright, "Decomposing the body: X Rays and the Cinema," in *Screening the Body: Tracing Medicine's Visual Culture* (Minneapolis: University of Minnesota Press, 1995), 115-119.

<sup>16</sup> Dumit, "Traveling Images," 144.

<sup>17</sup> PET stands for protons emission technology, another neuroimaging technology.

<sup>18</sup> Dumit, "Traveling Images," 144.

Psychological Association (APA), a professional organization designed to educate the public about psychology, illustrates this process. The article explains, “neuroimaging ... shed light – literally – on the human brain.”<sup>19</sup> Utilizing the symbol of light, the APA portrays neuroimaging as the representation of progress and the hope for solving brain-related problems. The APA’s article implies rhetorical meanings to brain images in addition to the scientific information they present. Besides implementing rhetors in the description, Dumit argues that scientific magazines also employ brain scans with evident distinctions and include understandable but shallow captions to present research results. The caption in its original article for Figure 3 – “Healthy twin (left) compared with twin with schizophrenia (right)” – illustrates the application of this compare and contrast technique.<sup>20</sup> On the one hand, the images are discernibly different. On the other hand, the caption establishes multiple contrasts: left, in comparison to right, and healthy, contrasting with schizophrenia. These facile explanations of brain scans reduce the complexity of mental disorders and create categories of different types of people, such as “*novice, practiced, depressed, healthy, retarded,*” as mentioned by Dumit.<sup>21</sup>



**Figure 3.** MRI brain scans of two twins from a journal article in *Advances in Psychiatric Treatment*. According to its original caption, the left image is the healthy brain, and the right one is from a schizophrenic patient.<sup>22</sup>

Dumit explicates the process by which simple, sometimes one-word, captions for brain scans can effectively implant in the public’s perception that differences in brain structure cause mental disorders and ultimately lead to different types of people. American comics theorist Scott McCloud analyzes the cartoon’s power in shaping people’s understanding of their identities because it is “a form of amplification through simplification that focuses our attention on an idea.”<sup>23</sup> Quoting Scott

<sup>19</sup> American Psychological Association, “Scanning the Brain.”

<sup>20</sup> Ibid.

<sup>21</sup> Dumit, “Traveling Images,” 139.

<sup>22</sup> James Woolley and Philip McGuire, “Neuroimaging in schizophrenia: what does it tell the clinician?” *Advances in Psychiatric Treatment*, vol. 11, issue 3 (May 2005), <https://doi.org/10.1192/apt.11.3.195>.

<sup>23</sup> Dumit, “Traveling Images,” 145.



McCloud, Dumit equates simplistic explanations of brain images to cartoons, suggesting that simplified captions of brain images also concentrate people's focus on the notion that fundamentally and biologically, there are different types of humans. That simplified captions leave more room for interpretation and imagination also buttresses this notion of different kinds of human by making brain images "more universal (generalizable to human nature)," according to Dumit.<sup>24</sup> Philosopher Angela K. Thachuk agrees with Dumit's argument, claiming that the biomedical model of mental illness further stigmatizes mentally ill patients.<sup>25</sup> Thachuk's and Dumit's concerns about the biomedical model reflect the power of brain imaging in swaying the public's understanding of mental illness.

The particular form and content of neuroimaging may further consolidate its power among the public. James Gorman, a volunteer of the Human Connectome Project James (which aims to establish a complete understanding of the structures and functions of the brain), shares his thoughts on MRI imaging in a *New York Times* article on January 7, 2014:

My genome may be easier to read, but a GATTACA database doesn't feel as intimate as a picture. A video of a colonoscopy would be too intimate by a long shot, and who wants to find himself in his gut, even if the microbiome is the big deal it seems to be?<sup>26</sup>

According to Gorman, compared to genomes and colonoscopy, brain images possess the proper amount of intimacy, thus rendering them more acceptable than other visual representations. Gorman also suggests that he intends to "find himself" in brain images, implying that brain images tend to demonstrate one's individuality. This appeals to the public's curiosity and desire to understand themselves. Brain scans produced by MRI and fMRI are convincing to the general public as they seem to be straightforward and satisfy the public's fascination with biomedical images and the search for self. Simple descriptions and the rhetors of the brain scans foster the notion among the general public that the spiritual self, including mental conditions, is embedded in the biological brain.

## Redefining Schizophrenia

Physicians and psychiatrists perceive that the advent and the advance of MRI technology led to huge progress in understanding brain abnormalities and neuropathology of psychiatric disorders. Reviewing more than 193 MRI findings from 1988 to 2000, Dr. Martha E. Shenton and her team state that the question of schizophrenia shifted from "Are there brain abnormalities in schizophrenia?" to "Are the brain abnormalities observed due to events that occurred pre- or perinatally (i.e. have a neu-

---

<sup>24</sup> Dumit, "Traveling Images," 147.

<sup>25</sup> Angela K. Thachuk, "Stigma and the Politics of the Biomedical Models of Mental Illness," *The International Journal of Feminist Approaches to Bioethic*, Vol. 4, No. 1 (Spring 2011): 151, <https://www.jstor.org/stable/10.2979/intjfemappbio.4.1.140>.

<sup>26</sup> James Gorman, "A Search for Self in a Brain Scan: A picture of your brain may hold your..." *New York Times* (1923-Current file), Jan 7, 2014, D3.

rodevelopment origin), are they progressive (neurodegenerative), or are they a combination of both?”<sup>27</sup> She also argues that fMRI facilitates “one of the most important changes ... [in the 1990s, which is moving] from examining isolated brain regions to examining interconnected neural networks.”<sup>28</sup> According to Dr. Shenton, MRI and fMRI appear to enable the continuous refinement of the definition of schizophrenia, first confirming schizophrenia as a brain disease and then probing its more complex nature. Examining the clinical application of neuroimaging, radiologist Dr. Xiaoqi Huang and his colleagues conclude that MRI and fMRI results “provide the basis for a major step forward towards the translational use of psychiatric imaging for” diagnosis and monitoring treatments, but more clinical studies of MRI and fMRI are needed.<sup>29</sup>

Unlike physicians and psychiatrists, patients and relatives have limited expertise in the scientific understanding of schizophrenia. They still focus on one of the earlier questions of schizophrenia in medicine, advocating to reclassify schizophrenia as a brain disease. This section primarily concentrates on patients’ and their families’ perspectives on the reclassification of schizophrenia in the recent two decades. It aims to show neuroimaging’s prevalent presence in the way in which patients and families are drawn to the biomedical model of mental illness. SARDAA (Schizophrenia and Related Disorders Alliance of America), an advocacy organization aiming to improve the lives of patients with schizophrenia and related disorders, launched the Brain Campaign in 2016 to alleviate stigmas about schizophrenia and to promote respect for mentally ill patients. The following is the pledge of the Brain Campaign written by SARDAA:

I will change my language from Behavioral or Mental disorder to Brain disorder.

I will educate others to change their language from Mental illnesses to Brain illness.

I will not use stigmatizing, discriminatory, or derogatory language related to Brain illnesses.

I will encourage my peers, family members, and colleagues to advocate when they see discriminatory language or practices in the media or in public.

I will advocate for people with Brain illnesses to have equal and comprehensive healthcare.

I will treat all people with Brain illnesses with respect and compassion.<sup>30</sup>

According to this pledge, SARDAA hopes to reshape the definition of schizophrenia by substituting the ‘brain’ for ‘mental.’ Through this adjustment in language, SARDAA appears to promote the neurological explanation of schizophrenia. SARDAA’s Brain Campaign has gained tremendous support from mental illness communities,

---

<sup>27</sup> Martha E. Shenton, and Chandlee C. Dickey, Melissa Frumin, and Robert W. McCrley, “A Review of MRI Findings in Schizophrenia,” *Schizophrenia Research*, vol. 49, 1-2 (April 2001): 21, [https://doi:10.1016/S0920-9964\(01\)00163-3](https://doi:10.1016/S0920-9964(01)00163-3).

<sup>28</sup> Shenton, Dickey, Frumin, and McCrley, “A Review,” 27.

<sup>29</sup> Xiaoqi Huang, Qiyong Gong, John A. Sweeney, Bharat B. Biswal, “Progress in Psychoradiology, the Clinical Application of Psychiatric Neuroimaging,” *The British Journal of Radiology*, vol. 92, 1101 (September 2019): 6, <https://doi:10.1259/bjr.20181000>.

<sup>30</sup> “Brain Campaign: Take the Pledge,” SARDAA, Mar 14, 2016, accessed Aug 15, 2020, <https://sardaa.org/brain-campaign-take-the-pledge/>.

including families with schizophrenia. Kathy Day, a mother of a schizophrenic patient, in her blog on May 08, 2018, urges that “It’s time to move schizophrenia and all brain-based illnesses into medicine.”<sup>31</sup> She underlines the urgency of modifying the categorization of schizophrenia and implies that there is sufficient evidence to endorse this modification.

The solid evidence, such as neuroimaging technologies’ findings that attest to schizophrenia’s relationship to brain abnormalities, acts as one of the foundational pillars of the supporters of the Brain Campaign. Linda Statlers, CEO and founder of SARDAA, discloses some of the motivations behind the Brain Campaign in her letter on May 21, 2018, to a government committee. She first refers to dementia, Parkinson’s disease, stroke, and brain tumors, which are all psychosis related illnesses classified as brain diseases, to argue that schizophrenia, also related to psychosis, should be considered a brain illness. She also alludes to neuroimaging findings multiple times as evidence for supporting the reclassification of brain illness. She states that MRI studies demonstrate a decrease in grey matter volume in people with schizophrenia (see Figure 3), and “a myriad of functional MRI and electrophysiologic studies have shown disturbances in brain and cognitive functioning.”<sup>32</sup> Though there is no evidence that the neuroimaging findings directly led to the Brain Campaign, these technologies definitely play a crucial role in supporting it and may have contributed to the initiation of the Brain Campaign.

Many schizophrenic patients and their relatives echo the reclassification of schizophrenia. Their views and experience indicate that they are particularly attracted to the biomedical model’s potential to allay prevailing prejudices against schizophrenia, increase everyday care for patients, and raise awareness of schizophrenia in other social conditions. Research from National Alliance for the Mentally Ill (NAMI) discloses that “When families belonging to the National Alliance for the Mentally Ill (NAMI) were asked what had helped them to cope with stigma, 73.2 percent indicated that ‘research findings which establish a biological basis for mental illness helped much or very much in dealing with stigma.’”<sup>33</sup> According to these families, a biological basis is very promising for destigmatizing mental illness. A *Washington Post* article on January 13, 2020, illustrates Aaron Dumsch and his families’ experience with schizophrenia. Aaron manifested full-blown schizophrenia in his senior year in college. His family fully accepted his condition one year later. Since then, his mother, also his main caregiver, Anita Dumsch, has learned about schizophrenia and ways to help people with this situation. During an interview, Anita describes her experience with stigmas, claiming that “It would be so much easier to say Aaron had

---

<sup>31</sup> Kathy Day, “Our Schizophrenia Blog: It’s time to reclassify schizophrenia,” May 08, 2018, <https://sardaa.org/its-time-to-reclassify-schizophrenia-as-a-brain-disorder/>.

<sup>32</sup> “Support Reclassification of Schizophrenia as a Neurological Disorder by SARDAA,” Sooner Than Tomorrow, Feb 28, 2019, <https://www.soonerthantomorrow.com/blog/2019/2/25/6zrk92dl8foipjgp7q1i4hyujwce51>.

<sup>33</sup> Joseph Dumit, “When explanations rest: ‘good-enough’ brain science and the new socio-medical disorders,” in *Living and Working with the New Medical Technologies: Intersections of Inquiry*, edited by M. Lock, A. Young, and A. Cambrosio (Cambridge: Cambridge University Press, 2000), 224.

brain cancer, because the empathy would be immediate. When I say, 'My son is mentally ill with schizophrenia,' it's as if I said leprosy."<sup>34</sup> From Anita's perspective, reclassification of schizophrenia contests the negative connotation of schizophrenia and can reduce stigma due to the social neutrality of physical conditions.

Specifically, families suggest that the biomedical model of schizophrenia justifies the existence of schizophrenia and redirects the guilt away from patients or families of schizophrenic patients. In her blog on September 18, 2018, Dandelion, who suffered from schizophrenia since she was 15, quoted a therapist that she used to see who once told her, "So starting from the moment when you didn't have [an] enemy anymore, you created others yourself to keep you in a state of persecution."<sup>35</sup> She imparts that the phrase "you created" constantly haunts her. On the one hand, according to Dandelion, this "you created" denotes that she, as a schizophrenic patient, is guilty of the mental illness she suffered from. On the other hand, the phrase also implies that schizophrenia is an imaginary illness rather than an actual disorder. Philosopher Angela K. Thachuk argues that "[o]nce localized in the body, the disorder somehow becomes more concrete and tangible, lending credence to the individual's experience."<sup>36</sup> Thachuk's explanation suggests that the biomedical model confronts the notion that mental illness is imaginary pain, and legitimizes patients' sufferings, thus countering the prejudice that patients are responsible for their mental conditions. Dumit, in particular, ascribes the biomedical model's ability to reduce stigma to neuroimaging technologies, claiming that "[b]rain imaging offers the promise of *showing* that the disorder really is in their brain and not in their heads."<sup>37</sup> Dumit identifies the transformation of the origin of mental illness from the head, a conceptual term, to the brain, a biological entity, presenting the way in which the biomedical model authorizes mental illness. He also suggests that neuroimaging adds credibility to this destigmatization process.

Legitimizing the existence of mental illness with the biomedical model acts as the foundation of the venture of fighting against stigmas. Relieving patients' self-blame comes as an additional benefit of this legitimization. Angela K. Thachuk, philosopher mentioned before, argues that the biological basis "undercut[s] stigmatizing assumptions that persons with mental illness are simply weak-willed."<sup>38</sup> According to Thachuk, the biomedical model challenges the notion that "you created" the illness mentioned in Dandelion's blog, thus unshackling patients from self-blame. In addition to empowering patients, the biomedical model is also perceived by schizophrenic families as a way to relieve their guilt. Their support for medical research of mental illness demonstrates their enthusiasm for the biomedical model. According to a researcher of the Brain Imaging Center at the University of California, Irvine, one schizophrenic family in Orange County had donated \$250,000 to the center to support research on brain images of schizophrenia. The researcher explains that until the late 1980s, a child's schizophrenic conditions were thought to be induced by maternal

---

<sup>34</sup> Abigail Jones, "What Schizophrenia Does to Families - and Why the Mental Health System Can't Keep Up," *Washington Post*, January 13, 2020, <https://www.washingtonpost.com/magazine/2020/01/13/what-schizophrenia-does-families-why-mental-health-system-cant-keep-up/?arc404=true>.

<sup>35</sup> Dandelion (@sirdandelion), "Schizophrenia and activism," Medium, Sep 18, 2018, <https://medium.com/@sirdandelion/schizophrenia-and-activism-5cb1b256c43b>.

<sup>36</sup> Thachuk, "Stigma and the Politics," 149.

<sup>37</sup> Dumit, "When explanations rest," 219.

<sup>38</sup> Thachuk, "Stigma and the Politics," 149.

wrongdoings. This prejudice was so deeply absorbed by the society that it still underlies many people's beliefs. The researcher points out that many families experienced self-blame when they saw a psychiatrist. Thus, they are thrilled by the biological interpretation of schizophrenia, which acts as the evidence that proves their innocence.<sup>39</sup>

Along with its potential to mitigate stigma, the biomedical model of schizophrenia also strengthens people's understanding of this illness, thus resulting in daily care improvement for schizophrenic patients. Patients' families believe that a biological basis of mental illness increases other people's understanding of patients' conditions. A survey conducted by bioethics scholars with patients suffering from depression illustrates that "sixty-three percent of patients reported that brain scans would aid their families to accept their condition better."<sup>40</sup> As Thachuk points out, with the notion that "seeing is believing," brain scans appear to be a powerful way to explain mental illness to patients and their families. Reliance on biological explanation also applies to schizophrenia patients and families. In the supporter profile containing stories of brain health advocates of One Mind, a non-profit organization founded by a schizophrenic patient's parents, Anna, the daughter of Gayle Ayers (from Reno, Nevada), was diagnosed with schizophrenia in high school in 2009. When involved in the ABBRC (Adolescent Brain Behavior Research Clinic) program, a program for youths who experience full-blown psychotic disorders, Anna received clinical assessments by ABBRC staff. By identifying biological changes in the brain, the assessments "undoubtedly helped the Ayres family understand how Anna's cognitive abilities and non-verbal skills were being impacted," as claimed by One Mind's supporter file.<sup>41</sup> After comprehending Anna's illness, Gayle was able to modify her homeschooling method, thus providing more appropriate care for Anna. One Mind utilizes Anna's story to argue that the biomedical model of schizophrenia can result in practical improvements in daily care for schizophrenic patients. The mother of a schizophrenic patient, Kathy Day, also seems to agree that the reclassification of schizophrenia offers better care for patients. In her blog (mentioned before), she argues that "nothing helps [schizophrenia] except meds... Med[icines] can help the underlying brain disorder that causes what appears to be behaviors."<sup>42</sup> In her interpretation, the severe behavioral symptoms of schizophrenia are rooted in the abnormalities in the brain. As Day implies, identifying schizophrenia as a behavioral disorder is insufficient as it tackles merely the surface of the problem rather than the root, whereas applying a biomedical model of schizophrenia enables patients to access the proper treatment: medicines.

In addition to the possibilities of reducing stigmas and improving care for schizophrenic patients, advocates of the reclassification of schizophrenia also discern benefits of the biomedical model in other areas, such as mental illness research and the healthcare and insurance system. In a case in 1987 in Arkansas, an insured father won the case by demonstrating that his daughter's bipolar disease was a physical dis-

---

<sup>39</sup> Dumit, "Traveling Images," 155.

<sup>40</sup> J. Illes, S. Lombera, J. Rosenberg and B. Arnow, "In the mind's eye: provider and patient attitudes on functional brain imaging," *Journal of psychiatric research* vol. 43, 2 (2008): 107-14. doi:10.1016/j.jpsychires.2008.02.008.

<sup>41</sup> "Brain Health Caretaker and Mom: One Mind Supporter Profile," One Mind, accessed Aug 17, 2020, <https://onemind.org/reduce-mental-illness-stigma/brain-health-caretaker-and-mom/>.

<sup>42</sup> Day, "Our Schizophrenia Blog."

ease. Because the father provided evidence and testimony that showed bipolar disorder is a physical disease caused by the brain, the court eventually ruled that his daughter's illness should be covered by the insurance company Blue Cross Blue Shield's (BCBS) plan, which provided hospitalization and medical treatments. In this case, the biomedical model of mental illness enabled patients with bipolar disorder to receive adequate health care covered by insurance.<sup>43</sup> The victory in this case illustrates the power of the biomedical model in scenarios other than in the mental illness community and implies the superiority of physical illness, thus suggesting a reason for promoting this model.

Similar to bipolar disorder described in the 1987 case, schizophrenia is also a mental illness that has received inadequate attention, treatments, and care. Linda Stalters considers the benefits of the biomedical model of schizophrenia in a wide range of applications, claiming in a 2019 article that defining schizophrenia as a neurological illness helps to "provide commensurate research and treatment with increased funding," and "provide beds instead of incarceration or homelessness by circumventing the Institutions for Mental Diseases (IMD) exclusion."<sup>44</sup> The IMD exclusion, established in 1905, prohibits reimbursement for hospitalized mentally ill patients.<sup>45</sup> Stalters argues that mentally ill patients can only get financial support from the government by identifying themselves as physically ill. Furthermore, Stalters also implies that schizophrenic patients are wrongly judged as criminals, presumably due to their unusual behaviors, as she claimed that reclassification of schizophrenia can also "eliminate criminalizing people with a brain illness."<sup>46</sup> Expanding the benefits of the reclassification of schizophrenia to research, the healthcare system, and homelessness issues, Linda Stalters explicates the significance of the reclassification of schizophrenia. The foundational basis of Stalters's argument is that physical illness is more convincing and receives greater attention. Hence, she believes that by reclassifying schizophrenia as a neurological illness, schizophrenia will also face less suspicion and receive proper care.

Nevertheless, some mentally ill patients raise concerns about neuroimaging technologies and the biomedical model it promotes. James Gorman, the volunteer of a brain research project introduced before, describes his anxiety at seeing the image of his brain after he underwent an MRI scan. He claims that his brain contains his "innermost secrets," and the result of his brain image is there for "anyone to see."<sup>47</sup> Gorman clearly perceives neuroimaging's power to intrude on privacy, and admits to feeling "relieved" as the secrets could not be understood yet. Journalist Tracy Thompson records her experiences with depression in her book, *The Beast: A Reckoning with Depression* (1995), describing her struggles with the biomedical model of mental disorders. She states that by understanding depression as a brain disease, she would "be simply a product of some chemical abnormality in a lumpy gray organ between

---

<sup>43</sup> Thachuk, "Stigma and the Politics," 150.

<sup>44</sup> Linda Stalters, "Why We Must Reclassify Schizophrenia Spectrum as a Neurological Brain Illness," *Future of Personal Health*, Nov 19, 2019, <https://www.futureofpersonalhealth.com/mental-health/why-we-must-reclassify-schizophrenia-spectrum-as-a-neurological-brain-illness/>.

<sup>45</sup> "The Medicaid IMD Exclusion and Mental Illness Discrimination," *Treatment Advocacy Center*, August 2016, <https://www.treatmentadvocacycenter.org/evidence-and-research/learn-more-about/3952>.

<sup>46</sup> Stalters, "Why We Must Reclassify."

<sup>47</sup> Gorman, "A Search for Self in a Brain Scan," D3.

[her] ears.”<sup>48</sup> Thompson indicates a sense of loss of autonomy by referring to herself as a “product,” as if she is merchandise rather than a person, lacking spiritual power.

Though some people view the downside of neuroimaging technologies, schizophrenic patients and their families generally think positively of neuroimaging technologies. They perceive the ability of neuroimaging technologies to enhancing understanding and empathy and their potential to bring about improvements in healthcare and insurance system for mental illness. The fundamental pillar underlay these perceived benefits is neuroimaging’s ability to provide a biological basis that supports classifying schizophrenia as a physical disease rather than a mental illness.

## Conclusion

By creating images of a subject’s brain structures and activity, MRI and fMRI allow scientists to study the structure and function of the brain and the way they relate to symptoms of mental disorders. When MRI and fMRI images enter the public arena, the interaction of these technologies and images with the general public endow neuroimaging with social implications. While there are concerns about neuroimaging’s ability to create oversimplified categories, intrude on privacy, and redefine autonomy of self, schizophrenic patients and their families give almost full credence to these technologies. They regard neuroimaging technologies as a credible source of evidence, which can mitigate stigmas about schizophrenia, improve daily care for patients, and in general, generate more attention for this illness to enhance care for patients in a variety of social systems. Compared to physicians, who explore the biological nature of schizophrenia and the medical application of MRI and fMRI, patients and relatives tend to concentrate on social implications of these technologies and their potential in improving humanistic care and social service for mentally ill patients.

In this paper, while presenting schizophrenic patients’ and their families’ optimism of neuroimaging technologies in the recent two decades in America, I have also shown neuroimaging’s power to disseminate the biomedical model of mental illness and impact the public’s understanding of schizophrenia. Nonetheless, this paper does not fully explore how patients’ and their families’ confidence in neuroimaging technologies reflect current social conditions. Their optimism may mirror the society’s obsession with progress and technologies. According to historian Merritt Roe Smith, more and more people attach their hopes to technologies, treating technology and science as the “great panacea for everyday problems.”<sup>49</sup> Schizophrenic patients and their families also seem to perceive neuroimaging technologies as a solution for destigmatizing schizophrenia and providing better care for patients. In addition to people’s optimistic views about technologies, I propose that patients’ and their families’ enthusiasm about these technologies reflects the disparity between how society treats physical health and mental health. Dumit argues that for patients, the option of waiting for more evidence to show their mental illness “is simply not livable given the current state of healthcare in the US.”<sup>50</sup> Without support from the healthcare sys-

---

<sup>48</sup> Dumit, “Traveling Images,” 161.

<sup>49</sup> Merritt Roe Smith, “Technological Determinism in American Culture,” in *Does Technology Drive History? The Dilemma of Technological Determinism*, ed. Merritt Roe Smith, and Leo Marx (Cambridge: MIT Press, 1994), 23.

<sup>50</sup> Dumit, “When explanations rest,” 221.

tem, mentally ill patients and their families seek any scrap of evidence and information, such as MRI and fMRI studies, that can empower them. Perhaps the greatest reason for patients' optimism about neuroimaging technologies is rooted in the social ignorance about mental health. Faith in technology, thus, becomes an alternative approach to achieving justice as a result of insufficient support from society.

## Bibliography

- American Psychological Association. "Scanning the Brain." Accessed August 20, 2020. <https://www.apa.org/action/resources/research-in-action/scan>.
- Borgelt, Emily L., Daniel Z. Buchman, and Judy Illes. "Neuroimaging in mental health care: voices in translation." *Frontiers in Human Neuroscience*, vol. 6, no. 293 (October 2012). <https://doi.org/10.3389/fnhum.2012.00293>.
- Cartwright, Lisa. "Decomposing the body: X Rays and the Cinema." In *Screening the Body: Tracing Medicine's Visual Culture*, 107-142. Minneapolis: University of Minnesota Press, 1995.
- Dandelion (@sirdandelion). "Schizophrenia and activism." *Medium*, Sep 18, 2018. <https://medium.com/@sirdandelion/schizophrenia-and-activism-5cb1b256c43b>.
- Day, Kathy. "Our Schizophrenia Blog: It's time to reclassify schizophrenia." May 08, 2018. <https://sardaa.org/its-time-to-reclassify-schizophrenia-as-a-brain-disorder/>.
- Dumit, Joseph. "Traveling Images, Popularizing Brains." In *Picturing Personhood: Brain Scans and Biomedical Identities*, 141-205. Princeton: Princeton University Press, 2004.
- Dumit, Joseph. "When explanations rest: 'good-enough' brain science and the new socio-medical disorders." In *Living and Working with the New Medical Technologies: Intersections of Inquiry*, edited by M. Lock, A. Young, and A. Cambrosio, 209-232. Cambridge: Cambridge University Press, 2000.
- Falkai, Peter, Andrea Schmitt, and Nancy Andreasen. "Forty Years of Structural Brain Imaging in Mental Disorders: Is it Clinically Useful or Not?" *Dialogues in Clinical Neuroscience*, vol. 20, 3 (September 2018): 179-186. <https://doi.org/10.31887/DCNS.2018.20.3/pfalkai>.
- Gabrielsen, Paul. "Quantity, Not Just Quality, in New Brain Scan Method." *Neuroscience News*, Nov. 5, 2013. Accessed Sep. 3, 2020. <https://neurosciencenews.com/neuroimaging-mri-quantify-brain-volume-571/>.
- Glover, Gary H. "Overview of functional magnetic resonance imaging." *Neurosurgery clinics of North America*, 22(2), 133-vii (April 2011): 133-139. <https://doi.org/10.1016/j.nec.2010.11.001>.
- Gorman, James. "A Search for Self in a Brain Scan: A picture of your brain may hold your..." *New York Times (1923-Current file)*, Jan 7, 2014, D3.
- Huang, Xiaoli, Qiyong Gong, John A. Sweeney, Bharat B. Biswal. "Progress in Psychoradiology, the Clinical Application of Psychiatric Neuroimaging." *The British Journal of Radiology*, vol. 92, 1101 (September 2019): 1-9. <https://doi.org/10.1259/bjr.20181000>.
- Illes, Judy, S. Lomber, J. Rosenberg and B. Arnow. "In the mind's eye: provider and patient attitudes on functional brain imaging." *Journal of psychiatric research*, vol. 43, 2 (2008): 107-14. <https://doi.org/10.1016/j.jpsychires.2008.02.008>.



- Jones, Abigail. "What Schizophrenia Does to Families - and Why the Mental Health System Can't Keep Up." *Washington Post*, January 13, 2020. <https://www.washingtonpost.com/magazine/2020/01/13/what-schizophrenia-does-families-why-mental-health-system-cant-keep-up/?arc404=true>.
- Joyce, Kelly. "Appealing Images: Magnetic Resonance Imaging and the Production of Authoritative Knowledge." *Social Studies of Science* 35/3 (June 2005): 437-462. <https://doi.org/10.1177/0306312705050180>.
- Kevles, Bettyann H. *Naked to the Bone: Medical Imaging in the Twentieth Century*. New Jersey: Rutgers University Press, 1997.
- Mark, Leo, and Merritt Roe Smith. "Introduction." In *Does Technology Drive History? The Dilemma of Technological Determinism*, edited by Merritt Roe Smith, and Leo Marx, ix-xiv. Cambridge: MIT Press, 1994.
- McGreevey, Sue. "Finding signs of life when it matters most." *The Harvard Gazette*, Jan. 20, 2017. Accessed Sep. 3, 2020. <https://neurosciencenews.com/neuroimaging-mri-quantify-brain-volume-571/>.
- One Mind. "Brain Health Caretaker and Mom: One Mind Supporter Profile." Accessed Aug 17, 2020. <https://onemind.org/reduce-mental-illness-stigma/brain-health-caretaker-and-mom/>.
- Parens, Erik, and Josephine Johnston. "Neuroimaging: Beginning to Appreciate Its Complexities." *The Hastings Center Report*, vol. 44, no. 2 (March-April 2014): S2-S7. <https://doi.org/10.1002/hast.293>.
- SARDAA. "Brain Campaign: Take the Pledge." Mar 14, 2016. Accessed Aug 15, 2020. <https://sardaa.org/brain-campaign-take-the-pledge/>.
- Shenton, Martha E., Chandlee C. Dickey, Melissa Frumin, and Robert W. McCrley. "A Review of MRI Findings in Schizophrenia." *Schizophrenia Research*, vol. 49, 1-2 (April 2001): 1-52. [https://doi.org/10.1016/S0920-9964\(01\)00163-3](https://doi.org/10.1016/S0920-9964(01)00163-3).
- Shenton, Martha E., Thomas J. Whitford, and Marek Kubicki. "Structural Neuroimaging in Schizophrenia from Methods to Insights to Treatments." *Dialogues in Clinical Neuroscience*, vol. 12-3 (September 2010): 317-32. <https://doi.org/10.31887/DCNS.2010.12.3/mshenton>.
- Smith, Merritt Roe. "Technological Determinism in American Culture." In *Does Technology Drive History? The Dilemma of Technological Determinism*, edited by Merritt Roe Smith, and Leo Marx, 1-35. Cambridge: MIT Press, 1994.
- Sooner Than Tomorrow. "Support Reclassification of Schizophrenia as a Neurological Disorder by SARDAA." Feb 28, 2019. <https://www.soonerthantomorrow.com/blog/2019/2/25/6zrk92dl8foipjgp7q1i4hyujwce51>.
- Stalters, Linda. "Why We Must Reclassify Schizophrenia Spectrum as a Neurological Brain Illness." *Future of Personal Health*, Nov 19, 2019. <https://www.futureofpersonalhealth.com/mental-health/why-we-must-reclassify-schizophrenia-spectrum-as-a-neurological-brain-illness/>.
- Thachuk, Angela K. "Stigma and the Politics of the Biomedical Models of Mental Illness." *The International Journal of Feminist Approaches to Bioethic*, vol. 4, no. 1 (Spring 2011): 141-163. <https://www.jstor.org/stable/10.2979/intjfe-mappbio.4.1.140>.
- Treatment Advocacy Center. "The Medicaid IMD Exclusion and Mental Illness Discrimination." August 2016. <https://www.treatmentadvocacycenter.org/evidence-and-research/learn-more-about/3952>.
- WHO. "Schizophrenia." Accessed Aug 24, 2020. <https://www.who.int/news-room/fact-sheets/detail/schizophrenia>.

Woolley, James, and Philip McGuire. "Neuroimaging in schizophrenia: what does it tell the clinician?" *Advances in Psychiatric Treatment*, vol. 11, issue 3 (May 2005). <https://doi.org/10.1192/apt.11.3.195>.



# Analysis of Testing Accuracy Threshold for COVID-19 through a Computational SIR Model

---

*Nabo Yu*

**Author Background:** *Nabo Yu grew up in China and currently attends The Webb Schools in Claremont, California in the United States. His Pioneer research concentration was in the field of computer science and titled “Networks: from Searching the World Wide Web to Slowing the Spread of Infectious Diseases.”*

## Abstract

There are certain thresholds in testing accuracy for pandemics above which the total number of people infected decreases dramatically. These effective threshold values are dependent on levels of social contact and other prevention measures. To investigate this claim, this study focuses on

1. Creating a computational spreading model representative of the spread of COVID-19;
2. Observing the effect of testing accuracy for the virus on the spread of the epidemic in our model and look for threshold behavior; and
3. Evaluating potential explanations for this behavior.

The results of this investigation suggest that:

1. Our hypothesis about thresholds caused by testing accuracy is correct;
2. Our model is capable of showing these thresholds when given certain parameters; and
3. Further research should be conducted on how different levels of social interaction would impact the effectiveness of testing in minimizing the effects of especially contagious pandemics, especially on whether there is a “window” for which increasing testing accuracy becomes particularly effective.

**Keywords:** COVID-19, testing accuracy, networks, threshold, basic reproduction number, SIR model

## 1. Introduction

### 1.1 Background

Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is a disease that has recently spread to communities around the world in the first quarter of 2020 (Mayo Clinic, 2020). It is a novel disease to which the human population does not have any effective immunity, and its significantly longer incubation period compared to many other common diseases

have made it difficult for individuals to recognize symptoms and isolate before infecting others (World Health Organization, 2020; Mayo Clinic, 2020; Guan, Ni, Hu, Liang, Ou, He, et al., 2020). Moreover, the disease has symptoms similar to that of the common cold and the common flu in mild and moderate cases, which make cases of COVID-19 difficult to identify before individuals carry them to higher-risk groups (Epidemiology Working Group for NCIP Epidemic Response, Chinese Center for Disease Control and Prevention, 2020). Among the most common methods of reducing the spreading and human costs of the coronavirus are social isolation/distancing, added personal hygiene and protective equipment, zoning, and testing for the virus (Mayo Clinic; CDC; Cascella, Rajnik, Cuomo, Dulebohn & Di Napoli, 2020). Among them, testing for the coronavirus has become a very controversial and popular topic given that many manufactured testing kits have low accuracy (BBC, 2020 Feb 13). Some countries have opted to not use tests that are less effective and suggest isolation for all individuals who display symptoms similar to that of COVID-19 instead (BBC, 2020 Apr 28; BBC, 2020 Mar 30). It is undoubted that higher social isolation will more effectively contain the virus, though the scale, length, and severity of such efforts have been unpredictable and are heavily debated given the financial restraints social isolation poses. Also, due to the structure of countries' welfare/healthcare systems, the financial impact on many individuals will ultimately affect other individuals in the society. (BBC, 2020 Apr 9). In April, a batch of kits bought from China were "found to be insufficiently accurate by a laboratory at Oxford University" and therefore mostly unused, resulting in a considerable waste of public funds (Kirkpatrick & Bradley, 2020). Due to inconsistencies in the results of coronavirus testing, the public in Europe and the United States have expressed great concern for the effectiveness of testing in controlling the virus (Gallagher). It is therefore pertinent to establish a comprehensive understanding of the impact testing accuracies has on the spreading of diseases.

In the study of networks, it is not uncommon to observe "threshold" effects: in the change of distances and clustering during the formation of small-world networks, in the emergence of scaling in random networks, or in the reproduction number of spreading processes (Kleinberg & Easley, 2010). "Threshold effect," as referred to in this paper, is the phenomenon where a dependent variable changes drastically around a certain value of the independent variable, as if from a stage to another. Compared to a spreading process without any intervention, testing individual nodes with symptoms and isolating if the test is positive should decrease the overall rate of infection and directly help the epidemic recede much faster. This will reduce the percentage of the population that will be infected when herd immunity becomes significant or will reduce the peak strain on the healthcare system (Kleinberg & Easley, 2010). The basic reproduction number, denoted by  $R_0$ , experiences a major threshold in its effect on the spreading process, as one that is higher than 1 would imply a growing rate of increase in the "infected" pool, while a number lower than 1 would imply a declining rate of increase, as the disease "dies off" (Iannelli, Kim, & Park, 1999; Kleinberg & Easley, 2010).

## 1.2 Hypothesis

With the ideas established in the last paragraph, we suggest that there are thresholds on a network spreading simulation around certain testing accuracies, where the number of total people infected during the pandemic will decrease dramatically. The relationship between testing accuracy and the total number of individuals infected may be similar to a logistic curve.

### 1.3 Rationale and Approach

To investigate the claim above, it would be suitable to use simulations compared to mathematical models, for the following reasons. First, simulations are more accurate depictions of real-life scenarios and effects compared to mathematical models. Their relative unpredictability renders them much more interesting to observe and more realistic to draw observations from. Second, they are more capable of expressing the full complexity of the effects different factors in the epidemiological model have on the spreading process. Using simulations, we observe the effects testing accuracy has on the spread of the epidemic in our model.

This study serves to fill gaps in current understanding of testing accuracy's true effect. These gaps were identified by news articles on testing practices and testing kit accuracy in the United Kingdom. It is for this reason that the model used in this study aims to approximate the testing practices of the United Kingdom at the time of writing.

### 1.4 Literature Review

On the topic of the coronavirus, it could be observed that there has been a drastic growth in research done surrounding its effects and containment. Also, over the past century, there has been an abundance of epidemiological studies done through the lens of network science, both on computational models and the structure of contact networks.

#### 1.4.1 Testing for COVID-19

At the time of writing, COVID-19 testing is still a highly discussed topic, and many technologies and public policies are subject to drastic change. As introduced previously, COVID-19 has similar symptoms with the common cold and the seasonal flu (CDC). Testing for COVID-19 has been widely studied and compared to other methods of diagnosis for the disease. In relatively earlier stages of the pandemic, Corman, Landt, Kaiser, Molenkamp, Meijer, Chu, et al. (2020) developed a diagnostic workflow for COVID-19 using reverse-transcription polymerase chain reaction (RT-PCR), and Ai, Yang, Hou, Zhan, Chen, Lv, et al. (2020) explored the correlation between results from the RT-PCR testing method and background CT scans to diagnose COVID-19 onset, with results that challenge the effectiveness of using the RT-PCR method alone. The controversy surrounding how much effort should be put into testing for the coronavirus, is a major consideration for this investigation into the effectiveness of testing for the virus.

Most reports on the epidemic and recommendations for the public suggest individuals should be tested when they show symptoms (CDC; WHO; Mayo Clinic; Guan, Ni, Hu, et al.). Though there is much debate specifically on the difference between different methods of testing and diagnosis for the coronavirus, the main focus of this study is on the effects of testing accuracy in general, upon becoming symptomatic.

### 1.4.2 Spreading Process and Epidemiological Analysis

There has been abundant research in epidemiological models as series of differential equations. For example, Ruan & Wang (2003) and Cai, Li, & Ghosh (2009) studied the behavior of infection in stage-based models of epidemics through differential equations, while Zaman, Kang, & Jung (2008) and Katriel (2010) focused on the effects of vaccination and immunity to reinfection in these models. McCluskey (2010) showed the endemic equilibrium to be stable for  $R_0 > 1$  in a SEIR model with infinite delay. Although analyses of epidemiological models through differential equations are highly informative and have been the subject of extensive study, they are not the main focus of this study, for reasons discussed in the Rationale and Approach section.

In comparison, in their analysis of epidemics, stochastic differential equation SIS epidemic models are closer to the computational model of this study than deterministic ones. In these models, there has been extensive effort in the study of the extinction and persistence of the disease, i.e. studying the effect of different parameters on the basic reproduction number threshold. Anderson, & May (1979) explored the endemic equilibrium of the disease in an SIR model in a two-part paper and re-examined the  $R_0$  threshold. Gray, Greenhalgh, Hu, Mao, & Pan (2011) simplified stochastic differential equations to determine the basic reproduction number when contributed to by multiple parameters, simplifying certain complexity. Gray, Greenhalgh, Mao, & Pan (2012) again studied determination of persistence and extinction, in an SIS epidemic model with Markovian switching. Tornatore, Buccellato, & Vetro (2005) studied the extinction of the disease under different conditions in a stochastic SIR model and its stability under different levels of noise; Ji & Jiang (2014) furthered their study on simplifying the complexity of various parameters and their study of noise in these models. Generally, it can be concluded that when introduced to certain levels of complexity and noise, mathematical determination of the basic reproduction number in the SIR model rapidly becomes highly difficult. In particular, most researchers in this topic are interested in learning about the persistence-extinction threshold of the disease when introduced to complexity, which is also the main focus of this study. However, this study will use computer simulations to show these thresholds and attempt to explain the phenomenon with methods that are much simpler mathematically.

### 1.4.3 Network Structure

The study of spreading processes on different network/graph topologies is still a developing field, though there are a few considerations in creating the underlying networks in the simulations. First, according to Newman, Strogatz, & Watts (2001), the random generation of networks representative of social networks has been studied extensively through networks of friendships and communication. However, it is difficult to experimentally determine the contact network by which diseases spread. Though random graphs with Poisson-distribution vertex degrees may be highly unrepresentative of stable, long-established social networks, many past models in network epidemiology have used the so-called fully mixed approximation, which is the assumption that contacts are random and uncorrelated. It is for this reason that this model has opted to use random graphs as the basis for the spreading process simulation.

## 2. Methodology

### 2.1 Spreading Process

Since the focus of this study is on reducing the effects of COVID-19, our computational simulation reflects and caters to characteristics of the novel coronavirus. Firstly, since it is not yet known whether and how reinfections for people who recover from the disease occur, a Susceptible-Infectious-Removed model without consideration for reinstating susceptibility is the basis of our model (CDC). Since it is not the focus of this investigation to consider the long-term effects of the epidemic, the model does not take into account regular births and deaths in the population. The above considerations encompass the simplest model that was first created for this study. In general, the simulation would initialize with the parameters:

- one undirected, weighted graph,
- an initial set of infected nodes,
- and an average number of days till each node is removed.

Then, the program would simulate a spreading process that follows the following: all initial infected nodes are infected on day one, and all newly infected nodes become infectious the next day. Each day, each infected node will try to infect its neighbors, and the success of the infection is dependent upon the “closeness” assigned to the edge between the two nodes; each infected node will also have a probability of being removed every day, calculated by  $\frac{1}{d_r}$ , where  $d_r$  is the average number of days to removal. This process ensures that the average number of days spent by each individual in the “infectious” pool is  $d_r$ , while following the variability of number of days to recovery for different symptoms from the disease, which has been shown in observations of coronavirus cases (WHO).

This general method of simulation described above has been used in a previous study (Friedman, Friedman, Johnson, & Landsberg, 2020).

The simple model described above (without intervention) was tested with random graphs, and the spreading process conforms with what one would expect from most conventional epidemic spreading models. The snapshots of nodes in each state throughout each period of the entire epidemic were checked manually for dozens of simulations and proven to be correct in simulating the process as described above.

### 2.2 Simulation of Testing

Our model reflects the real-world situation that it is trying to mimic: the incidence of testing depends on the appearance of symptoms. This mechanism is simulated the same way as removal: each infected node will have a probability of becoming symptomatic each day, calculated by  $\frac{1}{d_s}$ , where  $d_s$  is the average number of days to become symptomatic. The parameters for these two are set through empirical data compiled by the CDC and the WHO. To simulate the process of testing, our model “tests” nodes the day they become symptomatic. In this simulation, we only consider the ability of the tests to identify novel-coronavirus carriers, and disregard keeping track of symptomatic non-coronavirus patients. Therefore, in this case, our testing

accuracy dictates the percentage of symptomatic coronavirus carriers who are isolated, by using the testing accuracy to give each newly symptomatic patient that probability of being removed from the “infectious” pool. Tested individuals are not tested again and will only be removed from the “infectious” pool through the recovery probability.

It may be questioned whether our model should consider a delay between showing symptoms and testing, as testing on day one of symptoms may be unrealistic for some populations and there may also be a delay in returning results. However, our model simulates testing policy in the United Kingdom, which instructs citizens to self-isolate from the day of showing any suspicious symptoms to the day of receiving a negative test result (National Health Service, 2021). Therefore, this model will not consider the delays in testing, and the model’s results will already be conservative compared to real-world results.

### 2.3 Network/Graph Generation

For reasons discussed above in Literature Review, this model uses a  $G_{n,p}$  random graph, also known as an Erdős-Rényi graph or a binomial graph, where the average number of edges a node has is approximated by  $n \times p$ . However, in the generation of edge weights (previously referred to as “closeness” between two nodes), we employed a slightly different randomized approach, which aims to compensate for the deficiencies of the random graph structure and to mimic the distribution of weak/close social ties each node has. The edge weights are randomly generated conforming to an exponential distribution, where weak ties are much more common than stronger ties. These edge strengths are scaled according to a “top” parameter for the greatest edge strength. When a new graph is generated with  $m$  number of edges,  $m$  number of “closeness” values are generated according to the exponential distribution, and the maximum of these numbers are found. Then, all values in this set are scaled with a factor of  $\frac{t}{max}$ , where  $t$  is the “top” parameter, and  $max$  is the maximum of the set. These values are then randomly assigned to each edge. A generated set is graphed in Fig. 1 as a histogram.

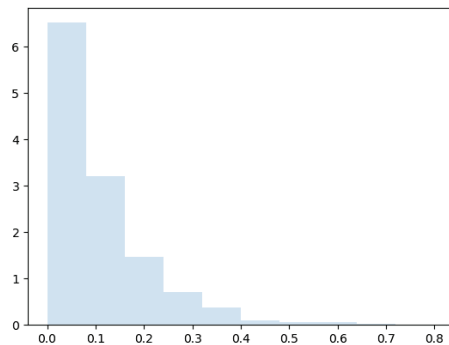


Figure 1. Histogram of edge weight distribution

The  $G_{n,p}$  random graphs were generated using python module NetworkX v2.4, and the “closeness” value arrays were generated using python module NumPy v1.18.

### 2.4 Running Simulations

Individual simulations were run and graphed over time with snapshots of individuals in each of the SIR states to ensure the reliability and correct functionality of the models. The simulations were then manipulated with different parameters, mostly through the full range of testing accuracies on different networks, which have different p-values for generating the graph. For each datapoint on a testing accuracy, the simulations are run up to 10 times on up to 10 multiple randomly generated



graphs (with same parameters) with randomized initial nodes. The “peak” of each simulation of the spreading—where the number of new cases were highest—was also recorded. More interestingly, the  $R_0$  values for each of these simulations were calculated, averaged for each of these datapoints generated with different testing accuracies, and graphed against testing accuracy. The exact method for determining  $R_0$  within the model is discussed below.

## 2.5 Calculation of $R_0$ from Simulations

As defined, the basic reproduction number ( $R_0$ ) is the total number of individuals an infectious person could infect over its infectious period, assuming that the population has no immunity, and all individuals except the infectious person themselves are susceptible (Kleinberg & Easley, 2010). When implemented in our model, the basic reproduction number is extremely difficult to determine mathematically but is relatively easy to obtain from the simulations themselves. During the simulation, we determine the success of an infection through a specific edge by the probability placed on the edge (“closeness”). In each period, each infectious node would first consider all of its neighbors for infection through this probability, regardless of whether the target neighboring node is susceptible. Later, the target neighboring node would only be moved to the infectious pool if it is susceptible and the “infection” was a success according to the probability assigned to the edge. Consider a node  $i$ , with  $n$  neighbors. If, for instance, we are in a later stage of the spreading process, and the effectual reproduction number has lowered, this effect is because node  $i$ ’s neighbors might not have the ability of being infected: they may have been infected before and have recovered, are isolated due to testing, or are infectious themselves. In this case, if we are to pretend as if node  $i$ ’s neighbors are all susceptible, we would only need to record the nodes that were successfully “infected” according to the probability on that edge. If this process is done throughout the infectious period of each node that were once infectious (with no repeating nodes), it would be possible to count how many neighbors each infectious node infected over its infectious period as if the population has no immunity, and all individuals except the infectious person themselves are susceptible. Averaging the total number of neighbors infected for all of the infectious nodes would yield the  $R_0$  throughout the simulation. This process is illustrated below with pseudocode:

```

If edge_weight > random value in [0, 1):
    Add neighbor to theoretical_infected_set
    If neighbor is susceptible:
        Add neighbor to infectious_pool
        Add neighbor to actual_infected_pool
...
If self is removed:
    Add count(theoretical_infected_set) to
    theoretical_infected_number_list
...
 $R_0 \leq \text{average}(\text{theoretical\_infected\_number\_list})$ 

```

Since the simulation is set to stop when there are either no infectious nodes left, or if all nodes are infectious, it is possible for the simulation to stop without

allowing the full infectious period of each node to run their course. To prevent miscalculation due to this mechanism, each count of the `theoretical_infected_set` is only logged if the node was removed in the end.

## 2.6 Presentation of Data

The graphs of individual simulations followed the number of total nodes in each pool and the number of newly symptomatic and isolated nodes during each day (each period). Isolated nodes are counted as members of the “removed” group. Taking the average of the total number infected and the full duration of the epidemic (from initial infection to the full elimination of the infectious pool/all other pools) proved to be a rather confusing way to represent the data, as it was later recognized with box plots that there were simulations in which almost no nodes were infected. These outliers seem to become more frequent as testing accuracies increase, which suggests they may be cases where the initial node was immediately isolated. Since the focus of this investigation is on the spreading processes, those outliers were mostly removed in the boxplot graphs presented. The basic reproduction number were also graphed against testing accuracy. The graphs are created with python module Matplotlib v3.2.2.

It should be noted here that, throughout the entire experimental process of this study, a vast array of graphs and data were generated. However, only data significant to the discussion are presented.

## 2.7 Parameters Used

Since this study is meant to simulate the spread of the coronavirus, and data on the pandemic are still being collected, it should be noted what values were used for important parameters.

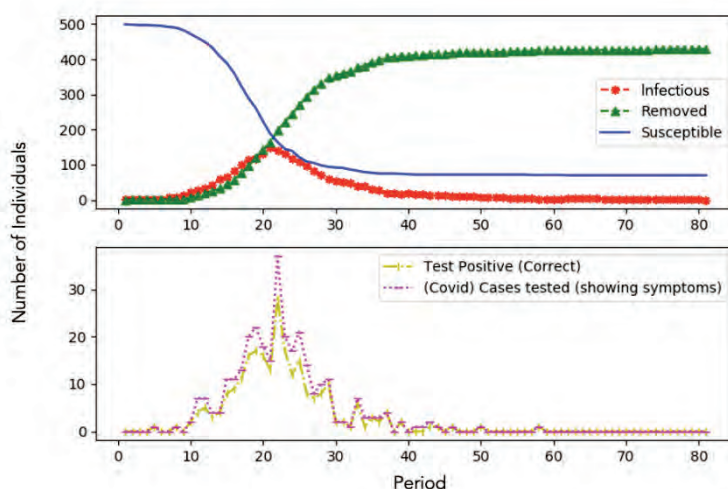
According to data from the CDC and the WHO, the incubation period for COVID-19 is around 4-6 days. The number used in most simulations in this study is 5 days.

According to WHO’s observations and CDC’s recommendations for discontinuation of isolation, mild cases of COVID-19 recover around 14-20 days, which covers most cases. The number used in most simulations in this study is 17 days. The reliability of using this number to produce real results is increased when considering the method of determining removal for each node described above in Spreading Process.

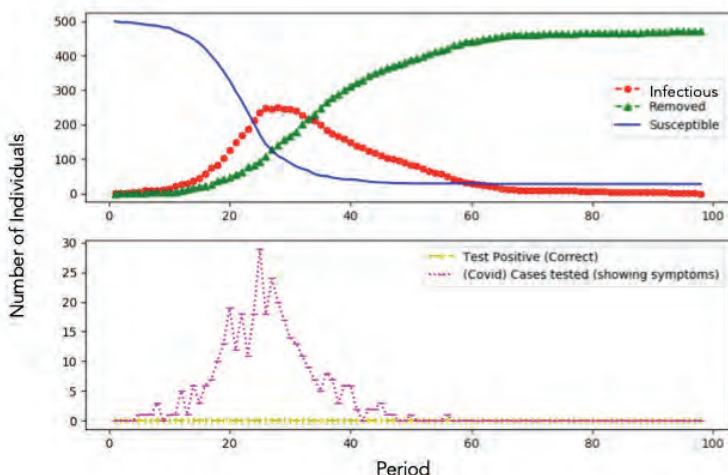
# 3. Results

## 3.1 Single Simulations

In Fig. 2, the single-simulation demonstration with testing accuracy 0.8, we can see the expected curves for each of the SIR pools, and the number of cases appearing in each time period. We can see that the number of new cases has its waxes and wanes, and the number of susceptible-infectious-removed individuals changes as expected. Note that in real life, the yellow line, “Test Positive (Correct),” is theoretically the number of new cases that would be reported in measured data.



**Figure 2.** Tracking the spreading process in the simulation with testing accuracy 0.8.



**Figure 3.** Tracking the spreading process in the simulation with no testing or isolation.

In Fig. 3, the testing accuracy is 0.0, meaning that there is no testing in this model. This is the simulation of a situation without any human intervention, and we can notice a relatively significant increase in the number of people who are eventually infected in the process. This will be discussed further in the following sections.

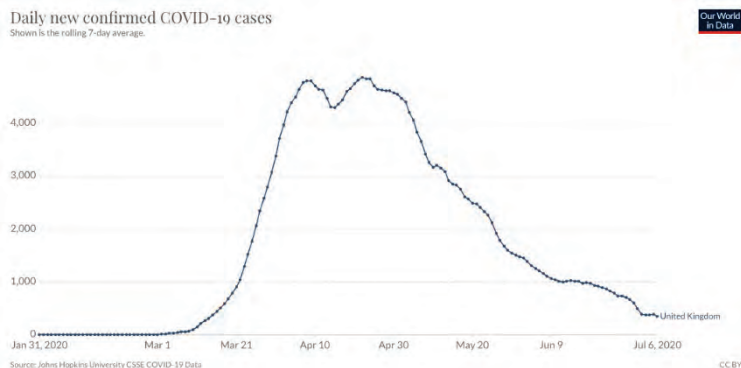
Fig. 2 and 3 are important in understanding the data presented in the following sections. The number of people who were eventually infected are equivalent to the number of people who are not susceptible at the end of the simulation. The “peak” is marked by the day in which there was the greatest number of individuals in the infectious pool. The data presented below focus on analyzing the outcomes of these single simulations with varying testing accuracy.

In generating Figures 5-1 to 7-2, we ran the simulation depicted in Figures 2 & 3 with varying testing accuracy, and recorded the number of people who were ever infected at any moment in time for each simulation (equivalent to the total number

of people who are in the “removed” pool, which could be seen in Fig. 2 & 3). We also recorded the average  $R_0$  throughout the simulation using the method described in “Calculation of  $R_0$  from Simulations.” To create the box plots in Figures 5-1 to 7-1, all box plots were the result of 10 new simulations with random starting nodes on each of 10 different, randomly-generated graphs, meaning each box plot represents 100 separate simulations. Similarly, the Figures 5-2 to 7-2 were generated by taking the average of 100 separate simulations’ average  $R_0$  values. All graphs are reproducible with our non-deterministic model by using similar parameters.

### 3.2 Validating Model with Real-World Data

As shown in Fig. 4, the real-world spreading of the virus in the United Kingdom has followed the usual curve expected by almost all epidemiological models, including our model. Using data compiled by Dong, Du, Gardner (2020), the graph of newly confirmed COVID-19 cases in Fig. 4 bears considerable resemblance to the positive tests section in Fig. 2 when considering Fig. 4 only plots the 7-day average. The periodic waxing and waning of confirmed cases in Fig. 2 are especially significant as indicators that our model adheres to realistic testing and spreading.



**Figure 4.** Real-world spreading in the United Kingdom (Jan-Jul 2020)

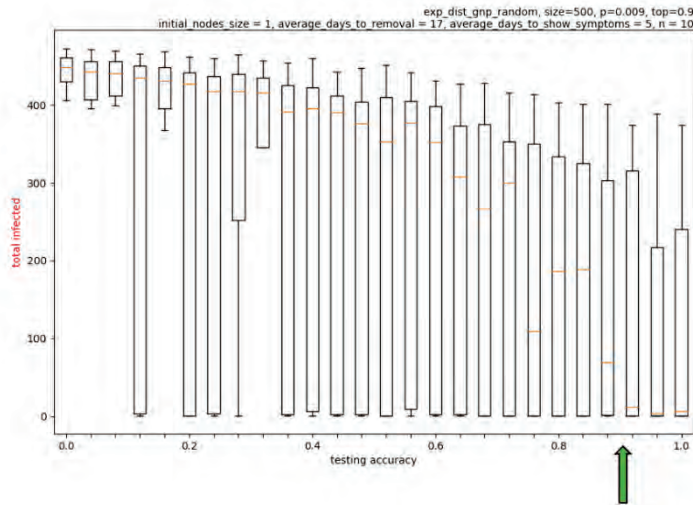
The time frame in Fig. 4 needs to be addressed. The data spans the beginning of the pandemic to July 6<sup>th</sup>, 2020, before the “2<sup>nd</sup> wave” of COVID-19 began in the UK and Europe in general. The focus on this timeframe is because it is most representative of spreading under the system of policies that our model strives to simulate—which the UK had in place during this general timeframe. After this time, many areas in the UK lifted restrictions prematurely in a “rush to reopen,” and were slow to respond to newly rising trends (Holder, Stevis-Gridneff & McCann, 2020). It should be noted once again that this study is not concerned with investigating the emergence of “2<sup>nd</sup> waves,” which does not have a formal scientific definition and is still subject of ongoing debate. While its causes are still being investigated, the virus’s resurgence in Europe in the autumn and winter of 2020 may be attributed to factors such as weather, holidays, political/economic pressure, and lifting social restrictions prematurely (Holder, 2020), all of which do not fall in the considerations of this investigation.

As seen in Fig. 5-1 to 7-2, the estimated  $R_0$  value for the spreading that our model simulates is slightly over 2 in general, which is in line with most recent estimates for the  $R_0$  value (Bauch, 2020). This confirmation further validates the network generation and realism of our model. In light of these observations, our

model is sufficient for describing real-world testing and accompanying policies and making inferences about testing accuracy.

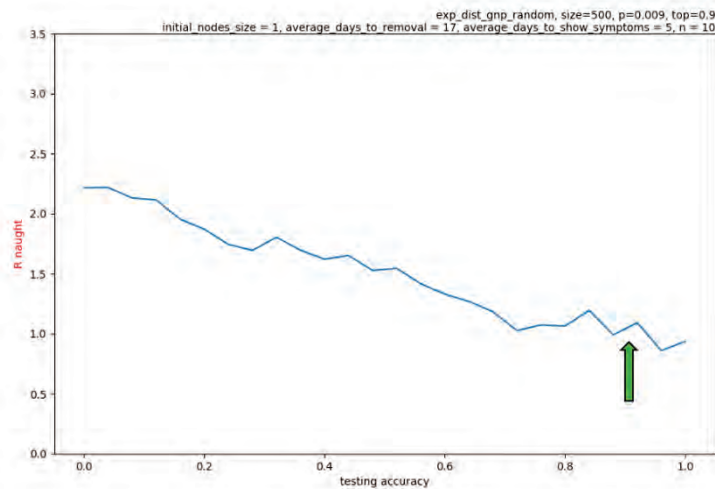
### 3.3 Threshold Behavior

The main significant observation made in this study is the emergence of threshold behaviors for different testing accuracies. When graphed using box plots, the data of total number infected against testing accuracy show a clear threshold behavior around certain testing accuracies, and this can be observed in networks with varying degrees of social interaction.

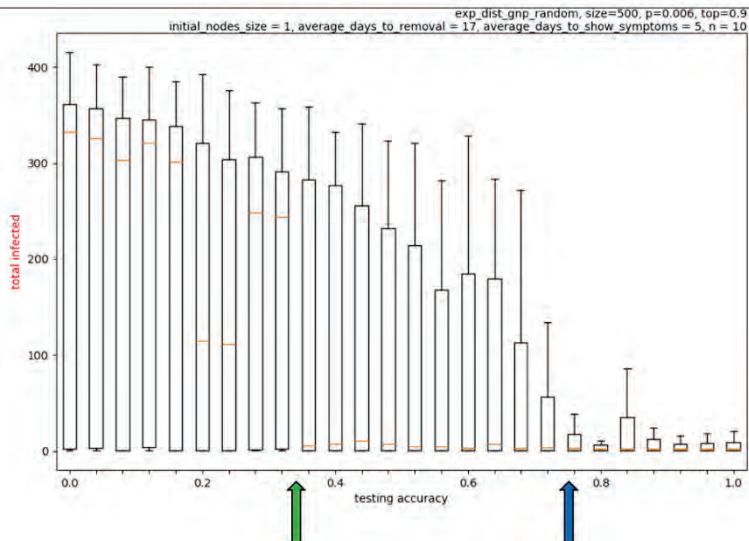


**Figure 5-1.** Total infected boxplot with  $np = 4.5$ . The green arrow marks a threshold.

In Fig. 5-1 and 5-2, the parameters determine that the average number of edges for each node in the network (average degree) is 4.5. This number is not very representative of real-world connections between individuals, as the random network does not very accurately depict friendships and connections in the real world, but instead represents the random contact network. We can see in this graph that a threshold could be identified at where testing accuracy is approx. 0.9; at this threshold, about half of the simulations indicate a “dying out” of the disease fairly quickly, and the upper quartile has steadily dropped to about half of the estimate for simulations without testing. This threshold coincides quite well with the transition across 1.0 for the average basic reproduction number in Fig. 5-2. It can be suggested that the distribution of  $R$  naught values over the 100 simulations would most likely have caused the great amount of variability in the number of infected for each simulation in Fig. 5-1. In general, the increase of testing accuracy has changed the behavior of the spreading process drastically around 0.9—a specific threshold—which is also reflected in the  $R$  naught quite precisely.

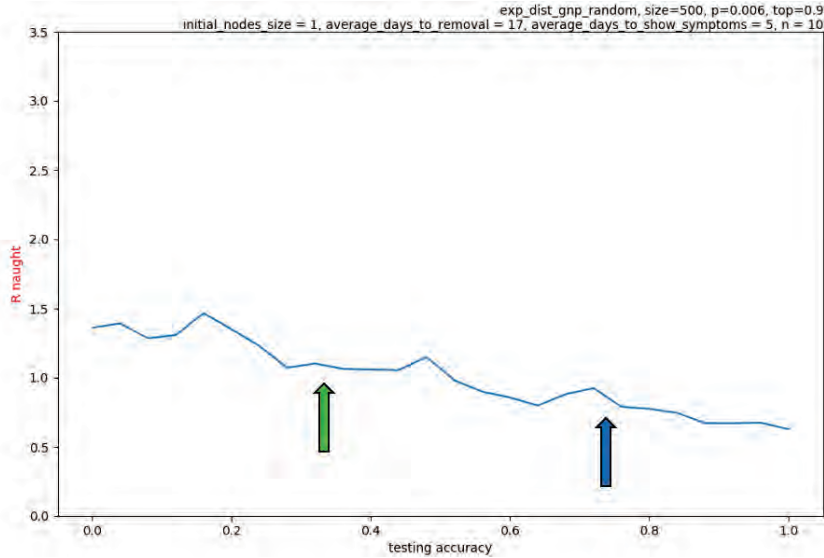


**Figure 5-2.** Average  $R$  naught plot with  $np = 4.5$ . The green arrow marks a threshold.

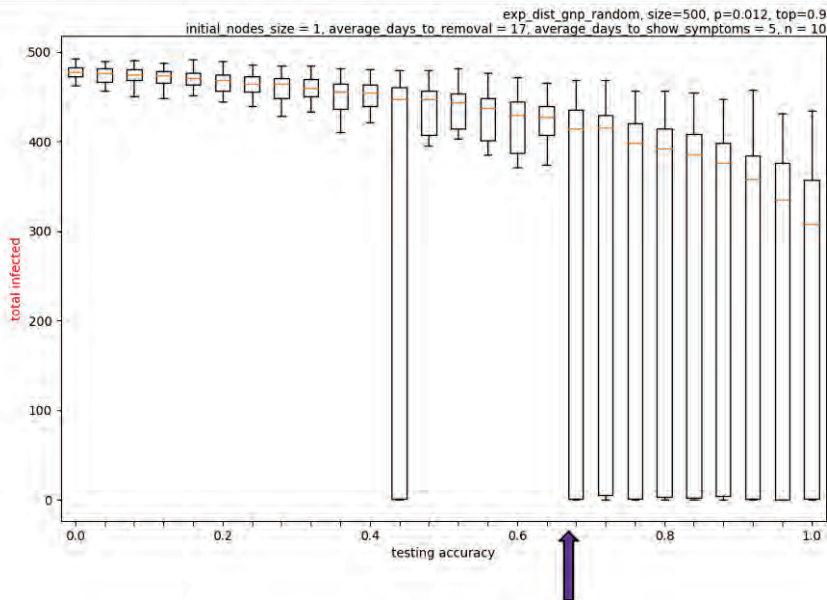


**Figure 6-1.** Total infected boxplot with  $np = 3$ . The arrows mark identified thresholds.

In Fig. 6, the average degree is dropped to 3. In this set of simulations, we observe more interesting dynamics involving two distinguishable thresholds. In Fig. 6-1, around where testing accuracy equals 0.35, we observe the type of threshold described in the previous paragraph—a threshold where the median effectively drops to 0 infected. This corresponds with the area in Fig. 6-2 where average  $R$  naught approximates to 1, like in Fig. 5-2. A second threshold could be observed around where testing accuracy equals 0.7 in Fig. 6-1, where the maximum drops to around 40 individuals, and where the upper quartile approaches 0. This second threshold in Fig. 6-1 corresponds to an approximate average  $R$  naught value of 0.75 in Fig. 6-2, which may indicate that the distribution of  $R$  naught values in these datasets rarely exceeded  $+0.25$  of the average value.



**Figure 6-2.** Average  $R_{naught}$  plot with  $np = 3$ . The arrows mark identified thresholds.

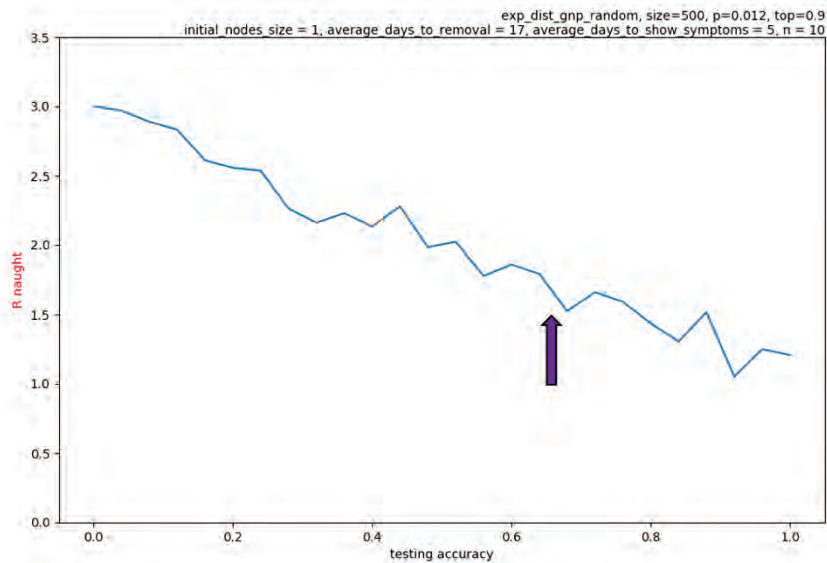


**Figure 7-1.** Total infected boxplot with  $np = 6$ . The arrow marks a threshold.

In Fig. 7-1 and 7-2, we observe the third type of testing accuracy thresholds. This threshold occurs around where testing accuracy is approx. 0.65. The data presented in Fig. 7-1 is most likely not precisely descriptive of the actual datasets around the said threshold, as the box plot was generated to omit outliers, and this threshold may be a threshold for where the plotting program determined certain datapoints to be outliers. Nevertheless, we can observe a threshold in that the lower quartile has been dropped to 0, indicating what is probably a shift in the composition of the set of  $R_{naught}$  values, with over a quarter of  $R_{naught}$  values becoming lower than 1. This corresponds to where the average  $R_{naught}$  equals



approx. 1.75 in Fig. 7-2. We see the average  $R$  naught value does not cross 1 in Fig. 7-2, which is reasonable.



**Figure 7-2.** Average  $R$  naught plot with  $np = 6$ . The arrow marks a threshold.

### 3.4 Real-World Implications of Threshold Behavior

Even though we have established that the model assembled in this study is not 100% realistic (see Appendix II: Drawbacks and Possible Improvements), there are some crude real-world implications that could be made from our findings aside from purely mathematical analyses. For example, using Figures 4-1 and 4-2, we can predict with some certainty that, if the contact number is 4.5, testing methods with over 90% accuracies will generate dramatically better results than those with lower accuracies.

Indeed, these contact numbers are grim. However, as discussed previously, these numbers are relatively inaccurate considerations in terms of individuals. It would be much more sensible to look at these numbers in terms of family-to-family interactions in a local community, as the relationship network of individuals are simplified in the random contact networks. Following this logic, we can make a few predictions:

Using Figures 5-1 and 5-2, we can predict that, if the average number of other households that a household randomly interacts with is about 3 and people test on symptomatic onset, the testing accuracy for the coronavirus needs only to be about 35% to create a 50% chance for quick and full eradication of the coronavirus. Under these circumstances, a testing accuracy slightly over 70% would lead to almost 100% certainty for quick and full eradication of the coronavirus. In comparison, using Figures 6-1 and 6-2, if the contact number is about 6 households, then it would be effectively useless to divert money and effort into any testing kits that are less than 65% accurate, and the community should look for kits that are at least 90% accurate, as that is a level beyond which results will improve considerably faster. In this sense predictions would imitate the social distancing guidelines in the UK more realistically.

All in all, there are many simplifications and intricate dynamics that makes this model only a very crude imitation of real life and should not be the sole basis for serious decisions. Much of these more intricate dynamics needs to be studied further



to revise this model to be more believable, as discussed in more detail in Appendix II: Drawbacks and Possible Improvements.

### 3.5 Other Observations and Speculations

Data from this model abundantly proves that testing, even with lower accuracies, can considerably impact the eventual number of individuals who are infected in an epidemic.

Also, the  $p$  value in the graph generation may slightly impact the concavity of the curve created with different testing accuracies. With higher  $p$  values, representing more social interaction, the infection distribution over the testing accuracy range tends to be slightly more convex, while with lower  $p$  values, representing fewer social ties, the infection distribution seems to be slightly more concave. However, this has been shown to vary between graphs, and cannot be established as conclusive. We hypothesize here that as testing accuracy increases, networks with sufficiently low or high levels of social interaction would receive diminishing, though still substantial results in increasing its ability to minimize the effects of the virus; while networks with intermediate levels of social interaction may yield increasingly substantial results as testing accuracy increases. A possible support for this claim is that in networks with lower levels of social interaction, the  $R$  naught is already relatively low, while for networks with very high levels of social interaction, there is a point at which higher testing accuracy cannot decrease the  $R$  naught sufficiently to make a considerable difference compared to lower testing accuracy. Graphs that might further this idea are in Appendix I: Graphs.

### 3.6 Summary

In general, the total number of patients infected decreases with higher testing accuracies. Provided that the epidemic was contained in the community, using parameters for COVID-19, there are three types of clear threshold behavior for the number of total infected when graphed against different testing accuracies:

1. Where 25% of  $RO$  values are below 1,
2. Where 50% of simulations have near 0 individuals infected (average  $RO$  drops below 1), and
3. Where almost all  $RO$  values fall below 1 (total number of individuals infected are limited to a small percentage of the population).

Lastly, from the data collected, it is indeterminable how exactly different levels of social interaction impact the effectiveness higher testing accuracy has in minimizing the effects of the coronavirus.

## 4. Theoretical Analysis of Phenomenon and Implications

### 4.1 Parameters and Levels of Impact

In containing the epidemic on a network, there is a certain point at which a high testing accuracy will be able to directly restrict the spread of the epidemic, at which point we might assume the  $R$  was reduced to below 1 sufficiently quickly. In our model, the idea of  $R_0$  and  $R$  as the primary measure of the spreading process is relatively unreliable, as they may fluctuate greatly throughout the simulation, and are

impacted by a significant number of parameters. However, we can consider a few ideas, as summarized by a conceptual expression:

$$R_0 \leftarrow np[1 - (1 - Q)^d]$$

*Note:* This expression above only demonstrates levels of impact in our model and is not a precise mathematical function.

Where  $n$  is the total number of nodes in the graph, and  $p$  is the p-value of the binomial graph.  $Q$  represents the average-effectual probability of infection (“closeness”) over a node’s edges, and  $d$  represents the average days before the nodes is removed. We consider the composition of  $d$  to be three parts:

$d_r$ : the average number of periods before a node is removed

$d_s$ : the average number of periods before a node becomes symptomatic

$t$ : testing accuracy

We notice here that the testing accuracy is limited greatly by  $d_s$  in our model’s representation of COVID-19, and that its impact on the  $R_0$  for any disease is highly dependent on  $d_s$ . However, the ability of the testing accuracy  $t$  at impacting  $R_0$  should not be readily dismissed, as it should be mentioned that a high testing accuracy would greatly reduce  $d$  if  $d_r$  is particularly high. The position of  $d$  as the exponent gives it a crucial role in controlling  $R_0$ . Secondly, we notice that  $Q$ , the probability of infection (“closeness”), plays a critical role in determining the scaling properties of  $d$ . When  $Q$  is sufficiently small, a higher  $d$  value would no longer create such a great impact as when  $Q$  is greater. We also observe that  $np$ , the average degree of each node, is an important scalar in affecting the entire spreading process, as we should keep in mind that  $R_0$  and  $R$  do not have a linear relationship with the actual numbers of increase in the “infectious” pool.

These observations conform with the types of threshold behaviors that we observe in our simulations. With a certain  $np$  as the scalar for  $R_0$ , we can see that, with too great a degree of social interaction, the impact of testing accuracy on the actual number of people infected is dwarfed significantly. The same rationale applies for simulations with very small  $np$ . These phenomena could be seen in Figures 4-1 through 7. Due to the length of COVID-19’s incubation phase, testing for the coronavirus has been made much more complex and challenging.

## 4.2 Implications for Further Research

In our results, these parameters and the explanation of their roles above mostly serve as an attempt at explaining the threshold behaviors of the computer simulations, and as a tool to put these ideas into perspective for a larger investigation into the dynamics of diseases simulated on the model assembled by this paper. Yet, there is much work that can be done on expanding the analysis of the different parameters’ effect on the basic reproduction number, and we may raise a few questions and hypotheses to supplement the main hypothesis raised in the Results section:

$Q$  is not explored much in this investigation; it would be interesting to see how  $Q$  may be determined from the model and how it would affect the ability of testing to change our results. In particular, as changes in  $Q$  may be observed in the real world as the degree of protection between individuals and personal hygiene, it would be interesting to see the effects of  $Q$  when taken together with  $np$ , which represents the width of each person’s contact network. Also, it would be very interesting if we could empirically understand the range of effects the incubation phase would impose on how the total number of individuals infected is changed by

testing accuracy. How does the model react when the difference between  $d_r$  and  $d_s$  is changed?

Also, it would be interesting to look further into the relationship between testing accuracy and  $R_0$ . The graphs in this paper seem to suggest a linear relationship, but may be limited by COVID-19 parameters for the incubation and recovery phase to fully explore the idea.

## 5. Discussion and Conclusion

### 5.1 General Implications of Results

The results of the simulations lead to meaningful confirmations and implications. First, testing accuracy plays a critical role in affecting the virus' reproduction number such that certain testing accuracy thresholds may appear for spreading processes, three of which has been summarized in the Results; by running our model, we can determine whether and where thresholds exist in given conditions. Second, the effectiveness of testing for a disease most likely depends considerably on the length of its incubation phase, but in general, our model corroborates the presumption that with higher testing accuracy, the reproduction number should decrease dramatically as the infectious node is rendered non-infectious much earlier. Finally, in the case of COVID-19, it should be further investigated how low, intermediate, and high levels of social interaction would impact the effectiveness of testing in minimizing the effects of the coronavirus, especially on whether there is a "window" for which increasing testing accuracy becomes particularly effective.

### 5.2 Further Connections

This study builds on previous studies in the field in that it provides a different approach to the complexity of different factors in an epidemic, through its computational model. Specifically, it looks at the effectiveness of testing in lowering the basic reproduction number of COVID-19, a disease characterized by its long incubation phase. It also uses mathematical analysis and simulation data to confirm the ability of testing in containing the coronavirus. This study furthers previous studies by connecting the ideas of social isolation with testing accuracy and has possibly offered a basic method of determining acceptable levels of testing accuracy based on the level of social isolation.

## References

- Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., ... & Xia, L. (2020). Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology*, 200642.
- Anderson, R. M., & May, R. M. (1979). Population biology of infectious diseases: Part I. *Nature*, 280(5721), 361-367.
- Bauch, C. T. (2020). Estimating the COVID-19 R number: A bargain with the devil? *The Lancet*. doi:10.1016/s1473-3099(20)30840-9
- BBC (2020, March 30). Coronavirus: Countries reject Chinese-made equipment. <https://www.bbc.com/news/world-europe-52092395>

- (2020, April 9). Coronavirus: Worst economic crisis since 1930s depression, IMF says. <https://www.bbc.com/news/business-52236936>
- (2020, April 28). Coronavirus: India cancels order for 'faulty' China rapid test kits. <https://www.bbc.com/news/world-asia-india-52451455>
- Cai, L. M., Li, X. Z., & Ghosh, M. (2009). Global stability of a stage-structured epidemic model with a nonlinear incidence. *Applied Mathematics and Computation*, 214(1), 73-82.
- Cascella, M., Rajnik, M., Cuomo, A., Dulebohn, S. C., & Di Napoli, R. (2020). Features, evaluation and treatment coronavirus (COVID-19). In *Statpearls [internet]*. StatPearls Publishing.
- Centers for Disease Control and Prevention. (2020, June 30). *Interim Clinical Guidance for Management of Patients with Confirmed Coronavirus Disease (COVID-19)*. <https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-guidance-management-patients.html>
- Corman, V. M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D. K., ... & Mulders, D. G. (2020). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance*, 25(3), 2000045.
- Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet*, 20(5), 533-534. doi:10.1016/s1473-3099(20)30120-1
- Epidemiology Working Group for NCIP Epidemic Response, Chinese Center for Disease Control and Prevention (2020). [The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) in China]. *Zhonghua Liu Xing Bing Xue Za Zhi*. 2020;41(2):145-151. doi:10.3760/cma.j.issn.0254-6450.2020.02.003
- Friedman, E., Friedman, J., Johnson, S., & Landsberg, A. (2020). Transitioning out of the coronavirus lockdown: A framework for zone-based social distancing. *arXiv preprint arXiv:2004.08504*.
- Gallagher, J. (2020, February 13). Are coronavirus tests flawed? BBC. <https://www.bbc.com/news/health-51491763>
- Gray, A., Greenhalgh, D., Hu, L., Mao, X., & Pan, J. (2011). A stochastic differential equation SIS epidemic model. *SIAM Journal on Applied Mathematics*, 71(3), 876-902.
- Gray, A., Greenhalgh, D., Mao, X., & Pan, J. (2012). The SIS epidemic model with Markovian switching. *Journal of Mathematical Analysis and Applications*, 394(2), 496-516.
- Guan, W. J., Ni, Z. Y., Hu, Y., Liang, W. H., Ou, C. Q., He, J. X., ... & Du, B. (2020). Clinical characteristics of coronavirus disease 2019 in China. *New England Journal of Medicine*, 382(18), 1708-1720.
- Holder, J., Stevis-Gridneff, M., & McCann, A. (2020, December 4). Europe's deadly second wave: how did it happen again? *The New York Times*. Retrieved from <https://www.nytimes.com/interactive/2020/12/04/world/europe/europe-covid-deaths.html>
- Iannelli, M., Kim, M. Y., & Park, E. J. (1999). Asymptotic behavior for an SIS epidemic model and its approximation. *Nonlinear Analysis: Theory, Methods & Applications*, 35(7), 797-814.
- Ji, C. & Jiang, D. (2014). Threshold behaviour of a stochastic SIR model. *Applied Mathematical Modelling*, 38(21-22), 5067-5079.
- Katriel, G. (2010). Epidemics with partial immunity to reinfection. *Mathematical biosciences*, 228(2), 153-159.

- Kirkpatrick, D. D. & Bradley, J. (2020, May 7). U.K. paid \$20 million for new coronavirus tests. they didn't work. *The New York Times*. <https://www.nytimes.com/2020/04/16/world/europe/coronavirus-antibody-test-uk.html>
- Kleinberg, J. & Easley, D. (2010). *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge UP.
- May, R. M., & Anderson, R. M. (1979). Population biology of infectious diseases: Part II. *Nature*, 280(5722), 455-461.
- Mayo Clinic. (2020, July 3). Coronavirus disease 2019 (COVID-19). <https://www.mayoclinic.org/diseases-conditions/coronavirus/symptoms-causes/syc-20479963>
- McCluskey, C. C. (2010). Complete global stability for an SIR epidemic model with delay—distributed or discrete. *Nonlinear Analysis: Real World Applications*, 11(1), 55-59.
- National Health Service. (2021, January 21). Your coronavirus test result. <https://www.nhs.uk/conditions/coronavirus-covid-19/testing-and-tracing/what-your-test-result-means>
- Newman, M. E., Strogatz, S. H., & Watts, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Physical review E*, 64(2), 026118.
- Ruan, S., & Wang, W. (2003). Dynamical behavior of an epidemic model with a nonlinear incidence rate. *Journal of Differential Equations*, 188(1), 135-163.
- Tornatore, E., Buccellato, S. M., & Vetro, P. (2005). Stability of a stochastic SIR system. *Physica A: Statistical Mechanics and Its Applications*, 354, 111-126.
- World Health Organization. *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)* (2020). Available at: <https://www.who.int/docs/default-source/coronaviruse/whochina-joint-mission-on-covid-19-final-report.pdf>
- Zaman, G., Kang, Y. H., & Jung, I. H. (2008). Stability analysis and optimal vaccination of an SIR epidemic model. *BioSystems*, 93(3), 240-249.

## Appendix I: Additional Graphs

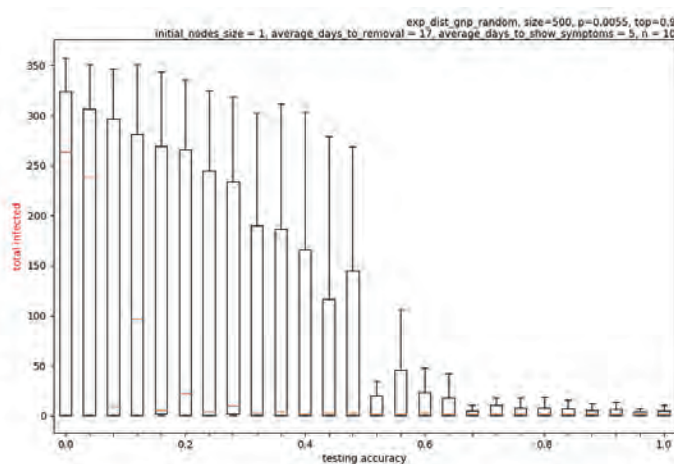
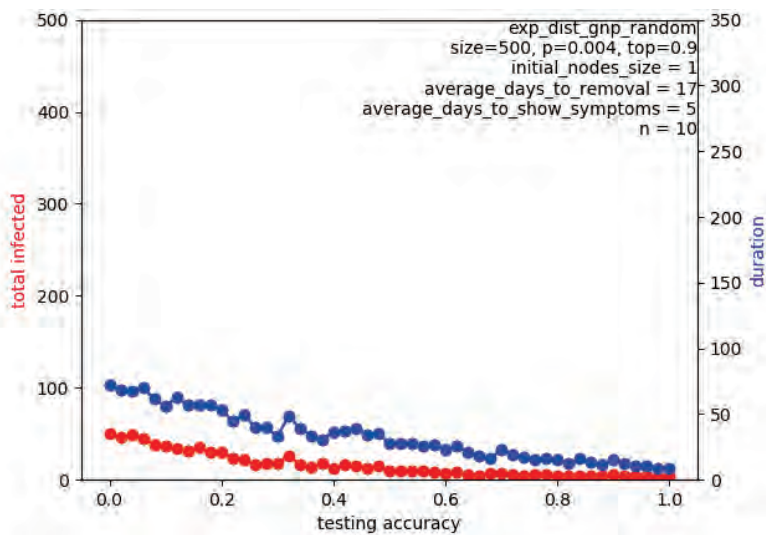


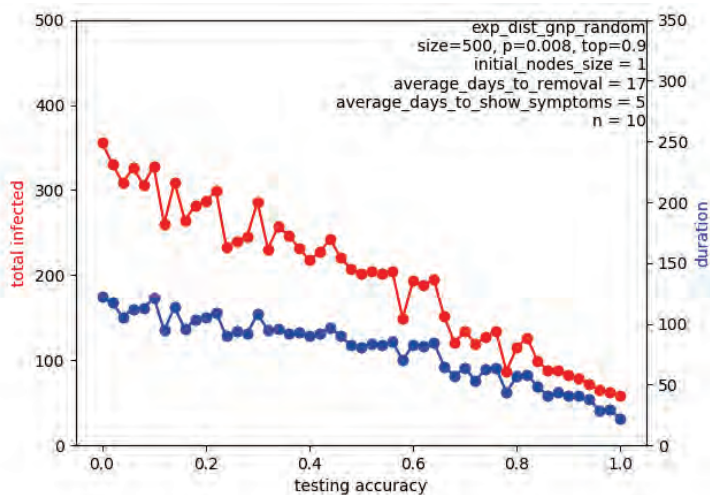
Figure 8. Total infected boxplot with  $np = 2.75$ .

In the following graphs, instead of using box plots to tackle the variability of our non-deterministic model, the total number of people infected is averaged. All data points were the result of 10 new simulations with random starting nodes on 10 different, randomly generated graphs, making each datapoint on the graphs the average value for 100 simulations.

It is possible to suggest that simply in terms of percentage, Fig. 10 has the greatest reduction with increasing testing accuracy, while Fig. 9 has diminishing returns. Also, while a testing accuracy higher than 50% was very important for Fig. 11's significant decrease in cases, the effects of higher testing accuracy in Fig. 12 seem to have become slightly lower in scale, though still very substantial. This idea is also present in the comparison between Fig. 8 and Fig. 7-1, though more data would need to be obtained to make further connections.



**Figure 9.** Total infected and duration with  $np = 2$ .



**Figure 10.** Total infected and duration with  $np = 4$ .

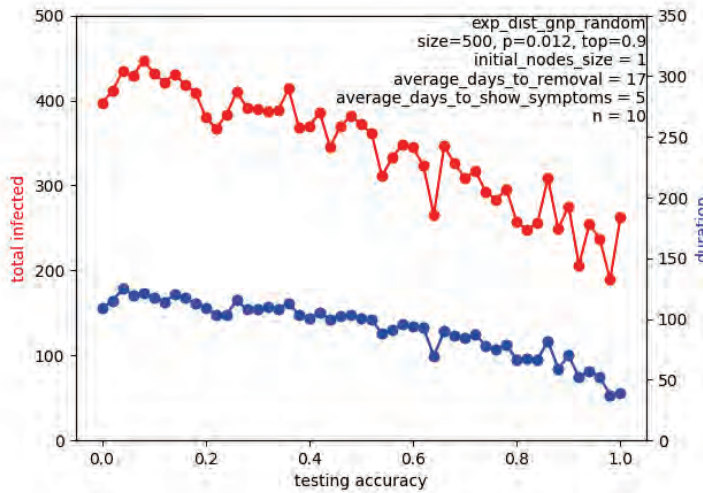


Figure 11. Total infected and duration with  $np = 6$ .

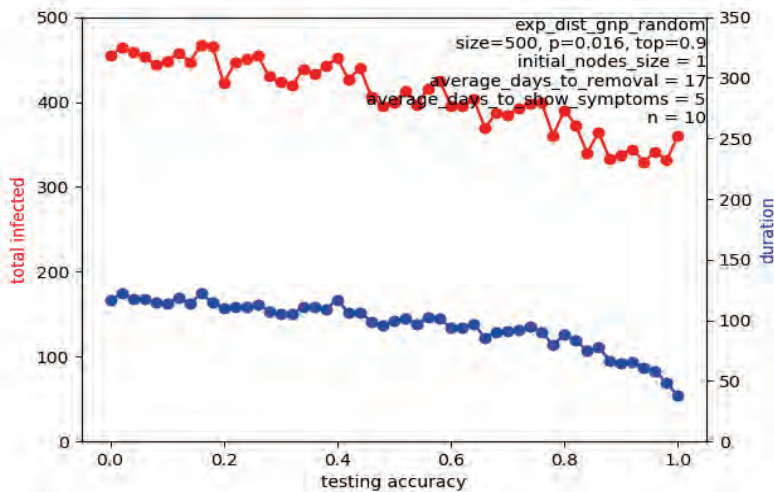


Figure 12. Total infected and duration with  $np = 8$ .

## Appendix II: Drawbacks and Possible Improvements

### Network Structure

As described by Newman, Watts, & Strogatz (2001), random graphs are not representative of long-term relationships and contact networks between individuals. It must be pointed out that this study does not take into account that this idea may suggest there are better underlying network structures to use for our simulations, since our model does simulate the spread of diseases over relatively long time periods. It may be helpful to compare the results from this study’s with results done on certain small-world networks or scale-free networks.

Also, this study does not consider the structure of families and their impact on the spreading process. It may be suggested that families could be simplified as

single nodes in random networks, but a more specific look into the effects of such social networks would be very helpful.

Along with family structure, there is the issue of public gatherings. This study disregards the possibility of public gatherings or public spaces with numerous visitors, such as supermarkets.

### **Simulation of Testing and Isolation**

In terms of testing, this study glosses over many details. First, there is the possibility of individuals receiving multiple tests over time as symptoms develop. Introducing this idea would dramatically complicate the model in terms of the impact testing accuracy has, but it is an idea that needs to be considered. Furthermore, this study does not consider the situation of isolation when moderate and severe symptoms develop, at which point healthcare workers are capable of identifying COVID-19 cases from other indicators.





# Cube-plate Tiling Numbers and Their Identities

---

Qianyu Guo

*Author Background:* Qianyu Guo grew up in China and attends Shanghai YK Pao School in Shanghai, China. His Pioneer research concentration was in the field of mathematics and was titled “Fibonacci Numbers and Visual Proofs.”

## 1. Introduction

Fibonacci numbers are a well-known sequence of integers recursively defined as  $f_0 = 1$ ,  $f_1 = 1$ , and  $f_n = f_{n-1} + f_{n-2}$  for every  $n > 2$  and  $n \in \mathbf{Z}$ . Yet in the book *Proofs that Really Count* by Art Benjamin and Jennifer Quinn, these numbers are redefined in a combinatorics context, where a linear board is tiled with squares and dominoes only. By conditioning on the last tile of every tiling of board with length  $n$ , the case of square will give  $f_{n-1}$  and the case of domino will give  $f_{n-2}$ . When they are added together, it becomes the defining equation of the Fibonacci sequence,  $f_n = f_{n-1} + f_{n-2}$ . Also, the first few numbers of board tiling are also the same as Fibonacci numbers. In this way,  $f_n$  represents the number of ways to tile a linear board with length  $n$  ( $n \in \mathbf{Z}$ ; in this paper, all variables are in  $\mathbf{Z}$  unless otherwise noted) with squares and dominoes only. With this visual definition, most Fibonacci identities can be proved with ease.

Tiling a board with squares, dominoes, and other polyominoes can be considered as 2-D tilings. The captivating properties of board tiling and its variants have been explored by many combinatorists. Naturally, we may consider 3-D tiling, where we “tile” a large box with different units of building blocks. Specifically, we want to tile a box with dimensions  $2 \times 2 \times n$ . There are a range of dimensions building blocks may have, but we will specifically address two of them:  $1 \times 1 \times 1$  cube, and  $1 \times 2 \times 2$  plate.

Define  $k_n$ , the cube-plate tiling numbers, as the number of ways to tile a  $2 \times 2 \times n$  box with  $1 \times 1 \times 1$  cubes, and  $1 \times 2 \times 2$  plates. The following table shows the first few values of  $k_n$ :

$n$	0	1	2	3	4	5	6
$k_n$	1	2	10	36	144	556	2172

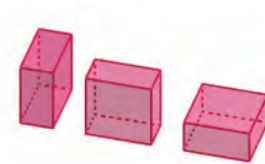
(Again,  $k_0$  is defined to be 1 because it is convenient and there is only one way to tile a  $2 \times 2 \times 0$  box, i.e., to use exactly 0 cubes and 0 plates. This sequence was

devised by the author with the help from his professor and recorded as A335559 in the On-line Encyclopedia of Integer Sequences (OEIS).)

At first glance, it is not obvious that  $k_n$  is produced from previous  $k$  values. However, since this problem originates from board tiling, we are confident in predicting that cube-plate tiling numbers follow some type of recursive relation. This essay will determine the recursive relation of  $k_n$  using combinatorics analysis and explore some of its identities in detail. While the proof of theorem 3 was inspired by *Proofs that Really Count*, all the theorems, corollaries and their proofs in this essay were solely devised by the author.

## 2. Recursive Relation of $k_n$

At first, we need to clarify what a  $2 \times 2 \times 1$  rectangular prism refers to. By stating  $2 \times 2 \times 1$ , it does not mean the rectangular prism has  $1 \times 2$  rectangle for both the front and top view, and a  $2 \times 2$  square for the side view. It can be in any of the three orientations below:

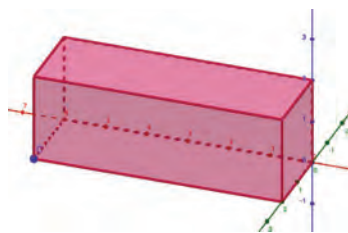


**Figure 1:** Three orientations of plates.

For convenience's sake, the family of these rectangular prisms with dimensions 2, 2, 1 is named *plates*. The three orientations of plates are *side plates*, *face plates*, and *flat plates*, from left to right, respectively.

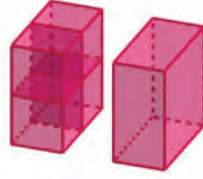
However, for the  $2 \times 2 \times n$  box, we consider its orientation fixed, i.e. always has  $2 \times 2$  square as the side view and  $2 \times n$  rectangle as both the front and top view, to keep the problem simple.

In the introduction, the first few values for  $k_n$  are given. Now we will have a closer look at how those values are obtained.



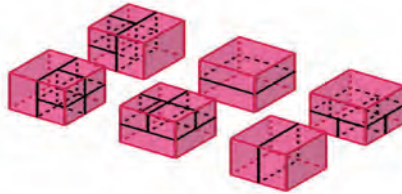
**Figure 2:** The orientation of box.

$k_1 = 2$  is obvious, as there are only two ways to tile a  $2 \times 2 \times 1$  box: with four  $1 \times 1 \times 1$  cubes, and one side plate (Figure 3).



**Figure 3:** Two ways to tile a  $2 \times 2 \times 1$  box.

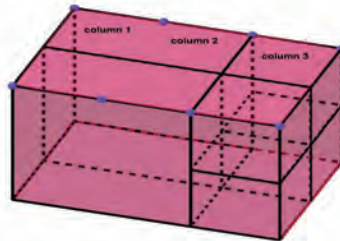
For  $k_2$ , things become a little more complicated. Firstly, we can cut a  $2 \times 2 \times 2$  box in the middle so it becomes two  $2 \times 2 \times 1$  boxes, each of which has  $k_1 = 2$  ways to tile. There are  $2 \times 2 = 4$  ways of this configuration. Yet face and flat plates can be used to tile a  $2 \times 2 \times 2$  box. For each, there are 3 additional ways. So, there are 6 ways in total, shown in Figure 4.



**Figure 4:** Six ways to tile a  $2 \times 2 \times 2$  box.

Therefore,  $k_2 = 4 + 6 = 10$ .

In the above exploration, we discover that the box may be *breakable*, i.e., cut into two smaller boxes, at column  $i$  where there are no face or flat plates connecting column  $i$  and  $i + 1$ , and *unbreakable* if there is a face or flat plate connecting column  $i$  and  $i + 1$ . For example, in the diagram below, we say that column 1 is unbreakable (since face plates cover both columns 1 and 2) and column 2 is breakable.



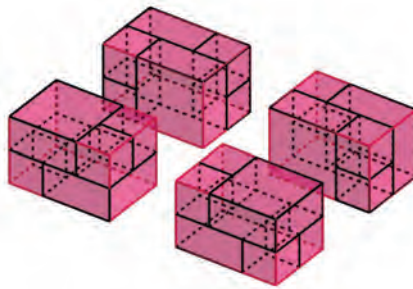
**Figure 5:** Breakability in  $2 \times 2 \times 3$  box.

The concept of breakability is useful in finding larger values of  $k_n$ . For  $k_3$ , consider the last column which is breakable.

(i) If it breaks at column 2, then it can be considered as a combination of  $2 \times 2 \times 2$  box and a  $2 \times 2 \times 1$  box, where there are  $2 \times 10 = 20$  ways.

(ii) If it breaks at column 1, then it can be considered as a combination of  $2 \times 2 \times 1$  box and an unbreakable  $2 \times 2 \times 2$  box. From the discussion of  $k_2$ , there are six ways to tile an unbreakable  $2 \times 2 \times 2$  box (using flat or face plates). So there are  $2 \times 6 = 12$  ways.

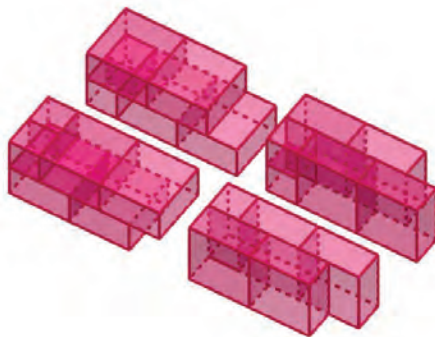
(iii) If it never breaks at columns 1 and 2, then there are four cases, as illustrated below:



**Figure 6:** Four cases that the tiling never breaks

Therefore,  $k_3 = 20 + 12 + 4 = 36$ .

Notice that for  $n \geq 3$ , there are always four ways that the tiling of  $2 \times 2 \times n$  box is unbreakable at any column  $i$  ( $1 \leq i \leq n - 1$ ), similar to  $n = 3$ , where every face or flat plates overlap each other:



**Figure 7:** Starting patterns of unbreakable tilings.

Using this important property and analysis of breakability, we may reach our first recursive relation of  $k_n$ :

**Theorem 1.** For every  $n \geq 3$ ,

$$k_n = 2k_{n-1} + 6k_{n-2} + 4 \sum_{i=0}^{n-3} k_i$$

*Proof.* Conditioning on the last column of  $2 \times 2 \times n$  box tiling that is breakable:

If it lastly breaks at column  $n - 1$ , there are  $k_{n-1} \times k_1 = 2k_{n-1}$  tilings.

If it lastly breaks at column  $n - 2$ , there are  $6k_{n-2}$  tilings, as 6 tilings of  $2 \times 2 \times 2$  box are unbreakable at column 1.

If it lastly breaks at column  $c$  ( $1 \leq c \leq n - 3$ ), there are  $4k_c$  tilings for each  $c$ , since 4 tilings of each  $2 \times 2 \times (n - c)$  box are unbreakable at column  $j$ ,  $1 \leq j < n - c$ .

If the tiling never breaks at any column  $i$  ( $1 \leq i \leq n - 1$ ), we can consider it breaking at column 0. According to our discovery, there are exactly four tilings for  $2 \times 2 \times n$  ( $n \geq 3$ ) box that never break. Therefore, there are  $4k_0$  tilings.

Adding these up, we have

$$k_n = 2k_{n-1} + 6k_{n-2} + 4k_{n-3} + 4k_{n-4} + \dots + 4k_1 + 4k_0 = 2k_{n-1} + 6k_{n-2} + 4 \sum_{i=0}^{n-3} k_i$$

□

There are some useful corollaries of this theorem.

**Corollary 1.**

$$\sum_{i=0}^n k_i = \frac{k_{n+3} - 2k_{n+2} - 6k_{n+1}}{4}$$

*Proof.* In Theorem 1, solving for the term  $\sum_{i=0}^{n-3} k_i$  gives us

$$\sum_{i=0}^{n-3} k_i = \frac{k_n - 2k_{n-1} - 6k_{n-2}}{4}$$

Substitute  $n - 3$  with  $n$ , we have

$$\sum_{i=0}^n k_i = \frac{k_{n+3} - 2k_{n+2} - 6k_{n+1}}{4}$$

□

Corollary 1 is the formula for the sum of first  $n + 1$  terms (or the series) of  $\{k_n\}$ . From Theorem 1, we may also deduce the following corollary:

**Corollary 2.**

$$k_n = 3k_{n-1} + 4k_{n-2} - 2k_{n-3} \text{ for } n \geq 3$$

*Proof.* By subtracting two adjacent terms of  $\{k_n\}$ , we get

$$k_n - k_{n-1} = 2k_{n-1} + 6k_{n-2} + 4k_{n-3} + 4 \sum_{i=0}^{n-4} k_i - (2k_{n-2} + 6k_{n-3} + 4 \sum_{i=0}^{n-4} k_i)$$

$$k_n - k_{n-1} = 2k_{n-1} + 4k_{n-2} - 2k_{n-3}$$

$$k_n = 3k_{n-1} + 4k_{n-2} - 2k_{n-3}$$

□

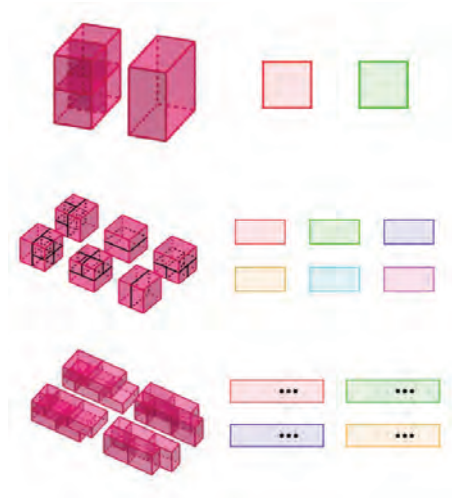
Corollary 2 is the recursive formula without the series of  $\{k_n\}$ . Using this corollary, we may easily compute the first few values of  $k_n$  in introduction by hand.

### 3. An alternative interpretation of $k_n$

The three-dimensional nature of cube-plate tiling makes it difficult to imagine and, therefore difficult to prove theorems using combinatorial argument. It is better to establish a bijection between cube-plate tiling and a certain type of board tiling, as seen here.

Any cube-plate tiling of  $2 \times 2 \times n$  box consists of smaller unbreakable  $2 \times 2 \times 1$  tilings,  $2 \times 2 \times 2$  tilings, or  $2 \times 2 \times k$  tilings with integer  $k \geq 3$ . As shown before, there are 2 types of unbreakable  $2 \times 2 \times 1$  tilings, 6 types of  $2 \times 2 \times 2$ , and 4 types of  $2 \times 2 \times k$ .

Consider cube-plate tiling of  $2 \times 2 \times n$  box and tiling of a length- $n$  board, using 2 colors of squares (or monominoes), 6 colors of dominoes, and 4 colors of linear length- $k$  polyominoes (henceforth abbreviated as “ $k$ -minoos”). Then, there is one-to-one correspondence between unbreakable  $2 \times 2 \times 1$  tiling and square, unbreakable  $2 \times 2 \times 2$  tiling and domino, and unbreakable  $2 \times 2 \times k$  tiling and  $k$ -minoos. In this way, we establish a bijection between cube-plate tiling of  $2 \times 2 \times n$  box and colored tiling of a length- $n$  board.



**Figure 8:** Bijection between “tiles” of two tiling problems.

Similarly, we define a cell  $i$  to be “breakable” if there is not a domino or  $k$ -mino covering both cell  $i$  and  $i + 1$ , and “unbreakable” if there is.

### 4. Further theorems of $k_n$ and proofs

Like Fibonacci numbers, cube-plate tiling numbers have quite a few curious identities. This section comes up with two other identities and uses tiling to prove them. All theorems and corollaries below can be found at A335559 in OEIS.

**Theorem 2.**

$$k_{m+n} = k_m k_n + 2k_{m-1} k_{n-1} + \frac{1}{4} (k_m - 2k_{m-1} - 2k_{m-2})(k_n - 2k_{n-1} - 2k_{n-2})$$

*Proof.* We will use the alternative interpretation of  $k_n$  to prove this identity. Consider tiling a length- $(m + n)$  board with squares of 2 colors, dominoes of 6 colors, and  $r$ -minoes of 4 colors. From the previous section, we know that there are  $k_{m+n}$  ways to achieve that.

On the other hand, conditioning on the breakability of cell  $m$ :

- (i) If it breaks at cell  $m$ , by multiplication principle, there are  $k_m k_n$  tilings;
- (ii) If it does not break at cell  $m$ , there must be a particular domino or  $r$ -mino ( $r \geq 3$ ) occupying both cell  $m$  and  $m + 1$ . Again, condition on the last cell this special domino or  $r$ -mino:

(a) If it ends at cell  $m + 1$ , then it can be both a domino or a  $r$ -mino. There are 6 colors for a domino, and 4 colors for a  $r$ -mino. Depending on the value of  $r$ , the number of cells to the left of this  $r$ -mino varies from 0 to  $m - 2$ . So there are

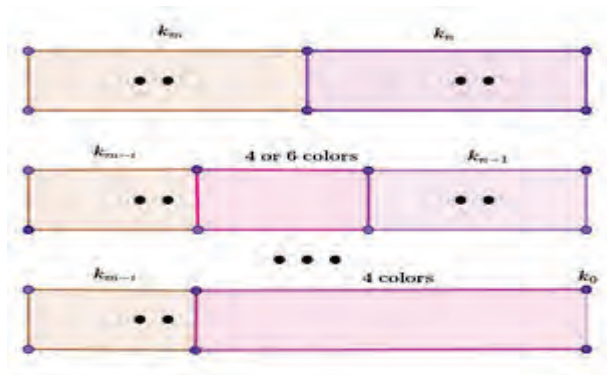
$$6k_{m-1}k_{n-1} + 4k_{n-1} \sum_{i=0}^{m-2} k_i$$

possible tilings in total.

(b) If it ends at any cell to the left of  $m + 1$ , it can only be a  $r$ -mino as it must cross over cells  $m$  and  $m+1$ . For each endpoint  $j$  such that  $m + 2 \leq j \leq m + n$ , depending on the value of  $r$ , there are from 0 to  $m - 1$  cells to the left of  $r$ -mino, 4 colors for the  $r$ -mino in the middle, and from 0 to  $n - 2$  cells to the right of  $r$ -mino. Therefore, there are

$$4 \sum_{i=0}^{m-1} k_i \sum_{j=0}^{n-2} k_j$$

possible tilings.



**Figure 9:** Conditioning on the special domino.

Adding these up, we have

$$\begin{aligned} & k_m k_n + 6k_{m-1}k_{n-1} + 4k_{n-1} \sum_{i=0}^{m-2} k_i + 4 \sum_{i=0}^{m-1} k_i \sum_{j=0}^{n-2} k_j \\ &= k_m k_n + 2k_{m-1}k_{n-1} + 4k_{n-1} \sum_{i=0}^{m-1} k_i + 4 \sum_{i=0}^{m-1} k_i \sum_{i=0}^{n-2} k_i \\ &= k_m k_n + 2k_{m-1}k_{n-1} + 4 \sum_{i=0}^{m-1} k_i \sum_{i=0}^{n-1} k_i \end{aligned}$$



possible colored tilings for a length- $m + n$  board in total.

Recall Corollary 1 and 2 from Section 3. Using those formulas gives us

$$\begin{aligned} & k_m k_n + 2k_{m-1} k_{n-1} + \frac{1}{4}(k_{m+2} - 2k_{m+1} - 6k_m)(k_{n+2} - 2k_{n+1} - 6k_n) \\ &= k_m k_n + 2k_{m-1} k_{n-1} + \frac{1}{4}(3k_{m+1} + 4k_m - 2k_{m-1} - 2k_{m+1} - 6k_m)(3k_{n+1} + 4k_n \\ &\quad - 2k_{n-1} - 2k_{n+1} - 6k_n) \\ &= k_m k_n + 2k_{m-1} k_{n-1} + \frac{1}{4}(k_m - 2k_{m-1} - 2k_{m-2})(k_n - 2k_{n-1} - 2k_{n-2}) \end{aligned}$$

Since the above formula and  $k_{m+n}$  both count colored tiling of length  $m + n$  board,

$$k_{m+n} = k_m k_n + 2k_{m-1} k_{n-1} + \frac{1}{4}(k_m - 2k_{m-1} - 2k_{m-2})(k_n - 2k_{n-1} - 2k_{n-2})$$

□

**Theorem 3.**

$$(28)(3^n) + 10 \sum_{i=0}^n 3^{n-i} k_i = 3k_{n+2} - 2k_n$$

*Proof.* This proof will use Corollary 2 as a primary definition of  $k_n$ , i.e., for every  $n \geq 3$ ,

$$k_n = 3k_{n-1} + 4k_{n-2} - 2k_{n-3},$$

with initial conditions  $k_0 = 1, k_1 = 2$ , and  $k_2 = 10$ . Assign each square a weight of 3, each domino a weight of 4, and each tromino a weight of  $-2$ , **except for the initial tile**. Define the weight of the tiling to be the product of weights of all the tiles used in that tiling. For example, a tiling of “tromino-domino-squaredomino-domino-tromino” has weight  $(3)(4)^3(-2)^2$ . Then, define  $k_n$  ( $n \geq 3$ ) to be the sum of the weights of all weighted tilings of length- $n$  board. In this way, for every  $n \geq 3$ , it can be easily proved that  $k_n = 3k_{n-1} + 4k_{n-2} - 2k_{n-3}$  by conditioning on the last tile. However, the initial conditions of  $k_1$  and  $k_2$  are not satisfied. Therefore, for the initial tiles specifically, a starting square has a weight of 2, a starting domino has a weight of 4, and a starting tromino has a weight of  $-2$ . In this way, we ensure that  $k_1 = 2$  and  $k_2 = 2 \times 3 + 4 = 10$  is well-defined. This special technique, again, is introduced in Section 3.5. of *Proofs that Really Count*.

In this context, consider the sum of weights of weighted tilings of an  $n + 3$ board ending with a square or tromino. At first, condition on the last tile of the board. The sum of the weights of tilings that end with a square is  $3k_{n+2}$ , and the sum of the weights of tilings that end with a tromino is  $-2k_n$ . So the desired sum of weights is  $3k_{n+2} - 2k_n$ .

On the other hand, condition on the location of last domino or tromino. We count the sum of weights in  $n + 3$ -tilings whose last domino or tromino begins at cell  $i + 1$ . Note that the domino cannot begin at  $n + 2$  as that violates the ending condition. Therefore, cells  $i + 1, i + 2,$  and  $i + 3$  consist of either a tromino (which has weight  $-2$ ) or a domino followed by a square (which has weight  $3 \times 4 = 12$ ). To the left of these three cells, cells  $1$  through  $i$  can be tiled arbitrarily, and the sum of weights is  $k_i$ . To the right of them follows  $n - i$  squares, with the weight  $3^{n-i}$ . So far, the sum of weight is  $\sum_{i=1}^n (12 - 2)3^{n-i}k_i = 10 \sum_{i=1}^n 3^{n-i}k_i$ .

The tilings that we ignore are those that begin with special weightings and followed by all squares. The sum of weights of tilings that begin with a tromino and followed by all squares is  $-2(3)^n$ . Otherwise, we have a 2-tiling followed by all squares, with sum of weights of tilings  $k_2(3)^{n+1} = 10(3)^{n+1} = 30(3)^n$ . Altogether, when we add up all the subtotals, the sum of weights of weighted tilings of an  $n + 3$ -board ending with a square or tromino can be represented as

$$30(3)^n - 2(3)^n + 10 \sum_{i=1}^n 3^{n-i}k_i = 28(3)^n + 10 \sum_{i=1}^n 3^{n-i}k_i$$

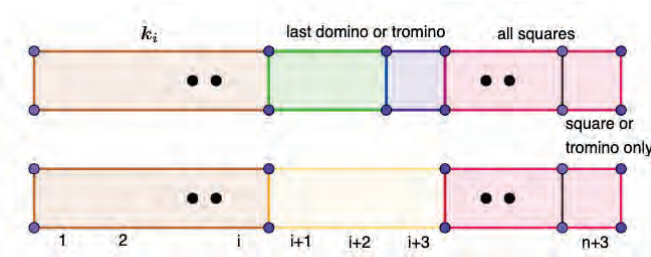


Figure 10: Conditioning on the last domino or tromino.

Since we were counting the desired sum of weights in two ways, we have

$$28(3)^n + 10 \sum_{i=1}^n 3^{n-i}k_i = 3k_{n+2} - 2k_n$$

□

## 5. Bibliography

Benjamin, Arthur T., and Jennifer J. Quinn. *Proofs That Really Count*. Mathematical Association of America, 2003.



# Uncovering the Invisibility: Protecting LGBT Groups in Armed Conflicts and After

---

*Qing Wan*

**Author Background:** *Qing Wan grew up in China and currently attends Nanjing Jinling High School International Department in Nanjing, China. His Pioneer research concentration was in the field of sociology/political science and titled "Gender and Political Conflict."*

## Introduction

The rights of lesbian, gay, bisexual, and transgender people (LGBT) are largely neglected in peacetime because they seldom fall into the prevalent notion of heterosexual norms, which are championed in nearly all places and cultures. These groups are also often victims of targeted violence, which they might attempt to minimize by not revealing their sexual orientation and identity. In times of armed conflict, LGBT community would face more challenges because of the volatile and violent nature of conflict (Park & Mykhyalyshyn 2016) that exposes everyone to danger, and in particular minority groups. One report from Williams Institutes records that soldiers from the Communist Party of Nepal regard LGBT people as easy targets and scapegoats: they can do "anything" to them without facing any consequence (2014, 57-59). This phenomenon thus forms a loop: targeted with violence, LGBT people strategically refrain from standing out, facilitating the impunity of those armed groups that target them. Therefore, when International Humanitarian Law (IHL) does not provide explicit content that illustrates the problems experienced by LGBT people during wartime, it worsens the already devastating consequences of their relative invisibility.

The fact that people are paying more attention to gender-based crimes against women and girls during armed conflict (United Nations 2017) is due to women's global mobilization and advocacy, which have defied stereotypes of women's weakness, passivity, and unique role as caretakers. Women have become the focus of attention because they let their voice be heard. For the LGBT community, which is equally vulnerable in times of conflict, to gain the same attention, they must become more visible.

This paper first casts an overarching view of the general lack of effort in providing special provisions for LGBT groups in IHL. The paper then proceeds to an overview of the general vulnerability of LGBT people in regard to several forms of discrimination and violence throughout different stages of conflicts, from pre-conflict to post-conflict. An examination of the factors that lead to such vulnerability emphasizes the necessity for institutions and policy makers to fill the existing gaps that hinder the protection of LGBT groups. Lastly, the paper proposes some of the potential ways that would better address, if not mitigate, the vulnerability of LGBT

groups during and after armed conflict due to their invisibility, as well as the roles that UNHCR and human rights doctrine can play in it.

## **Existing Gaps within the International Humanitarian Law Framework**

IHL is especially designed to manage and regulate behavior throughout armed conflicts, both international and domestic. The general purpose of IHL is to draft rules to ensure that the actions in conflicts are monitored and protection is provided for combatants and civilians from violating the norm of distinction. IHL is the first thing that commanders of armed forces should make reference to in regard to how to act when encountering enemy groups and civilians.

Due to the important role that IHL plays during and after armed conflicts, it is necessary for it to be as inclusive as possible so that it can capture as many participants in armed conflicts as possible to provide them with sufficient protection. Even though IHL prohibits discrimination and violence based on “race, color, religion or faith, sex, birth or wealth, or any other similar criteria” under Common Article 3 of the Geneva Conventions, it fails to make an explicit reference to both sexual orientation and gender identity. Part of the reason why it pays little regards to LGBT group is the fact that IHL was drafted and compiled after WWII, when homosexuality was still considered illegal. Even the 1998 Rome Statute, the foundation of the International Criminal Court (ICC), does not explicitly indicate sexual orientation and identities. Suffice to say that LGBT groups will continue to suffer from threats and difficulties until international law specifies that sexual orientation and identity cannot be used to discriminate others during and after conflict. As a matter of fact, the lack of emphasis on the significance of protecting LGBT groups weakens the law’s effectiveness, thus risking creating the illusion to states and combatants that LGBT rights are relatively unimportant compared with those of women and girls, who have begun to gain more attention from the international community. As a result, equal protection across gender groups is not ensured as long as LGBT groups do not receive better consideration.

Through the Refugee Convention, the United Nation High Commissioner of Refugees (UNHCR) has been gradually incorporating sexual orientation and identity into its agenda. This sets an example for IHL to follow, as it informs participants of conflict that it is explicitly unlawful to target LGBT group and to help establish the credibility of the illegal nature of such violence and discrimination when the perpetrators are indicted at international courts. The report on LGBT rights by Human Dignity Trust (2015, 15) indicates that UNHRC proposes the idea that sexual orientation should be considered as part of the agenda:

UNHCR’s programmes usually tend to focus on sexual violence against women, while other forms of psychological, economic, or socio-cultural gender-based violence are less commonly and comprehensively addressed. In particular, the absence of an appropriate guiding policy on how to address and respond to SGBV [sexual and gender-based violence] against lesbian, gay, bisexual and transgendered people of concern (LGBT) remains a serious problem. In addition, the sexual abuse of boys and men is often neglected, under-reported and hardly addressed by any of UNHCR’s programmes.

In 2012, the UNHCR adopted specific guidelines in commitment to the protection of the LGBT groups, which were applied to the more recent refugee emergency in Latin America: “These Guidelines provide substantive and procedural guidance on the determination of refugee status of individuals on the basis of their sexual orientation and/or gender identity, with a view to ensuring a proper and harmonized interpretation of the refugee definition in the 1951 Convention” (ibid, 16, UNHCR 2012, 2).

The extensive effort that the UNHCR is putting forth to protect LGBT people is gradually mainstreaming the LGBT groups in political protection; the IHL framework should follow a similar pattern. In other words, the crucial components regarding the protection of LGBT are present in IHL, but they need to be compiled and vocalized by the members in governments, the United Nations, and other international organizations, thus prohibiting the violence on the basis of sexual orientation and identity. Hence, the extent to which LGBT groups are marginalized and discriminated against needs to be considered by the international community, making it a necessity to highlight the particular vulnerability that LGBT people face in armed conflicts.

## **Forms of Vulnerability that LGBT People Experience in Conflicts**

### **A Continuum of Violence**

The failure to provide to the security needs of LGBT people often leaves them more vulnerable than others, making them targets of violence, especially sexual violence. For example, in Peru, the Truth and Reconciliation Commission Report (2003) suggests that members of the Communist Party of Peru and guerrilla groups committed murders of gay men and rape of lesbian women, and at the same time, publicly promoted homophobic ideas. What is more important is the continuation of violence, which does not begin with the start of the conflict nor end with ceasefire. This continuum proceeds at the micro scale of individuals and the macro level of history. For individuals, hiding their identities or escaping from current war zones does not eliminate the exposure to discrimination and violence.

In societies that are healing from the trauma of armed conflict, discrimination and violence continue, as Bargent (2014, 2) suggests, showing how the entrenched gang violence in Honduras has clearly become a more dominant force in discriminating against LGBT people from armed groups in the civil war. As for refugees who belong to the LGBT group, exposure to violence persists across various phases and different settings of their flight. According to Littauer (2007, 3), alleged gay or lesbian people who have been persecuted in pre-war Syria, continue to be targeted in both war and after. In peacetime, raids on private meeting places were conducted by local police enforcement. People who were arrested, mostly LGBT people, allegedly faced prison sentences and brutal treatments such as police beating. We learn from Krause (2015, 12-15) that refugees from the Syrian conflict encounter violence in rural areas, where they escaped from, and in larger cities in Lebanon, where they hoped to find a society that is hospitable for them.

On the continuum of violence at a macro scale of history, several particular forms of violence stand out, in which the LGBT groups bear worse outcomes than other groups of participants in armed conflicts. Targeted violence is often times utilized by armed forces against LGBT people throughout the course of history. One of the largest scales of such systematic targeted violence was conducted by Nazi Germany, resulting in thousands of deaths of LGBT people (Jellonnek 1990). When it comes to more contemporary targeted violence, Iraq and Syria would be the most visible and prominent ones. During conflicts in Iraq, those Islamic States executed men who are homosexual and threw them off buildings (OutRight Action International 2014, 3-4). They targeted people who are not self-identifying as LGBT. In the same conflicts, according to Human Rights Watch (HRW 2009, 12-14), men and women, who are accused of being either too feminine or masculine, are suspected to be gay or lesbian on the basis of their behavior, clothing, and hair style. Violence and death threats are then exerted on them by armed forces such as the Mahdi Army and the League of the Righteous. Refugees from Syria (Allouche 2016, 1-16) point out that they are forced to purchase protection from local security forces with fees and sexual services, even though they do not belong to LGBT groups. Such exploitation of LGBT groups by less organized armed forces, combined with the systematic violence and perpetration, sometimes result in more lethal consequences. This can be seen from the previous situations in Iraq and Syria.

### **Targeted Violence and Detention**

One of the contexts in which LGBT people are extremely vulnerable is detention. The Office of the High Commissioner for Human Rights (OHCHR 2014, 2) concluded that the reason why LGBT people would be detained in the first place is that they are persecuted fairly arbitrarily on the basis of on their sexual orientation and identity. While in detention, detainees are confined in a relatively limited space deprived of liberty, making them more prone to ill-treatments and violence. In fact, sexual abuse towards LGBT groups captured by state and non-state forces are not uncommon, and such violence may be initiated by both wardens of the detention facilities and fellow victims who share the same space with LGBT detainees (Margalit 2018, 255). Efforts have been made to lower the risk of assaults by holding LGBT detainees separately from other detainees or wardens who show hostility towards them. This practice has been adopted by the Israeli Prison Service (2018), which released its policy regarding transgender prisoners in its system: these prisoners are separated from others in the first five days of their prison life, which minimizes the risk of violence and assault because professionals examine the situation of each prisoner to personalize their arrangement by placing them in either male or female cells.

As sound as it may seem, such management cannot be always achieved, because resources are scarce and the assessment process regarding personal history and well-being is too complicated to truly take care of every detainee, especially during armed conflicts. In other words, the personal safety of LGBT would become the responsibility of authorities in prison and detention facilities. However, the condition of the detention staff varies, especially with regard to proper training and instructions for specifically interacting with LGBT detainees and recognizing their special needs and protection against potential risks. Thus, in many areas in conflict, prisons fail to offer the desperately needed aid to LGBT detainees. For example, many detention facilities do not let LGBT detainees choose the sex of the guards who are responsible for searching them, nor can the search itself be guaranteed to be

respectful of the privacy and rights of the detainees. In addition, the accommodation of the detainees, according to the “Mandela Rules” (2015), tends to be divided into only male and female, regardless of other conditions such as cultural backgrounds and language of detainees. Such neglected elements may result in the lack of social support and minor frictions and tension among detainees, though the situation might escalate and culminate into large scale riots. Even though international organizations such as the International Committee of the Red Cross (ICRC) can intervene in detention facilities directly, they are allowed only to provide external aid to the detainees. By strengthening the rules and revising existing policies, these international forces should be capable of ensuring treatments that are humane and respectful to all detainees, especially those belonging to LGBT groups.

### **Imposing Conservative Gender Norms**

The issue of gender norms, in addition to targeted violence, can also arise in armed conflicts. Not only religious conservatives, but also an array of political groups and nationalists use homophobic violence and discrimination as a tactic to humiliate LGBT people and demonstrate their superiority and dominance over the victims since heterosexual men are often associated with power and strength. This violence is often conducted in the form of social cleansing, which is to force everyone to behave according to heterosexual norms, which are considered as masculine and make the societal group look powerful (HRW 2020, 15-17). The motivation for imposing gender norms and practicing social cleansing of LGBT people may also stem from the need of armed forces to resonate with and gain support and recognition from local residents, the majority of who regard LGBT groups as an anomaly that deserve such punishment and violence. Such social cleansing happens in South Africa, as Strudwick (2014, 1) indicates, where lesbian women are subject to “corrective rape” during occasional armed conflicts because they are forced to “cure” their wrong sexual orientation. The women who are especially vulnerable to corrective rape are those who belong to lower socioeconomic backgrounds. Research conducted by First Aid (2009, 8) reveals that “it is black lesbian from township—who lack sufficient support systems and are already disadvantaged by cultural, economic, and social discrimination—who are particularly at risk. While 44% of white lesbians from the Western Cape lived in fear of sexual assault, 86% of their black counterparts felt the same.”

Similar patterns occurred in Columbia civil war. A report from the National Center of Historical Memory (2015, 26-29) shows that local armed forces would eradicate any presence that bothered them or choose to correct the behavior such as same sex relationship that seems to go against mainstream norms. The Revolutionary Armed Forces of Colombia (FARC) were involved in violence against LGBT people by exercising conservative gender norms, including unbearable assaults on lesbian women such as sexual slavery and rape (Gill 2009, 313-25). As for gay men, members of FARC compel them to box for entertainment. In order to protect LGBT people, local doctors have to let LGBT patients pretend to be heterosexual in medical records to avoid gender norm-based violence from the revolutionary forces (Verdad Abierta 2014, 1). The severity of the situation, however, has led to some attention from the government authority. The Colombian Special Jurisdiction for Peace, a judicial component in Columbia, has started to investigate violent cases that are gender and sexual-based. Its effort in bringing justice mainly concentrates on preventing the



potential exerted norms by capturing the context of the violence instead of prioritizing perpetrators' individual motivations (Valiñas 2020, 1).

Dominant gender norms can also hinder states from implementing sufficient help and protection for LGBT communities. Thus, as Law (1988, 187-194) points out, many states are rather reluctant in providing LGBT people with protection, which is often time strongly opposed by local political forces. Despite the vacuum in necessary support, some states even blocked UN human rights experts who are specialized in studying gender and conflicts. The rejection of collaboration thus reflects the extent to which heterosexual norms are embedded in the culture of the states and the damages the norms cause to the LGBT communities within these states.

## **Factors Leading to the Particular Vulnerability of the LGBT Community**

### **Social Marginalization and Exclusion**

Violence and discrimination against LGBT groups may stem from deeply ingrained mainstream culture and misconceptions about gender. Superstitions and rumors may spread, claiming that homosexuality is a contagious disease, as in modern day Armenia after the collapse of the Soviet Union (Public Information and Knowledge NGO Armenia 2016, 11). This increases the marginalization of LGBT people since people have an instinct to exclude and stay away, instead of actively helping, from other people when they see the others as potential threats to their interests or health.

In religious-related conflicts such as in Northern Ireland, lesbian women are considered as traitors and threats to the society because they fail to contribute to the national cause by producing offspring, which is an essential duty in a patriarchal society (Duggan 2012, 43-53). In times of conflict, many armed groups portray themselves as inherently masculine and heterosexual and their enemy forces as subordinate and homosexual, which is essentially an insult to LGBT groups living under the control of these armed forces (Sremac & Ganzevoort 2015, 158).

Homophobic ideas are magnified after conflict when people are desperate to restore pre-conflict rules and orders, hoping to create better life conditions and community relations. Since established rules are based on heterosexual conventions, the existence of LGBT people undermines and challenges the authority of social order (Myrntinen, Naujoks & El-Bushra 2014). For instance, according to Igrutinović, Sremac and van den Berg (2015, 204-25), in post-war Bosnia, violence against LGBT groups are conducted by nearly all social and religious groups, including Muslim, Serb, Croat, and Wahhabis. Ironically, their unanimous resentment towards LGBT groups reinforces their nationalism sentiments, thus creating a sense of unity, which is essential in the process of re-establishment of the post-war society.

Exclusion by family members and communities can also result in the marginalization of LGBT groups in the context of conflicts. In patriarchal places such as Syria, many LGBT people receive insults and death threats from their male family members because they think a homosexual would disgrace and stigmatize the whole family (HRW 2020, 27-29). In Guatemala, Honduras, and other South American countries where conflicts among heavily-armed gangster groups are prevalent and LGBT people are subject to persecution and exploitation, across different age groups, LGBT children tend to suffer the most from exclusion since they are pushed out of

their family in their early stages of life, losing care and protection from their family and the opportunity for education. As a result, with their childhood trauma and relatively low social status, they often resort to committing thefts and robberies for subsistence (UNHCR 2015, 15-19).

In peacetimes, as teenagers discover their true sexual orientations as homosexual and proceed to adulthood, pressure from families to have heterosexual marriages put LGBT people in a situation where they are prone to violence from their partners. In forced marriages, LGBT people's heterosexual partners compel them to act in a way that conforms to their social obligations, including sexual activities (Myrntinen, Naujoks & El-Bushra 2014, 14). Domestic violence against LGBT increases after conflict. According to Myrntinen and Daigle's research on lesbian groups in Nepal after the civil war, one of their surveys suggests that the proportion of women experiencing intimate partner violence is an astonishing 70%, further corroborating their vulnerability to abuse from families. As for people who do not succumb to forced marriage, 20% of lesbian women are not allowed to participate in community gatherings, and 25% are forced to be displaced due to discrimination by community members (2017, 28).

Thus, with the shame and stigma pinned on them, LGBT people are more reluctant to talk about their experiences in war and postwar, or to seek help voluntarily, leading to limited statistics and therefore inadequate corresponding attention. This vicious circle results in further marginalization of LGBT groups (HRW 2020, 43). Therefore, it is important to inform organizations and policy makers about the exclusion of LGBT people and the difficulties this exclusion brings to them.

### **Lack of Social and Economic Resources**

In peacetime, sexual rights are a major concern for LGBT groups because they can have a profound influence on almost every aspect of social life, including housing, employment, and education. The lack of these rights places them in disadvantaged situations. LGBT people are prone to harassment even in seemingly peaceful settings that are shielded from outside chaos such as schools. Based on a 2014 survey in Iranian secondary schools, around 38.5% of the 834 students surveyed experience various forms of bullying. For LGBT students, the case is worse since school faculties and employees are surrounded by negative discussions of sexual orientations, thus failing to prevent and interrupt the bullying from happening (OutRight Action International 2016, 32-34). LGBT people do not get equal job opportunities either, even in United States, one of the most gender-inclusive countries in the world. Surveys by the Williams Institute (2007, 6) on LGBT employees suggest that up to "17 percent reported that they were fired or denied employment based on their sexual orientation, and 19 percent reported receiving unequal pay or benefits."

When it comes to situations of armed conflict, the circumstances that LGBT people face are much harsher than those in peacetime because their hardships are no longer limited to the aspects described above. In other words, LGBT people face issues that are crucial to their survival. For instance, in regard to physical health, according to Nyanzi (2013, 455-460), health workers in Uganda during the conflict with Congo refused to offer treatment to LGBT victims of war due to their repulsion of homosexuality. In Egypt, where conservative Islamic culture is prevalent and enhanced after the turn of the century, patients with HIV are regarded as indecent and overindulgent, and HIV is considered proof of homosexuality, resulting in widespread discrimination against LGBT people (HRW 2008, 1). In that context, the

stigma of being promiscuous hampers their search for physical and psychological help when they are sick or experiencing traumatic events. Their degradation of health, on the other hand, is often not noticed, which is why humanitarian organizations should specifically arrange human resources to focus on the largely neglected health of LGBT groups (HRW 2018, 4-6).

As for economic independence, LGBT tend to struggle in war-torn areas where recessions continue for years, if not decades. In Colombia, transgender women have few job opportunities provided to them because of their low education and discriminated identities, and the few jobs available are usually low-paid and laborious. A similar situation occurs in Syria, where lesbian women and gay men are compelled to become sex workers for financial reasons, thus putting themselves in a more vulnerable position because it is difficult for them to obtain safety measures when having sexual activities. Hence, they are more likely to be infected with sexually transmitted diseases such as AIDS and syphilis (Allouche 2016, 5-15). Moreover, Syrians who escaped from sexual violence during the war report having limited support and facilities from organizations in Lebanon because of inadequate funding and lack of professional training of the staff to understand the special needs of LGBT people. Consequently, pertinent skills of accommodating resources for LGBT people is much needed (HRW 2020, 18).

## **Potential Ways to Mitigate the Vulnerability of LGBT Groups in Conflict**

### **Establishing Asylum and Providing Social Services for LGBT Groups**

When it comes to the protection of LGBT refugees, it is necessary to ensure them a safe space. This is not easy, however. As discussed in previous sections regarding the continuum of violence, when facing persecution based on sexual orientation in one place, LGBT people are forced to flee abroad in seek of safety. But the truth is, violence against them follows them along through various of stages of their escape. Thus, international cooperation is required, which is in turn based on measures taken at the domestic level. It is in this direction that the current framework of asylum can be further improved. The OHCHR notes the lack of procedures for LGBT people who seek for asylum:

Even in countries that recognize these grounds for asylum, practices and procedures often fall short of international standards. Review of applications is something arbitrary and inconsistent. Officials may have little knowledge about or sensitivity towards conditions facing LGBT people...Refoulement of asylum-seekers fleeing persecution places them at risk of violence, discrimination, and criminalization. (2011, 13)

Hence, the issue of refoulement of asylum-seekers is where international cooperation plays a role. Abiding to international laws, states are allowed to transfer refugees only when the receiving place is able to provide humane treatment and protection. Also, states are obliged to not hand people over to other states where imprisonment, death threats, and ill-treatment may be imposed on them (UNHCR 2012, para 42-46). Domestic policies have also been established to respond to the

vulnerable position of LGBT refugees. Many countries have opened their borders to accept these refugees. For example, the UK has claimed to incorporate Syrian LGBT people into its Syria resettlement program (Wheaton 2015, 1). Even though there are still many challenges when attempting to offer asylum seekers a fairer treatment and abiding to local practices at the same time, taking account of the previous persecution experience might create a measurable standard in treating asylum seekers (Dustin 2018, 1). Also, the scope of acceptance of asylum seekers is continually expanding. It is estimated that more than forty countries have similar policies to provide asylum for LGBT refugees that are targeted based on sexual orientation and identity.

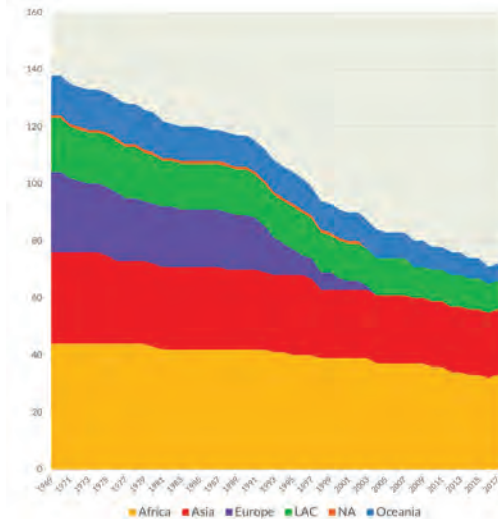
In addition to a safe space, corresponding social services are needed to help LGBT refugees integrate into war-free societies and start their new lives. Many of the services are provided by non-profit organizations. The Jewish Family & Community Services Easy Bay (JFCS East Bay) is one example. It helps resettlement for LGBT refugees in United States through providing affordable housing and offering mental health programs to encourage them to embrace the new environment that has less prejudice and discrimination on their sexual orientation. Other than that, considering the deprivation of education of many of the LGBT people due to their marginalization by the society, JFCS East Bay also enrolls them in classes that teach life skills and English so that they can find employment and blend into society (JFCS East Bay 2015, 1). Similarly, the Los Angeles LGBT Center (2020) is also committed to create a “vital social safety net for the LGBT community.” Its services tend to be at the local level, including assistance with immigration processes, participation in local LGBT cultural activities, and support on educational resources.

### **State-Building Against Homophobia Sentiments**

One of the reasons why the protection of LGBT groups often fall short is the discriminative nature of some domestic laws, hindering them from exercising their human rights. When a conflict comes to an end, it presents an opportunity for the new state to incorporate protective measures for LGBT people, as well as policies that inhibit homophobia, in its state-building plans. During implementation, however, these plans face a challenge: the majority of locals might not champion the new policies that erase homophobic laws. Even though the law of belligerent occupation, which manages the occupying power, states that for the good of the states and local population, necessary changes in laws need to be made to facilitate the economic and ideological development of post-conflict societies, local residents contend that abolishing anti-LGBT policies does not reflect their interests when the majority holds a hostile view to LGBT groups (HRW 2018, 40). A common perception is that equal rights for both heterosexuals and homosexuals promote conflict, for they deviate from the mainstream religious conventions and ideas (Buchanan 2015, 1). Therefore, changes in policies could lead to tensions, even violence, between locals and state government, eradicating the peace that came at a hard cost.

These potential oppositions do not outweigh the principles that international humanitarian law establishes, which is to offer humane treatment to all people, including LGBT groups. Several precedents have proved the viability of the changes in legislation. In India, after 17 years of protests, campaigns, and relentless struggles, the government repealed the notorious Section 377 of the Indian Penal Codes that “violated the rights to equality, non-discrimination, privacy, and dignity,” and abolished the 1860 law that criminalized LGBT groups as well (Narrain 2018, 2-3). Furthermore, in Angola, which used to be a Portuguese colony, the government voted

to rescind the original penal codes that discriminated homosexuality. In response to this policy, an UN official encourages other countries to follow Angola's steps: "All other countries that still criminalize homosexuality, must observe these processes of decriminalization as motivation to examine their own legal frameworks, and to bring themselves to full compliance with this human rights imperative" (UN 2019, 2).



**Figure 1.** UN Member Countries that Criminalize Homosexuality (Yang 2019).

Providing a global and historical perspective, Figure 1 demonstrates the change in the number of UN member states that criminalize LGBT people. The number began at 138 in 1969, and is almost halved by the end of 2017. The average number of countries that decriminalize LGBT each year is around 1.4, with a peak of 5 countries in 1998. In the span of the time since 1969, Europe and North America fully abandoned LGBT discriminative laws, while Africa and Asia have the most countries that criminalized homosexuality, with 44 and 32 countries respectively. Noticeably, prominent changes occurred in Latin America and Oceania, with the number of countries discriminating against LGBT groups decreasing by more than 50%. As seen from the chart, the overall trend leans towards the decriminalization of LGBT groups.

For those countries that still criminalize homosexuality, the majority of which are in Africa, the role of the international communities can be significant. UNHCR, for example, acts as a prime protector of the LGBT victims when the protection from state authority is largely absent. Over a certain time period, however, the government of those countries, such as Kenya, with the presence of UNHCR, has been more involved in regulating the refugee system than those without such as Uganda. This also in turn increases UNHCR's protection capability in the region (Pincock 2020, 7-8).

Therefore, with such a global trend, even though local culture and perceptions may hamper the progress of decriminalizing LGBT people on a national level, governments and local authorities can at least attempt to revise or limit the enforcement of the existing discriminative laws. Through this manner, homophobia

sentiment is able to be first contained, then gradually erased as the process of legislative changes and state-building continues.

## Conclusion

LGBT groups face intensified violence and discrimination on the basis of their sexual orientation during and after armed conflicts, making them one of the most vulnerable minority groups. Recognizing their particular vulnerability, many armed forces exacerbate their initial targeted violence to LGBT groups to avoid rightful accountability. LGBT vulnerability is compounded when international laws do not provide special protection for them since the consequences of homophobia and the hardship LGBT people encounter are underrated and thus invisible to the international community. As a result, it is necessary to appeal to governments, organizations, and legislators and reveal the true conditions that LGBT people face during armed conflicts, including targeted violence against them and the constraint set by conservative gender norms. Furthermore, numerous factors that result in their vulnerability should also be emphasized so that the issue can be mitigated from the source. The exclusion of LGBT groups from mainstream society and their lack of economic and social resource make them invisible in a society that pays little regards to minority groups.

After becoming cognizant of the situation of LGBT groups, states and institutions should be reminded of their capability in providing them with protection and services. For the time being, states can establish asylums in order to create a risk-free space that block potential violence and discrimination. It is the role of organizations to provide social services such as life skills and language training to help LGBT refugees adapt to a new society. In the long term, states should then take action to gradually reverse homophobic ideology by both repealing discriminative laws and promoting equal rights for everyone, including LGBT groups. In order to further accelerate the process, activists that promote human and LGBT rights can spread those rights internationally to motivate domestic citizens, who in turn prompt the government to act accordingly due to the concern over reputation (Keck & Sikkink 1998, 16-25). Such a persuasion mechanism should be considered when incorporating international cooperation into the cause of emphasizing the rights of LGBT people. The true effectiveness of it would even be greater depending on social contexts and in combinations with other methods.

## References

- Action Aid (2009). "Hate Crimes: The Rise of 'Corrective 'Rape in South Africa". ActionAvalablattat:[https://www.actionaid.org.uk/sites/default/files/publications/hatecrimes\\_the\\_rise\\_of\\_corrective\\_rape\\_in\\_south\\_africa\\_september\\_2009.pdf](https://www.actionaid.org.uk/sites/default/files/publications/hatecrimes_the_rise_of_corrective_rape_in_south_africa_september_2009.pdf)
- Allouche, Sabiha (2018). "IR and the Instrumentalization of Sexuality in the Reproduction of Political Leadership in Lebanon: On being Queer Syrian and Refugee in Contemporary Lebanon," *International Studies Association (ISA) Annual Convention, Atlanta*.

- Bargent, James (2014). "Police, Gangs behind Abuse of Honduras' LGBTI Community," *InSight Crime*. August 22, 2014. Available at: <http://www.insightcrime.org/news-briefs/police-gangs-abuse-honduras-lgbti-community>
- Buchanan, Rose (2015). "Robert Mugabe Tells UN General Assembly: 'We Are Not Gays!'" *The Independent*. September 29, 2015. Available at: <https://www.independent.co.uk/news/people/robert-mugabe-tells-un-general-assembly-we-are-not-gays-a6671316.html>
- Dinstein, Yoram (2009). *The International Law of Belligerent Occupation*. Cambridge University Press.
- Duggan, Marian (2012). *Queering Conflict: Examining Lesbian and Gay Experiences of Homophobia in Northern Ireland*. Farnham: Ashgate
- Dustin, Moira (2018). "Many rivers to cross: The recognition of LGBTQI asylum in the UK." *International Journal of Refugee Law* 30.1 : 104-127.
- Gill, Lesley (2009). "The Parastate in Colombia: Political Violence and the Restructuring of Barrancabermeja", *Canadian Anthropology Society*, 51(2): 313-25.
- Haeyoun, Park and Iaryna Mykhyalyshyn (2016). "LGBT People Are More Likely to Be Targets of Hate Crimes than Any Other Minority Group," *The New York Times*, June 16, 2016. Available at: [www.nytimes.com/interactive/2016/06/16/us/hate-crimes-against-lgbt.html](http://www.nytimes.com/interactive/2016/06/16/us/hate-crimes-against-lgbt.html).
- Human Dignity Trust (2015). "Criminalising Homosexuality and LGBT Rights in Times of Conflict, Violence and Natural Disasters". London: Human Dignity Trust.
- Human Rights Council (HRC) (2014). "Report of the Independent International Commission of Inquiry on the Syrian Arab Republic". UN Doc. A/HRC/25/65.
- Human Rights Watch (2008). "Egypt: Stop Criminalizing HIV", *New York: HRW*. Available at: [www.hrw.org/news/2008/02/04/egypt-stop-criminalizing-hiv](http://www.hrw.org/news/2008/02/04/egypt-stop-criminalizing-hiv)
- Human Rights Watch (2009). "They Want us Exterminated': Murder, Torture, Sexual Orientation and Gender in Iraq." New York: HRW. Available at: <https://www.hrw.org/sites/default/files/reports/iraq0809web.pdf>
- Human Rights Watch (2018). "More Arrests in Egypt's LGBT Crackdown, but No International Outcry," *New York: HRW*.
- Human Rights Watch (2018). "Scared in Public and Now No Privacy: Human Rights and Public Health Impacts of Indonesia's Anti- LGBT Moral Panic". HRW, July 2018. Available at: [www.hrw.org/sites/default/files/accessible\\_document/indonesia0618\\_lgbt\\_web.pdf](http://www.hrw.org/sites/default/files/accessible_document/indonesia0618_lgbt_web.pdf)
- Human Rights Watch (2020). "'They Treated Us in a Monstrous Way.' Sexual Violence Against Men, Boys, and Transgender Women in Syrian Conflict." New York: HRW. Available at: <https://www.hrw.org/report/2020/07/29/they-treated-us-monstrous-ways/sexual-violence-against-men-boys-and-transgender>
- Igrutinović, Danica, Srdjan Sremac, and Mariecke Van Den Berg (2015). "Pride Parades And/or Prayer Processions: Contested Public Space in Serbia Belgrade Pride 2014," *Journal of Empirical Theology* 28 (2): 204-225
- Israel Prison Service (2018). "Admission of Transgender Detainees and Guidelines for Their Admission: Policy." Available at: <https://tinyurl.com/y66ktxl>
- Jellonnek, Burkhard (1990). *Homosexuals under the Hakenkreuz: The Persecution of Homosexuals during the Third Reich*. Verlag Ferdinand Schöningh.
- Jewish Family & Community Services East Bay (2015). *LGBT Refugee Services*. Available at: <https://jfcs-eastbay.org/lgbt-refugee-services/>

- Keck, Margaret and Kathryn Sikkink (1998). *Activists Beyond Borders: Advocacy Networks in International Politics*. Cornell University Press.
- Krause, Ulrike (2015). "A Continuum of Violence? Linking Sexual and Gender-based Violence during Conflict, Flight, and Encampment", *Refugee Survey Quarterly*, 34 (4): 1-19.
- Law, Sylvia (1988). "Homosexuality and the Social Meaning of Gender," *Wisconsin Law Review*, Vol. 1988, (2): 187-202.
- Littauer, Dan (2010). "Syria's Attack on Gay People must End," *The Guardian*, July 7, 2010. Available at: <http://www.theguardian.com/commentisfree/2010/jul/07/homosexual-syria-persecution-lgbt-rights>
- Los Angeles LGBT Center (2020). "Culture & Education". Available at: <https://lalgbtcenter.org/culture-education>
- Margalit, Alon (2018). "Still a Blind Spot: The Protection of LGBT Persons During Armed Conflict and Other Situations of Violence," *International Review of the Red Cross*, 100 (1-2-3): 237-65.
- Myrntinen, Henri, Jana Naujoks and Judy El-Bushra (2014). "Re-thinking Gender in Peacebuilding". London: International Alert.
- Myrntinen, Henri, and Megan Daigle (2017). *When merely Existing is a Risk: Sexual and Gender Minorities in Conflict, Displacement and Peacebuilding*. International Alert.
- Narrain, Arvind (2018). "Right to Love: Navtej Singh Johar v. Union of India: A Transformative Constitution and the Rights of LGBT Persons". Bangalore: National Printing Press.
- National Center of Historical Memory (2015). "Annihilating the Difference: Lesbians, Gays, Bisexuals and Transgender People in the Context of the Colombian Armed Conflict." Bogotá: CNMH
- Nyanzi, Stella. 2013. "Homosexuality, Sex Work, and HIV/AIDS in Displacement and Post-Conflict Settings: The Case of Refugees in Uganda," *International Peacekeeping*, 20(4): 450-468.
- Office of the United Nations High Commissioner for Human Rights (OHCHR) (2014). "The Gambia: Zeid Criticizes Harsh Legal Amendment, Violence and Arrests Targeting Gay Men and Lesbians". OHCHR.
- OHCHR (2011). "Discriminatory Laws and Practices and Acts of Violence against Individuals based on their Sexual Orientation and Gender Identity". OHCHR.
- OutRight Action International (2014). "When Coming Out is a Death Sentence: Persecution of LGBT Iraqis". New York: OutRight Action International.
- OutRight Action International (2016). "Human Rights Report: Being Lesbian in Iran". New York: OutRight Action International
- Pincock, Kate (2020). "UNHCR and LGBTI refugees in Kenya: The limits of 'protection.'" *Disasters*. <https://doi.org/10.1111/disa.12447>
- Public Information and of Knowledge NGO (PINK) Armenia (2016). "From Prejudice to Equality: Study of Societal Attitudes towards LGBTI People in Armenia". Yerevan: PINK Armenia, June, 2016. Available at: <http://www.pinkarmenia.org/wp-content/uploads/2016/06/From-Prejudice-to-Equality-English.pdf>
- Sremac, Srdjan and Ruud Ganzevoort (2015). *Religious and Sexual Nationalisms in Central and Eastern Europe: Gods, Gays and Governments*. Brill.



- Strudwick, Patrick (2014). "Crisis in South Africa: The Shocking Practice of Corrective Rape Aimed at Curing Lesbians." *The Independent*, January 4, 2014. Available at: <http://www.independent.co.uk/news/world/africa/crisis-in-south-africa-the-shocking-practice-of-corrective-rape-aimed-at-curing-lesbians-9033224.html>
- The Williams Institute (2014). "Gender Identity and Sexual Orientation Discrimination in the Workplace." The Williams Institute. Available at: <https://escholarship.org/uc/item/9qson354>
- The Williams Institute (2014). "Surveying Nepal's Sexual and Gender Minorities: An Inclusive Approach". Bangkok: UNDP. Available at: <http://williamsinstitute.law.ucla.edu/wp-content/uploads/Nepal-Survey-Oct-2014.pdf>
- Truth and Reconciliation Commission (2003). "Final Report". Lima: CVR. Available at: <https://www.cverdad.org.pe/ingles/ifinal/>
- UNHCR (2012). "Guidelines on International Protection No. 9: Claims to Refugee Status based on Sexual Orientation and/or Gender Identity within the Context of Article 1A(2) of the 1951 Convention and/or its 1967 Protocol relating to the Status of Refugees". UNHCR. Available at: <http://www.unhcr.org/50ae466f9.html>
- UNHCR (2012). "Sexual Orientation and Gender Identity Guidelines". UN Doc HCR/GIP/12/09 October 13, 2012. Available at <http://www.unhcr.org/refworld/docid/50348afc2.html>
- UNHCR (2015). "Women on the Run: First-hand Accounts of Refugees Fleeing El Salvador", *Guatemala, Honduras, and Mexico*. Geneva: UNHCR.
- United Nations (UN) (2015). "UN Standard Minimum Rules for the Treatment of Prisoners (Mandela Rules)", *UNGA Res. 70/175, 2015*.
- United Nations (UN) (2017). "Report of the Secretary-General on Conflict-Related Sexual Violence". Geneva: United Nations. Available at: <https://www.un.org/sexualviolenceinconflict/wpcontent/uploads/2019/04/reports-2019-280/Annual-report-2018.pdf>.
- United Nations (2019). "UN welcomes Angola's Repeal of Anti-Gay Law, and Ban on Discrimination based on Sexual Orientation," UN, January, 2019. Available at: <https://news.un.org/en/story/2019/01/1031292>
- United Nations Security Council (2015). "Conflict-related Sexual Violence: Report of the Secretary-General". S/2015/203.
- Valiñas, Marta (2020). "The Colombian Special Jurisdiction for Peace: A Few Issues for Consideration When Investigating and Adjudicating Sexual and Gender-based Crimes." *Journal of International Criminal Justice* 18.2 (2020): 449-467.
- Verdad Abierta (2014). "The 'Paramilitaries' Killed Us for Being Homosexual," November 25, 2014. Available at: <http://www.verdadabierta.com/victimassccion/asesinatos-colectivos/5524-los-paras-nos-mataban-por-ser-homosexuales>
- Ware, Wesley (2011). "Locked Up and Out: Lesbian, Gay, Bisexual and Transgender Youth in Louisiana's Juvenile Justice System". New Orleans: Juvenile Justice Project of Louisiana.
- Wheaton, Oliver (2015). "David Cameron: We will Resettle LGBT Refugees from Syria and Iraq." *The Metro*. December 9, 2015. Available at: <http://metro.co.uk/2015/12/09/david-cameron-we-will-resettle-lgbt-refugees-from-syria-and-iraq-5555194/#ixzz3u1qvVmyF>

Yang, Daryl. 2019. "Global Trend on the Decriminalisation of Consensual Same-Sex Sexual Acts (1969-2019)". ILGA World: State-Sponsored Homophobia 13th edition.



# The Relationship between the Anti-Inflammatory Strength of Butyrate and the Progression of Colorectal Cancer

---

*Rohan Kartha*

**Author Background:** *Rohan Kartha grew up in the United States and currently attends Westwood High School, in Austin, Texas, USA. His Pioneer research concentration was in the field of biology and titled "The Gut Microbiome."*

## Abstract

Colorectal cancer (CRC) is a type of cancer that begins in the large intestine or rectum. CRC has been linked to many risk factors, including inflammatory bowel disease, family medical history, excessive consumption of processed meat, and high alcohol consumption. Although there is not one singular causal factor for CRC, an association has been established with the dysbiosis of the gut microbiome, a collection of trillions of bacteria, fungi, and other microorganisms colonizing the gastrointestinal tract. Similarly, it has been proposed that CRC-induced inflammation of the epithelium creates a microenvironment unfavorable to the survival of several taxa in the gut microbiome. Butyrate, a product of bacterial fermentation that functions as the primary energy source for colonocytes, has been widely studied for its immunoregulatory and anti-carcinogenic properties. However, recent studies of butyrate in inflammatory bowel disease (IBD) models suggest that the anti-inflammatory effects of butyrate may be overridden by other factors as the disease progresses. In this paper, a study is proposed to test the hypothesis that as CRC progresses, the anti-inflammatory effects of butyrate will be overridden due to the CRC-induced loss of butyrate-producing bacteria from the gut microbiome.

## 1. Introduction

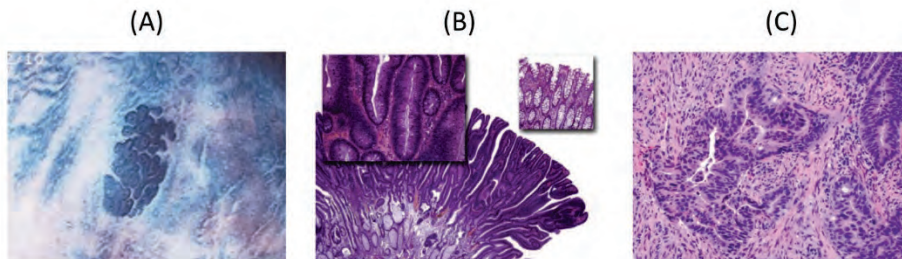
### 1.1. Familial Adenomatous Polyposis

One of the most-studied hereditary forms of CRC is familial adenomatous polyposis, which is inherited through an autosomal dominant mutation in the adenomatous polyposis (APC) gene [1]. The APC gene encodes adenomatous polyposis coli, a protein which plays an important role in the Wnt- $\beta$ -catenin pathway and whose absence results in a nuclear buildup of  $\beta$ -catenin, a transcription factor for genes promoting intestinal epithelial crypt cell proliferation [2]. Later adenoma and

carcinoma development is enhanced by mutations in the tumor suppressor gene p53 and the oncogene KRAS [3]. Additionally, a recessive form of familial adenomatous polyposis, known as MYH-associated polyposis, is inherited through mutations in the MUTYH gene, which encodes MYH glycosylase, an enzyme that corrects base-pairing errors during DNA replication [4].

### 1.2. Histological Structures of Familial Adenomatous Polyposis

There are three main histological structures associated with the progression of familial adenomatous polyposis: aberrant crypt foci, adenomas, and adenocarcinomas [5,6]. Aberrant crypt foci are neoplastic lesions that serve as precursors to adenomatous polyps [5]. The crypts are situated above the intestinal mucosal layer and are composed of unusually thickened epithelial tissue, which is stained well by methylene blue. Thus, aberrant crypt foci can be easily distinguished from normal intestinal crypts. Colorectal adenomas are benign glandular neoplasia in the large intestine [7] of which two of five common types, adenomatous and villous/tubulovillous adenoma, are most likely to turn cancerous [6]. Interestingly, pseudoinvasion of the muscularis mucosae is displayed by some adenomas [8]. Adenocarcinomas, associated with further progression of the disease, are metastatic glandular neoplasia in the colon and its surrounding tissues that penetrate the muscularis mucosae and extend into the submucosa [8, 9].



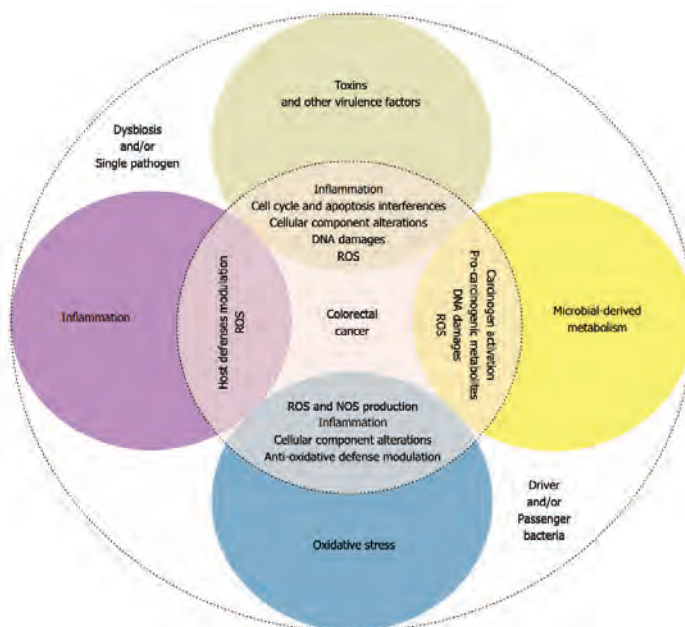
**Figure 1.** The three main histological structures associated with different stages of colorectal cancer: aberrant crypt foci, adenoma, and adenocarcinoma. **(a)** An aberrant crypt focus stained by methylene blue showing dark blue thickened epithelium and loss of cellular organization [10]. **(b)** Villous adenoma exhibiting low-grade dysplasia. The top right image is normal mucosal tissue [11]. **(c)** Adenocarcinoma with moderate differentiation and developed glandular structures [12].

### 1.3. Colorectal Cancer and Gut Microbiome

The gut microbiome is a collection of more than 100 trillion microbes, including bacteria, fungi, and phages, that colonize the gastrointestinal tract and perform a variety of functions essential to human survival [13]. In the colon, which contains the body's largest microbial community, most bacteria are obligate anaerobes and members of phyla *Firmicutes*, *Bacteroidetes*, *Actinobacteria*, *Proteobacteria*, or *Verrucomicrobia* [14]. Recently, a link between dysbiosis of the gut microbiome and susceptibility to CRC has been discovered. One study associated the risk of CRC in humans with a decreased abundance of the class *Clostridia* and an enrichment of the genera *Fusobacterium*, *Porphyromonas*, and *Atopobium* [15]. The risk of CRC has

additionally been linked with increases in intestinal epithelial levels of *S. bovis*, *H. pylori*, *B. fragilis*, *E. faecalis*, *C. septicum*, *F. nucleatum*, and *E. coli* [16]. However, these trends are not well established; for example, in one study, a higher abundance of OTUs in class *Clostridia* was exhibited by human patients with adenomas [17]. Thus, while a general association between the gut microbiome and CRC has been discovered, the exact nature of the relationship and specific taxa involved are not clearly defined.

One mechanism through which the gut microbiome might induce CRC is its immunoregulatory properties. As the microbiome is remodeled, a proinflammatory response is induced, leading to the transformation of the epithelium into an environment more conducive to carcinogenesis. For example, it was shown that in NOD2-deficient mice, the subsequent dysbiosis mediated by the lack of the immune system protein increased the risk of colorectal carcinogenesis [18]. An associated theory is the “driver-passenger hypothesis”, which proposes that bacteria-triggered epithelial cell DNA damage will lead to carcinogenesis, and that the subsequent change in the epithelial environment will promote the growth of bacteria that support tumorigenesis [9]. Other possible mechanisms have been summarized in the figure below.



**Figure 2.** Proposed mechanisms through which the gut microbiome induces onset of CRC. Notable mechanisms in addition to inflammation include bacterial-derived genotoxins and microbial metabolites. For example, some strains of *H. pylori* produce cytotoxin-associated gene A, which increases inflammation and cancer progression [16].

#### 1.4. Butyrate, Gut Microbiome, and Colorectal Cancer

Butyrate is an organic, four-carbon short chain fatty acid (SCFA) produced by the bacterial fermentation of indigestible, resistant starches and dietary fibers [19]. Important butyrate producers in the gut include *B. pullicaecorum* 25-3 T, *F. prausnitzii*, *Roseburia hominis*, *Roseburia inulinivorans*, *Anaerostipes caccae*, and

*Eubacterium hallii* [20]. Additionally, butyrate producers have been found to have lower abundance in patients with CRC. For example, a fourfold decrease in the levels of butyrate producers *E. rectale* and *F. prausnitzii* in the stool of CRC patients was observed [21].

Butyrate, rather than glucose, is primarily oxidized by colonocytes and has been shown to be responsible for up to 86% of the total oxygen consumption in rat colonocyte cultures [22]. Among other roles, butyrate strengthens the intestinal barrier, regulates the growth and differentiation of epithelial cells, regulates colonocyte ion transport, and, perhaps most importantly, plays an anti-inflammatory role in the colonic epithelium and mucosa [19]. Butyrate has been linked to the suppression of proinflammatory factors interleukin-8 (IL-8), interleukin-1b (IL-1B), tumor necrosis factor alpha (TNF $\alpha$ ), and nitric oxide (NO), as well as the increased expression of anti-inflammatory factors interleukin-10 (IL-10) and Prostaglandin E<sub>2</sub> (PGE<sub>2</sub>) [23, 24]. However, the relationship between butyrate and the immune system is bidirectional. A recent study found that the treatment of HCT116 colorectal cancer cells with interleukin-1 $\beta$  significantly decreased the colonocytic oxidation of butyrate and levels of NADH [25]. There are two main proposed mechanisms for how butyrate exerts its immunoregulatory properties: (1) as a histone deacetylase inhibitor (HDACi); (2) as a ligand to G-Protein Coupled Receptors (GPCRs) [23]. Butyrate has been shown to target class I and II histone deacetylases, promoting histone acetylation and increasing the transcription of tumor suppressor genes such as the p21 (Waf1/Cip1) gene [26, 27]. Similarly, butyrate has been shown to act as a ligand for intestinal epithelial cell GPCRs including GPR43, GPR41, and GPR109A, with some receptors being associated with tumor suppressor pathways [27]. Thus, due to its anti-inflammatory properties and CRC's association with an inflamed epithelial state, butyrate has been labeled as an anti-carcinogenic compound, shown to inhibit CRC progression in a variety of contexts and models [28].

However, recent studies have suggested that the anti-carcinogenic and anti-inflammatory roles of butyrate may not be as clear as previously thought. A study conducted in C57BL/6 mice with IBD administered n-butyrate to alleviate inflammation in the intestine [29]. However, mice who both were and were not administered butyrate showed similar weight loss, suggesting that the anti-inflammatory effects of butyrate were overridden by tissue reparative factors caused by the progression of IBD. Similarly, another study conducted with Caco-2 cells found that butyrate only suppressed IL-8 expression after cells had been stimulated with Pam3CSK4; if cells were pretreated with butyrate, IL-8 expression increased [30]. It is clear that the extent to which butyrate plays an anti-inflammatory role is dependent on the biochemical context in which it is produced or administered.

### 1.5. Research Question and Hypothesis

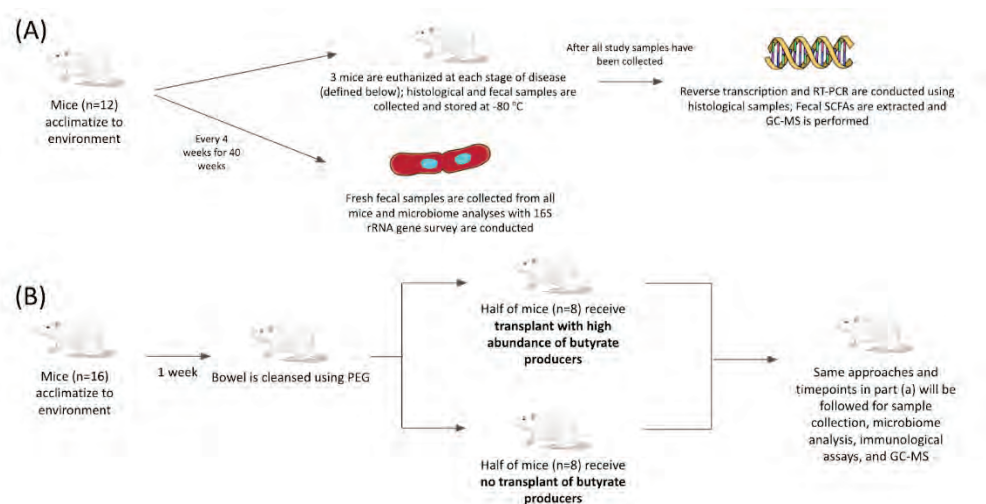
Thus, in the proposed study, the strength of the anti-inflammatory and anti-carcinogenic properties of butyrate, and how they might be augmented or overridden by other factors as CRC and its accompanying effects on the intestinal epithelial environment progress, will be analyzed. It is hypothesized that in a murine model, the magnitude of the anti-inflammatory effects generated by butyrate will differ according to the number of butyrate-producing bacteria in the gut microbiome and the stage of CRC present. Specifically, a significant anti-inflammatory role will be played by butyrate in the early stages of CRC because it is produced in greater abundance, but as the disease progresses, the production of butyrate will decrease,

due to a decrease in the abundance of butyrate-producing bacteria. As a result, the anti-inflammatory effects of butyrate will be mitigated.

## 2. Methods and Materials

### 2.1. General Approaches

Two approaches are proposed to elucidate the anti-inflammatory and anti-carcinogenic properties of butyrate, and how they might be overridden as the biochemical milieu of the colon changes. The first approach, an observational study, will track the progression of CRC in a murine model susceptible to onset of Familial Adenomatous Polyposis and, at different timepoints within the disease, characterize intestinal inflammatory markers, cancer progression, and fecal short-chain fatty acids, in an attempt to draw a correlation between the stage of CRC and the effects of butyrate. The second approach, a manipulative study, will alter the levels of butyrate-producers in a germ-free murine model and track changes in the inflammatory response, tumor progression, and the gut microbiome at different timepoints.



**Figure 3.** A summary of the two experimental approaches. **(a)** The first, observational 40-week study. **(b)** The second, experimental 40-week study.

### 2.2. Animal Model

Fabp1Cre;Apc<sup>15lox/+</sup> mice will model butyrate's anti-inflammatory and anti-tumorigenic properties in a colorectal cancer model. The starting age of the mice will be 4 weeks, at which age Fabp1Cre;Apc<sup>15lox/+</sup> has no intestinal tumor development, but does have significant loss of functional Apc alleles in colonocytes via prior Cre-mediated exon 15 deletion [31]. Mice will be randomly assigned to treatment groups. To prevent contamination and maintain specific pathogen-free conditions, all treatment groups will be housed in flexible film isolators under identical conditions. Additionally, feed, bedding, and drinking water will be autoclaved to ensure sterile

conditions. The approach of Zhai et al. will be followed by maintaining a living temperature of 23°C, relative humidity of 40%-70%, and a 12h/12h light/dark cycle [32]. Water and feed will be applied as needed, and the mice will be maintained on conventional laboratory diets with sufficient quantities of resistant starches and dietary fibers.

### **2.3. Creation of Germ-Free State**

The second experiment necessitates a manipulation of the abundance of butyrate-producing taxa in the gut microbiome. To this end, the approach of Wrzosek et al. will be followed to create a germ-free gut state in the *Fabp1Cre;Apc<sup>15lox/+</sup>* mouse at the beginning of the experiment [33]. The mice will first fast for 1 hour with access to water. Next, the murine gut will be cleansed with four successive rinses of 200 µL and 425g/L polyethylene glycol (PEG) at twenty-minute intervals. PEG will be administered through a sterilized oral-gastric gavage. After bacterial transplantation, gut microbiome composition will be characterized from feces, and the mice will be given food and water ad libitum.

### **2.4. Taxa and Transplantation of Butyrate-Producers**

In the second experiment, the composition and health of butyrate-producers in the murine gut microbiome will be manipulated using germ-free mice. The specific butyrate-producers and culture methods described by Geirnaert et al. will be used [20]. *B. pullicaecorum 25-3 T*, *F. prausnitzii*, *Roseburia hominis*, *Roseburia inulinivorans*, *Anaerostipes caccae*, and *Eubacterium hallii* will be used in this study. The strains will be plated at pH 6 in anaerobic M2GSC medium and incubated for 20-40h at 37°C. The day before experimentation, the bacteria will be inoculated in liquid M2GSC medium and incubated at 37°C. Subsequently, the cultures will be centrifuged to remove supernatant and concentrate butyrate-producers. Next, the cultures will be used to inoculate the germ-free mouse through oral-gastric gavage; bacteria will be added in equal ratios and will only be added once.

### **2.5. Gut Microbiome Characterization**

The approach of Choi et al. will be followed to characterize the gut microbiome through DNA extraction and 16S rRNA gene sequencing [34]. Fecal samples will be collected every four weeks and stored in sterilized conditions at -80°C. The FastDNA SPIN Kit for Soil will be used to extract DNA from the fecal samples, and will amplify the V3 and V4 regions of the 16S rRNA gene through PCR using primers 341F and 805R. The following conditions for amplification will be used: 3 min at 95°C, 25 cycles for 30 sec at 95°C, 30 sec at 55°C, 30 sec at 72°C, and 5 min at 72°C. The amplified DNA will be purified, and its quality will be measured. DNA sequencing will be conducted with an Illumina MiSeq sequencing system. To quantify microbial diversity, the Shannon index, species richness, species dominance, and UniFrac will be measured. The relative abundances of the 6 butyrate-producers will also be measured. All data analysis will be performed using the software QIIME2.

### **2.6. Progression of Colorectal Cancer**

The progression of CRC in the *Fabp1Cre;Apc<sup>15lox/+</sup>* mice will be estimated using external indicators such as weight loss. A certain number of mice will be euthanized at each stage of CRC. After euthanization, there are three histological structures



whose presence and abundances in the samples will be analyzed to determine the true stage of CRC: aberrant crypt foci, tubular and villous adenomas, and adenocarcinomas. Aberrant crypt foci (ACF) will be identified by staining murine histological samples with the dye methylene blue [5]. Stored tissue samples will be rinsed with deionized water, stained with 0.05% methylene blue solution for 3 min, and have excess dye removed with deionized water [35]. The number of crypts per ACF and the number of ACF per cm<sup>2</sup> of tissue will then be analyzed using a light microscope [36]. As delineated by Boivin et al., adenomas will be characterized by their growth pattern, histologic type, and grade of dysplasia, and the adenocarcinomas by grade of differentiation and histologic type [8].

## **2.7. Extracting SCFAs and GC-MS**

The levels of the SCFAs butyrate, acetate, and propionate in murine fecal samples will be gauged every four weeks throughout both experiments. To extract SCFAs and characterize levels in fecal samples, the approach previously described by He et al. will be followed [37]. Fecal samples of 30 mg, previously stored at -80°C, will be diluted in water to make a 300 mL solution. After the sonication and centrifugation of the solution, the supernatant will be derivatized by pentafluorobenzyl bromide and mixed with acetone solution and buffer. After a period of incubation and cooling, 150 µL of hexane will be added to the mixture, which will then be centrifuged, and supernatant will be collected. A Thermo ITQ 1100 GC-Ion Trap MS Instrument, a 1310 autosampler, and Thermo Trace 1310 gas chromatography system will be used for GC-MS analysis. Two GC columns, DB-225 ms and DB-5 ms, will be used with a helium carrier gas flow rate of 1.0 mL/min and will have an initial temperature of 40°C. The temperatures of the inlet, ion source, and transfer line will be 220°C. The following conditions for column temperature will be used: 0.5 min at 40°C, increase to 170°C at a rate of 10°C/min, 0.5 min at 170°C, increase to 220°C at a rate of 5°C/min, 5 min at 220°C. 1 µL of supernatant will be injected into GC-MS in splitless mode, and mass spectrum will be generated in selected ion monitoring mode.

## **2.8. Proinflammatory and Anti-inflammatory Markers**

To accurately gauge the inflammatory effects of butyrate, the levels of proinflammatory markers interleukin-8 (IL-8), interleukin-1b (IL-1B), tumor necrosis factor alpha (TNF $\alpha$ ), and nitric oxide (NO), and anti-inflammatory markers interleukin-10 (IL-10) and Prostaglandin E<sub>2</sub> (PGE<sub>2</sub>) will be measured in histological intestinal samples. To this end, the approach previously demonstrated by Zhai et al. will be followed: every four weeks, mice will be euthanized, and colon tissue samples will be dissected and stored at -80 °C [32]. RNA will be isolated from histological samples using Trizol reagent, and will be reverse-transcribed to cDNA using the ImProm-II Reverse Transcription kit. Next, the RT-PCR will be conducted using the QuantStudio RT-PCR Instrument and SYBER Green Supermix to amplify DNA and measure the levels of expressed genes associated with the synthesis of specific cytokines and eicosanoids.

### 3. Discussion

#### 3.1. Expected Results

In the first experiment of this study, it is predicted that the concentrations of proinflammatory markers will increase and that the concentrations of anti-inflammatory markers, abundance of butyrate producers, and concentrations of fecal SCFAs will decrease with the progression of CRC. This result is expected because a loss in abundance of butyrate-producers has been associated with CRC [21], and butyrate has been linked to the suppression of proinflammatory factors and the expression of anti-inflammatory factors [23,24]. However, several other outcomes could occur. For example, an unchanging gut microbiome could be accompanied by a decrease in fecal SCFA levels, which could likely be attributed to a decrease in the metabolic rate of butyrate and other SCFA producers. Similarly, an increase in SCFA levels and a stable gut microbiome could be linked to increased bacterial metabolic activity. If we are unable to detect changes in either SCFAs or the gut microbiome, it is likely that experimental limitations, such as collecting SCFAs and bacteria from feces rather than directly from the gut, have significantly influenced the data, since the community of the gut lining is known to be different from that of the feces.

In the second experiment, slower progression of CRC, lower concentrations of proinflammatory markers, higher concentrations of anti-inflammatory markers, greater abundance of butyrate-producers, and higher levels of fecal SCFAs in mice who receive the butyrate producer transplant is expected to be observed. This result is expected because butyrate regulates the health of the intestinal epithelium [19] and has been linked to lower levels of proinflammatory factors and higher levels of anti-inflammatory factors [23,24]. Many alternative outcomes and explanations are similar to those of the observational phase. Additionally, it is possible that a similar progression of CRC could be observed in mice with different types of transplants; possible explanations include experimental limitations and the interactions of butyrate producers with remnant bacteria from bowel cleansing.

#### 3.2. Justifications and Limitations

The *FabplCre;Apc<sup>15lox/+</sup>* mouse model holds several advantages over more conventional CRC murine models. In a recent study that compared *FabplCre;Apc<sup>15lox/+</sup>* mice with *Apc<sup>Δ15/+</sup>* mice, *FabplCre;Apc<sup>15lox/+</sup>* displayed a longer lifespan (7.6 months), developed colorectal tumors rather than the small intestinal tumors characteristic of *Apc<sup>Δ15/+</sup>* and other models, and grew low-grade adenomas in 91% of mice, high-grade adenomas in 50% of mice, and adenocarcinomas in 50% of mice [38]. Therefore, *FabplCre;Apc<sup>15lox/+</sup>* mice are suitable models for this study, which spans many stages of colorectal cancer. However, to keep animal numbers low, the stage of cancer present in mice will be judged through external signals like weight loss, which are not exact indicators of cancer progression and may be prone to error. Because CRC will progress at different rates in different mice, one alternative is to euthanize mice at fixed time points; however, this would greatly increase animal numbers.

The butyrate-producers that will be used in this study, *B. pullicaecorum* 25-3 T, *F. prausnitzii*, *Roseburia hominis*, *Roseburia inulinivorans*, *Anaerostipes caccae*, and *Eubacterium hallii*, have been shown to successfully colonize the intestinal lumen, significantly increase the synthesis of butyrate, and improve epithelial barrier function in an *in vitro* inflammatory bowel disease (IBD) model

[20]. However, because successive rinses of PEG have previously been shown to only remove 90% of gut bacteria [33], even though all butyrate producers will be added to PEG-cleansed mice in equal ratios, it is likely that some of the added taxa may not survive in the intestinal environment and will be outcompeted by remnant bacteria or other added butyrate producers. One alternative to PEG bowel cleansing is using germ-free mice, which would avoid any interactions between the butyrate-producers and other microbes.

Quantifying SCFA levels and characterizing the gut microbiome from fecal samples greatly reduces the number of euthanizations and mice needed for this study. Additionally, the use of fecal microbiota as a proxy for the gut microbiome has been shown to be effective; Suzuki and Nachman found that 93.3% of core OTUs in *Mus musculus* were shared between fecal samples and the lower gastrointestinal tract [39]. More accurate SCFA and microbiome measurements would necessitate euthanization and the collection of samples directly from the gut.

## 4. Conclusion

A bidirectional relationship between CRC and the dysbiosis of the gut microbiome has been established by previous literature. In this paper, a study is proposed to elucidate the extent to which the anti-carcinogenic and anti-inflammatory abilities of the SCFA butyrate are overridden as CRC progresses in a *Fabp1Cre;Apc<sup>5lox/+</sup>* murine model. Methods utilized include 16S rRNA gene surveys, gas-chromatography-mass spectrometry, real-time polymerase chain reaction, and histological staining; variables measured include gut microbiome diversity, expression of inflammatory factors, and fecal SCFA content. It is hoped that this study will improve the current understanding of the mechanisms underlying colorectal carcinogenesis and suggest outlets for future research.

## References

1. Half E, Bercovich D, Rozen P. Familial Adenomatous Polyposis. *Orphanet J Rare Dis.* 2009 Oct 12;4(1): 22. doi: 10.1186/1750-1172-4-22 PMID: 19822006
2. Bogaert J, Prenen H. Molecular genetics of colorectal cancer. *Ann Gastroenterol.* 2014;27(1): 9–14. PMID: 24714764
3. Fearon ER. Molecular Genetics of Colorectal Cancer. *Annu Rev Pathol.* 2011;6(1): 479–507. doi: 10.1146/annurev-pathol-011110-130235 PMID: 21090969
4. NIH Genetics Home Reference. Familial adenomatous polyposis [Internet]. 2013 October [updated 2020 Aug 17; cited 2020 Sep 2] Available from: <https://ghr.nlm.nih.gov/condition/familial-adenomatous-polyposis#genes>
5. Alrawi SJ, Schiff M, Carroll RE, Dayton M, Gibbs JF, Kulavlat M, et al. Aberrant Crypt Foci. *Anticancer Res.* 2006 Jan-Feb;26(1A): 107–19. PMID: 16475686
6. University of Michigan Medicine. Colon and Rectal Polyps [Internet]. [cited 2020 Sep 2] Available from: <https://www.uofmhealth.org/conditions-treatments/digestive-and-liver-health/colon-and-rectal-polyps>
7. Stanford Medicine. Adenoma of the Colon and Rectum [Internet]. 2010 Jan 31 [updated 2014 Jan 19; cited 1 Sep 2020]. Available from: <http://surgpathcriteria.stanford.edu/gitumors/colorectal-adenoma/index.html>

8. Boivin GP, Washington K, Yang K, Ward JM, Pretlow TP, Russell R, et al. Pathology of Mouse Models of Intestinal Cancer: Consensus Report and Recommendations. *Gastroenterology*. 2003 Mar;124(3): 762–77. doi: 10.1053/gast.2003.50094 PMID: 12612914
9. Sausa E, Iraola-Guzmán S, Willis JR, Brunet-Vega A, Gabaldón T. Microbiome and colorectal cancer: Roles in carcinogenesis and clinical potential. *Mol Aspects Med*. 2019 Oct;69: 93–106. doi: 10.1016/j.mam.2019.05.001 PMID: 31082399
10. Takayama T, Katsuki S, Takahashi Y, Ohi M, Nojiri S, Sakamaki S. Aberrant Crypt Foci of the Colon as Precursors of Adenoma And Cancer. *J Surg Oncol*. 2008 Sep 1;98(3): 207–13. doi: 10.1002/jso.21106 PMID: 18623110
11. Mohamed M, Schofield JB. The pathology of colorectal polyps and cancers (including biopsy). *Surgery*. 2014;32(4): 165–71. doi: 10.1016/J.MPSUR.2014.02.004
12. Fleming M, Ravula S, Tatishchev SF, Wang HL. Colorectal carcinoma: Pathologic aspects. *J Gastrointest Oncol*. 2012 Sep;3(3): 153–73. doi: 10.3978/j.issn.2078-6891.2012.030. PMID: 22943008.
13. Thursby E, Juge N. Introduction to the human gut microbiota. *Biochem J*. 2017 Jun 1;474(11): 1823–1836. doi: 10.1042/BCJ20160510 PMID: 28512250
14. Louis P, Hold GL, Flint HJ. The gut microbiota, bacterial metabolites and colorectal cancer. *Nat Rev Microbiol*. 2014 Oct;12(10): 661–72. doi: 10.1038/nrmicro3344 PMID: 25198138
15. Ahn J, Sinha R, Pei Z, Dominianni C, Wu J, Shi J, et al. Human Gut Microbiome and Risk for Colorectal Cancer. *J Natl Cancer Inst*. 2013 Dec 18;105(24): 1907–11. doi: 10.1093/jnci/djt300 PMID: 24316595
16. Gagnière J, Raisch J, Veziant J, Barnich N, Bonnet R, Buc E, et al. Gut microbiota imbalance and colorectal cancer. *World J Gastroenterol*. 2016 Jan 14;22(2): 501–18. doi: 10.3748/wjg.v22.i2.501 PMID: 26811603
17. Zackular JP, Rogers MAM, Ruffin MT, Schloss PD. The Human Gut Microbiome as a Screening Tool for Colorectal Cancer. *Cancer Prev Res (Phila)*. 2014 Nov;7(11): 1112–21 doi: 10.1158/1940-6207.CAPR-14-0129 PMID: 25104642
18. Couturier-Maillard A, Secher T, Rehman A, Normand S, De Arcangelis A, Haesler R, et al. NOD2-mediated dysbiosis predisposes mice to transmissible colitis and colorectal cancer. *J Clin Invest*. 2013 Feb;123(2): 700–11. doi: 10.1172/JCI62236 PMID: 23281400
19. Canani RB, Di Costanzo M, Leone L, Pedata M, Meli R, Calignano A. Potential beneficial effects of butyrate in intestinal and extraintestinal diseases. *World J Gastroenterol*. 2011 Mar 28;17(12): 1519–28. doi: 10.3748/wjg.v17.i12. 1519 PMID: 21472114
20. Geirnaert A, Calatayud M, Grootaert C, Laukens D, Devriese S, Smagghe G, et al. Butyrate-producing bacteria supplemented in vitro to Crohn's disease patient microbiota increased butyrate production and enhanced intestinal epithelial barrier integrity. *Sci Rep*. 2017 Sep 13;7(1): 11450. doi: 10.1038/s41598-017-11734-8 PMID: PMC5597586
21. Balamurugan R, Rajendiran E, George S, Samuel GV, Ramakrishna BS. Real-time polymerase chain reaction quantification of specific butyrate-producing bacteria, *Desulfovibrio* and *Enterococcus faecalis* in the feces of patients with colorectal cancer. *J Gastroenterol Hepatol*. 2008 Aug;23(8 Pt 1): 1298–303. doi: 10.1111/j.1440-1746.2008.05490.x PMID: 18624900

22. Roediger WE. Utilization of nutrients by isolated epithelial cells of the rat colon. *Gastroenterology*. 1982 Aug;83(2): 424–9. PMID: 7084619
23. Hamer HA, Jonkers D, Venema K, Vanhoutvin S, Troost FJ, Brummer RJ. Review article: the role of butyrate on colonic function. *Aliment Pharmacol Ther*. 2008 Jan 15;27(2): 104–19. doi: 10.1111/j.1365-2036.2007.03562.x PMID: 17973645
24. Vinolo MAR, Rodrigues HG, Nachbar RT, Curi R. Regulation of Inflammation by Short Chain Fatty Acids. *Nutrients*. 2011 Oct;3(10): 858–76. doi: 10.3390/nu3100858 PMID: 22254083
25. Johnstone M, Bennett N, Standifer C, Smith A, Han A, Bettaieb A, et al. Characterization of the Pro-Inflammatory Cytokine IL-1b on Butyrate Oxidation in Colorectal Cancer Cells. *J Cell Biochem*. 2017 Jun;118(6): 1614–21. doi: 10.1002/jcb.25824 PMID: 27922186
26. Davie JR. Inhibition of histone deacetylase activity by butyrate. *J Nutr*. 2003 Jul;133(7 Suppl): 2485S–2493S. doi: 10.1093/jn/133.7.2485S PMID: 12840228
27. McNabney SM, Henagan TM. Short Chain Fatty Acids in the Colon and Peripheral Tissues: A Focus on Butyrate, Colon Cancer, Obesity and Insulin Resistance. *Nutrients*. 2017 Dec;9(12): 1348. doi: 10.3390/nu9121348. PMID: 29231905
28. Wu X, Wu Y, He L, Wu L, Wang X, Liu Z. Effects of the intestinal microbial metabolite butyrate on the development of colorectal cancer. *J Cancer*. 2018;9(14): 2510–17. doi: 10.7150/jca.25324 PMID: 30026849
29. Chang PV, Hao L, Offermanns S, Medzhitova R. The microbial metabolite butyrate regulates intestinal macrophage function via histone deacetylase inhibition. *Proc Natl Acad Sci U S A*. 2014 Feb 11;111(6): 2247–52. doi: 10.1073/pnas.1322269111 PMID: 24390544
30. Weng M, Walker WA, Sanderson IR. Butyrate Regulates the Expression of Pathogen-Triggered IL-8 in Intestinal Epithelia. *Pediatr Res*. 2007 Nov;62(5): 542–6. doi: 10.1203/PDR.ob013e318155a422 PMID: 17805211
31. Karim BO, Huso DL. Mouse models for colorectal cancer. *Am J Cancer Res*. 2013;3(3): 240–250. PMID: 23841024
32. Zhai S, Qin S, Li L, Zhu L, Zou Z, Wang L. Dietary butyrate suppresses inflammation through modulating gut microbiota in high-fat diet-fed mice. *FEMS Microbiol Lett*. 2019 Jul 1;366(13): fnz153. doi: 10.1093/femsle/fnz153 PMID: 31295342
33. Wrzosek L, Ciocan D, Borentain P, Spatz M, Puchois V, Hugot C, et al. Transplantation of human microbiota into conventional mice durably reshapes the gut microbiota. *Sci Rep*. 2018;8: 6854. doi: 10.1038/s41598-018-25300-3. PMID: 29717179.
34. Choi JG, Huh E, Kim N, Kim DH, Oh MS. High-throughput 16S rRNA gene sequencing reveals that 6-hydroxydopamine affects gut microbial environment. *PLoS One*. 2019;14(8): e0217194. doi: 10.1371/journal.pone.0217194 PMID: 31404072
35. McGinley JN, Thompson MD, Thompson HJ. A Method for Serial Tissue Processing and Parallel Analysis of Aberrant Crypt Morphology, Mucin Depletion, and Beta-Catenin Staining in an Experimental Model of Colon Carcinogenesis. *Biol Proced Online*. 2010;12: 118–130. doi: 10.1007/s12575-010-9032-x. PMID: 21406072.
36. Roncucci L, Stamp D, Medline A, Cullen JB, Bruce WR. Identification and quantification of aberrant crypt foci and microadenomas in the human colon.

- Hum Pathol. 1991 Mar;22(3): 287–94. doi: 10.1016/0046-8177(91)90163-j PMID: 1706308
37. He L, Li F, Yin X, Bohman P, Kim S, McClain CJ. Profiling of Polar Metabolites in Mouse Feces Using Four Analytical Platforms to Study the Effects Of Cathelicidin-Related Antimicrobial Peptide in Alcoholic Liver Disease. *J Proteome Res.* 2019 Jul 5;18(7): 2875–2884. doi: 10.1021/acs.jproteome.9b00181 PMID: 31188604
  38. Robanus-Maandag EC, Koelink PJ, Breukel C, Salvatori CF, Jagmohan-Changur SC, Bosch CAJ, et al. A new conditional Apc-mutant mouse model for colorectal cancer. *Carcinogenesis.* 2010 May;31(5): 946–52. doi: 10.1093/carcin/bgq046 PMID: 20176656
  39. Suzuki TA, Nachman MW. Spatial Heterogeneity of Gut Microbial Composition along the Gastrointestinal Tract in Natural Populations of House Mice. *PLoS One.* 2016;11(9): e0163720. doi: 10.1371/journal.pone.0163720 PMID: 27669007



# Ship of Theseus: Looking at Hagia Sophia's Cultural Complexity in the 21st Century

---

*Ruoxuan Fang*

**Author Background:** *Ruoxuan Fang grew up in China and currently attends Miss Porter's School in Farmington, Connecticut, USA. Her Pioneer research concentration was in the field of art history/architecture and titled "The Medieval Mediterranean: Confluence of Cultures."*

## 1. Introduction

On July 24th, 2020, Istanbul's Hagia Sophia (fig.1) welcomed its first group of Muslims for prayer since it adopted its secular identity in 1935. The identity shift is not new to the history of Hagia Sophia as it has undergone several phases of change ever since it was first built as a Christian church.

The current building is the third cathedral on the site. The first edifice of Hagia Sophia was constructed in the 4th century by Byzantine Emperor Constantine I and his son Constantine II, and had undergone a few cycles of dilapidation and restoration before it became the domed Orthodox Patriarchal Basilica built by Justinian I in the 6th century.<sup>1</sup> After the Ottomans conquered Constantinople in 1453, Hagia Sophia was repurposed into a mosque by the sultan Mehmed II and began its centuries-long Islamization. It was secularized into a museum in 1935 by Ataturk, the founder of the Turkish Republic, and Friday prayers for Muslims were recently resumed by the current Turkish president Recep Tayyip Erdogan.<sup>2</sup> Following the Turkish court's annulment of Ataturk's secularization, Erdogan's actions has attracted support from Turkey's conservative Muslims, as well as criticisms from all around the globe, claiming this is an act that not only underplays Hagia Sophia's legacy of both Christianity and Islam as well as that of the Byzantine and Ottoman Empire, but also furthers the religious divide in the Turkish State.<sup>3</sup>

---

<sup>1</sup> Current church was ordered by Emperor Justinian; designed by Greek scientists Isidorus of Miletus and Anthemius of Tralles; dome rebuilt by Isidore the Younger; church converted into a mosque By Ottoman Sultan Mehmed II. Current church built 532 - 537 AD over a site where two earlier church were built; the original in 360 AD and the second in 415 AD; both burned to the ground, in 404 AD and 532 AD, respectively; converted into a Mosque circa 1453. Image: April, 2010. Interior, Nave, Omphalion. [https://library.artstor.org/asset/ASITESPHO-TOIG\\_10313843839](https://library.artstor.org/asset/ASITESPHO-TOIG_10313843839).

<sup>2</sup> The Editors of Encyclopaedia Britannica, "Hagia Sophia," *Encyclopædia Britannica* (Encyclopædia Britannica, inc., August 5, 2020), last modified August 5, 2020, accessed August 9, 2020, <https://www.britannica.com/topic/Hagia-Sophia>.

<sup>3</sup> "Turkey Reconverts Istanbul's Hagia Sophia Museum into a Mosque," *News | Al Jazeera* (Al Jazeera, July 10, 2020), last modified July 10, 2020, accessed August 18, 2020, <https://www.aljazeera.com/news/2020/07/turkey-court-revokes-hagia-sophia-museum-status-200710131419431.html>; Judith Herrin, "Opinion | Converting Hagia Sophia into a Mosque Is an Act of Cultural Cleansing," *The Washington Post* (WP Company, July 15, 2020), last modified July 15, 2020, accessed August 16, 2020, <https://www.washingtonpost.com/opinions/2020/07/15/converting-hagia-sophia-into-mosque-is-an-act-cultural-cleansing/>.

In light of this incident, this paper looks at the impact of Hagia Sophia's reconversion by discussing its political and religious significance throughout history, specifically the intertwined relationship of rulership and religion that has framed the identity of Hagia Sophia over time. The changes made to its decorative program also correlate to the alterations in Hagia Sophia's function. By examining its multicultural heritage, the research conducted upholds the notion that Hagia Sophia's present-day religious affiliation obscures its cultural and historical complexity.

## 2. Historical Context - Timeline of the Complete History

### 2.1 Hagia Sophia in the Byzantium

The first basilica of Hagia Sophia was built by Constantine I on the foundations of a pagan temple and was dedicated to Christianity by his son Constantine II in the fourth century. It included a wooden roof and was burnt down in a riot in 404 C.E., during the reign of Emperor Arkardios. The reconstruction of Hagia Sophia was completed in 415 C.E. by Theodosios II and was ruined in the fire of the "Nika Insurrection" under Justinian I's rule. Justinian commissioned the third edifice of Hagia Sophia in 532 C.E., as a part of his rebuilding of Constantinople. This final construction of Hagia Sophia was completed in 5 years as the centerpiece of Justinian's restoration.<sup>4</sup>

The 6th-century rebuilding of Hagia Sophia was a monumental architectural achievement designed by the Byzantine theoreticians Isidorus and Anthemius.<sup>5</sup> Based on a traditional Orthodox basilica, the architects added a grand dome supported by triangular pendentives as a device to expand the interior space.<sup>6</sup>

Hagia Sophia's early decorations were removed in 726 C.E. due to Iconoclasm doctrines, which oppose visual representations of religious images; its mosaic features of religious figures, such as the Virgin and Child, were gradually added since the 9th century as religious icons grew to be accepted by the majority. In 1453, after the siege of Constantinople put the Byzantine era to an end, the Byzantine Hagia Sophia had become a part of the Islamic Ottoman Empire.<sup>7</sup>

### 2.2 Hagia Sophia 1453-1935<sup>8</sup>

Hagia Sophia has undergone multiple phases of architectural alterations, especially after its consecration to Islam in 1453. To preserve Hagia Sophia's Byzantine past,

<sup>4</sup> Fergus M. Bordewich, "A Monumental Struggle to Preserve Hagia Sophia," last modified December 1, 2008, accessed August 9, 2020, <https://www.smithsonianmag.com/travel/a-monumental-struggle-to-preserve-hagia-sophia-92038218/>.

<sup>5</sup> JAMES STEVENS CURL "Isidorus of Miletus ." A Dictionary of Architecture and Landscape Architecture. *Encyclopedia.com*. (August 11, 2020). <https://www.encyclopedia.com/education/dictionaries-thesauruses-pictures-and-press-releases/isidorus-miletus> "Anthemius of Tralles." Complete Dictionary of Scientific Biography. *Encyclopedia.com*. (August 11, 2020). <https://www.encyclopedia.com/science/dictionaries-thesauruses-pictures-and-press-releases/anthemius-tralles>

<sup>6</sup> Emma Wegner, "Hagia Sophia, 532–37;" The Editors of Encyclopaedia Britannica, "Hagia Sophia Bio - Britannica."

<sup>7</sup> Bordewich, "Struggle to Preserve Hagia Sophia."

<sup>8</sup> The overall chronology is taken from Robert Mark, A. S. Cakmak, and Gülru Necipoğlu, "The Life of an Imperial Monument: Hagia Sophia after the Byzantium," in *Hagia Sophia from the Age of Justinian to the Present* (Cambridge: Cambridge University Press, 1992), pp. 195-225.



Mehmed II imposed minimal physical changes despite declaring its new Islamic identity. After clearing the former church of a few of its Christian symbols and relics as well as plastering over the mosaics located within sight during prayers, a marble minbar, a mihrab, Muslim relics, and the first two minarets were added. Mehmed placed the mihrab at an angle facing Mecca with carpet strips conforming to the same direction to prevent any structural changes. Emblems of conquest were added to the converted mosque over time as a reminder of Islam's victory over Christianity. Most Byzantine features were preserved to serve the same purpose.

Hagia Sophia's first large-scale renovation took place in the 16th century under the reign of Selim II from 1572 to 1574. To ensure the stability of this centuries-old structure, Selim II not only ordered the surrounding houses, latrines, and kitchens to be destroyed, but also repaired the decayed buttresses and one of the wooden minarets built by Mehmed II. He then commissioned two additional minarets to represent his rulership. Aiming to enhance his relationship with Hagia Sophia, Selim II issued firmans that charged architectural experts with inspecting, restoring, and renovating the building. To further consolidate his connection to this prestigious monument, Selim II made Hagia Sophia his funerary mosque. This act of making Hagia Sophia a mausoleum site was followed by the succeeding sultans, thus declaring their acknowledgment of Hagia Sophia's special status.

The second wave of radical renovation came under the reign of Ahmed I in the 17th century, during which Hagia Sophia's decorative skin was subjected to significant changes. In addition to the renewal of inscriptions added by Mehmed II, Ahmed I ordered Islamic features, such as the royal tribune, the minbar, and the mihrab, to be repainted. The walls of the interior and exterior were whitewashed and many mosaics and frescoes, such as those of the Apostles, were covered by paint. Changes on the main dome included the replacement of lead sheets on the exterior and the covering of the central Pantokrator image depicting Christ with large Arabic letters. Though Ahmed's renovation revealed the growing iconoclastic sentiment, the mosaic Virgin and Child on the apse were preserved as well as those of the seraphim and archangels, which were considered elements aligned with the Quran. More Islamic scripts were added to forward the Islamization.

When Hagia Sophia's Byzantine identity became less apparent in the 18th century due to its reinforced connections to Islam, iconoclastic views peaked as more mosaics were whitewashed, including the archangels and the Virgin and Child on the apse, leaving only the Seraphim on the pendentives. These changes were accompanied by the exterior Islamic structures added by Mahmud I.

In the 19th century, Abdulmecid I encouraged the uncovering of the mosaics during the restoration of Hagia Sophia by the Fossati brothers as the Ottoman Empire embarked on the path of Westernization. Though he advocated for the visibility of the mosaics outside the prayer hall and the coverings of the other mosaics to be removable, he was countered by the more conservative crowd who upheld the dogmatic Islamic traditions. Eventually, all the mosaics were whitewashed except for the Seraphim, the faces of which were replaced with star medallions (fig.2). As conservative Islamic views continued to infuse the state, the Islamic elements in Hagia Sophia became excessively pronounced. The eight massive roundels of Arabic scripts (fig.3), with the names of Allah, the prophets, and the caliphates that were added during the Fossati restoration from 1847 to 1849, proclaim the Islamic domination forcefully, interrupting the balance of the interior.

### 2.3 Hagia Sophia 1934-July 2020

With the dismantling of the Ottoman Empire following WWI, Hagia Sophia was then subjected to more changes. Mustafa Kemal Atatürk, the agnostic founder of the Turkish Republic, aimed to secularize the Turkish state after he rose to power in 1923.<sup>9</sup> The secularization of Hagia Sophia in 1934 was one of the measures taken to uproot the state's religious affiliation. Its conversion into a museum in 1935, preceded by Thomas Whittemore's restoration of the Byzantine mosaics,<sup>10</sup> symbolized Atatürk's creation of a unified Turkish state that represented all demographics instead of one Islamic cohort.<sup>11</sup> Hagia Sophia's reconversion mainly stemmed from Atatürk's need for symbolic political propaganda to promote his secular political ideology.<sup>12</sup> He invoked Hagia Sophia's special prestige and the resulting reparation and the restoration of the mosaics resurfaced the clues leading to its Byzantine past. By displaying Byzantine Christian elements together with the Islamic features, Hagia Sophia embodies the coexistence of two distinct rival cultures. The ability to see both the newly restored Byzantine mosaics and the distinctly Islamic features simultaneously placed Hagia Sophia on a global stage and attracted scholars who were eager to uncover the original identity of this mosque-turned-museum. Atatürk considered his secularization of Hagia Sophia as a humanistic approach:

It will be reasonable to convert it [Hagia Sophia] into a museum, which will be open to the visits of all nations and religions rather than letting it belong to a single religion and a single group, and it will particularly be appropriate to collect the Byzantine works in this new museum.<sup>13</sup>

The museumization of Hagia Sophia seemed to neutralize the attributions of this religious building for the first time in history by presenting its dual identity, regardless of the chronology of specific religious affiliation glazed over with a new secular profile. The radical redefinition of this centuries-old mosque by Atatürk put this monument at the center of global controversy. "Hagia Sophia is a pawn in the game of intrigue between the secular and religious parties," Robert Ousterhout, a scholar of Byzantine art history, commented on the aftermath of Atatürk's conversion: "there's an alarmist response on both sides. They always assume the worst of each other. Secularists fear that religious groups are part of a conspiracy funded from

<sup>9</sup> Biography.com Editors. "Mustafa Kemal Atatürk," *Biography.com* (A&E Networks Television, June 22, 2020), last modified June 22, 2020, accessed August 14, 2020, <https://www.biography.com/political-figure/mustafa-kemal-ataturk>.

<sup>10</sup> Thomas Whittemore, a Byzantinist from Boston, founded the Byzantine Institute of America in 1930. From 1931 to 1938, Whittemore's team, led by E. J. W. Hawkins, cleaned the mosaics of the inner narthex, vestibule, gallery, apse, and north tympanum wall. Butler, Lawrence E. "Whittemore, Thomas." *Grove Art Online*. 2003; Accessed 19 Aug. 2020. <https://www.oxfordartonline.com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-7000091446>.

<sup>11</sup> Tanyeri-Erdemir, Tuğba. "Archaeology as a Source of National Pride in the Early Years of the Turkish Republic." *Journal of Field Archaeology* 31, no. 4 (2006): 381.

<sup>12</sup> Ceren Katıpoğlu, and Çağla Caner-Yüksel. "Hagia Sophia 'Museum': A Humanist Project of the Turkish Republic." In *Constructing Cultural Identity, Representing Social Power*, 214. Pisa University Press, n.d. [https://www.academia.edu/23804647/Hagia\\_Sophia\\_Museum\\_A\\_Humanist\\_Project\\_of\\_the\\_Turkish\\_Republic](https://www.academia.edu/23804647/Hagia_Sophia_Museum_A_Humanist_Project_of_the_Turkish_Republic), 214.

<sup>13</sup> Quoted in Bilsel Candaş, Ceren Katıpoğlu, and Çağla Caner-Yüksel, "Hagia Sophia 'Museum': A Humanist Project of the Turkish Republic," in *Constructing Cultural Identity, Representing Social Power* (Pisa: Plus, 2010), 214.

Saudi Arabia, while religious people fear that the secularists want to take their mosques away from them.”<sup>14</sup> The debate continued as religious groups argued that Hagia Sophia’s conversion into a museum marked the demise of its existence as an active religious institution, whereas the secularists advocated for the nationalism of the Turkish Republic Hagia Sophia entailed.<sup>15</sup> This transitional period also fueled the Orthodox Christians’ plea of returning Hagia Sophia to its original Christian function. Chris Spirou, the president of the Free Hagia Sophia Council of America, pleaded against “the ruthlessness, the insensitivity and the barbaric behavior of rulers and conquerors towards human beings and their rights” at the Congressional Human Rights Caucus in Washington D.C. in 2007.<sup>16</sup> The Muslim counterpart’s opposing sentiments have been long-lasting. Necip Fazıl Kısakurek, a renowned Turkish poet mourned that Hagia Sophia’s conversion had “put the Turks’ essential spirit inside a museum.”<sup>17</sup>

As a universally recognized world heritage site, Hagia Sophia continues to be a center of dispute in large part due to its embodiment of both Christian and Islamic history. Many of its conversions throughout history have been due to violent conquest; all have generated debates over religious and political orientations. Given the passionate debate the building continues to spark, it is clear that Hagia Sophia still acts as an emblem of cultural complexity.

### 3. Hagia Sophia’s Cultural Ambivalence

#### 3.1 The Cross-Cultural Dialogue

When each culture claimed ownership of Hagia Sophia, this prestigious piece of architecture became the agent that broadcasted the message each dominant party desired to proclaim. As the values of the rulers change over time, the message is no longer static. We see the struggle between the iconoclasts and the “iconodules” in the Byzantine era and repetitive acts of covering and revealing the Byzantine mosaics by multiple Muslim Ottoman emperors.<sup>18</sup> Although different political and religious authorities tried to legitimize their narrative of Hagia Sophia while negating the others, they are simultaneously “mutually reinforcing” each other’s heritage.<sup>19</sup> This was enhanced after the Hagia Sophia’s secularization, as the museum exhibited the elements from both religions, facilitating a dialogue of both religions in the same space and time.

##### 3.1.1 The Christian Legacy

The restoration of the Byzantine mosaics by Thomas Whittemore in the 20th century on the interior amplified the dimensions of Hagia Sophia by revealing its Christian

<sup>14</sup> Bordewich, “Struggle to Preserve Hagia Sophia.”

<sup>15</sup> Robert S Nelson, Margaret Olin, and Robert S Nelson, “Tourists, Terrorists, and Metaphysical Theatre at Hagia Sophia,” in *Monuments and Memory, Made and Unmade* (Chicago: The University of Chicago Press, 2003), 69.

<sup>16</sup> Bordewich, “Struggle to Preserve Hagia Sophia.”

<sup>17</sup> Selim Koru. “Turkey’s Islamist Dream Finally Becomes a Reality.” *The New York Times*, July 14, 2020. <https://www.nytimes.com/2020/07/14/opinion/hagia-sophia-turkey-mosque.html>.

<sup>18</sup> Necipoğlu, “The Life of an Imperial Monument,” 202-225.

<sup>19</sup> Bigelow, Anna. “Hagia Sophia’s Tears and Smiles.” In *Istanbul*, edited by Nora Fisher-Onar, Susan C. Pearce, and E. Fuat Keyman, 112–28. Living with Difference in a Global City. Rutgers University Press, 2018. <https://doi.org/10.2307/j.ctt2ovxq27.12>.

identity. The designs have far exceeded their materiality, that is, the significance of the mosaics and other elements on the interior is greater than a mere literal reading of their subject matter, for they provide a connection between the earthly and spiritual realms.<sup>20</sup>

The nave floor (fig.4) creates a grand setting that frames the divine connection of Hagia Sophia. Fabio Berry speculates in his essay *Walking on Water: the Cosmic Floors in Antiquity and the Middle Ages* that the marble panels in Hagia Sophia resemble the concept of “frozen water,” which alludes to the Bible verse in Job 37:10 “the breath of God produces ice, and the broad waters become frozen.”<sup>21</sup> Moreover, the wavy patterns of the marble slabs mimic the celestial sea above which God’s throne hovers, emphasizing God’s presence. This could also be a reiteration of the world’s “watery genesis,”<sup>22</sup> which leads the visitors into the illusion that brings them closer to divinity.

The spiritual association is an overarching theme of Hagia Sophia. One of the most prominent mosaics uncovered is the Virgin and Child on the apse semi-dome (fig.5).<sup>23</sup> Overseeing the central hall of worship, the Theotokos<sup>24</sup> mosaic is illuminated by the light pouring through the windows at the bottom of the apse. With a height of 4.39 m (from the top of her hood to the left foot), the Virgin Mary sits on two layers of cushions supported by a jeweled bench with her feet on a jeweled steppingstone. The child Jesus, of approximately 2m (without the halo) in height, sits on Mary’s lap in a golden robe, which matches the golden background symbolic of his divine relation with heaven. The unveiling of the Ascension of Christ mosaic was a statement put forth by emperors Basil I and Michael III in the late 9th century, signaling the end of iconoclasm and the return of figural imagery.<sup>25</sup> This was echoed by a Greek inscription on the apse that is no longer visible, which stated: “the images which the imposters have cast down, here pious emperors have set up again.”<sup>26</sup> Aside from its role as a riposte against iconoclasm, this classic biblical scene elicits its more profound purpose. Flanked by Archangel Gabriel and Michael on the bema arches (fig.6), the ascension of Christ represents the moment in which the earth is joined with heaven. Because angels serve as messengers of God, their presence is an indication that Hagia Sophia embodies the reconciliation between divinity and humanity.<sup>27</sup>

<sup>20</sup> James, Liz. “Matters of Materiality in Byzantium. The Archangel Gabriel in Hagia Sophia, Constantinople.” *Konsthistorisk Tidskrift/Journal of Art History* 86, no. 3 (July 3, 2017): 145–57. <https://doi.org/10.1080/00233609.2017.1330279>.

<sup>21</sup> Bible Hub. “Job 37:10.” Accessed August 7, 2020. <https://biblehub.com/job/37-10.htm>.

<sup>22</sup> Barry, Fabio. “Walking on Water: Cosmic Floors in Antiquity and the Middle Ages.” *Art Bulletin* 89, no. 4 (December 2007): 627–56.

<sup>23</sup> Mango, Cyril, and Ernest J. W. Hawkins. “The Apse Mosaics of St. Sophia at Istanbul. Report on Work Carried out in 1964.” *Dumbarton Oaks Papers* 19 (1965): 113–51. <https://doi.org/10.2307/1291228>.

<sup>24</sup> Theotokos is a term in Eastern Orthodox for Virgin Mary. The Editors of Encyclopaedia Britannica, “Theotokos,” *Encyclopædia Britannica* (Encyclopædia Britannica, inc., December 1, 2017), last modified December 1, 2017, accessed August 16, 2020, <https://www.britannica.com/topic/Theotokos>.

<sup>25</sup> Hagia Sophia Research Team. “The Virgin and Child Mosaic.” Accessed August 7, 2020. <https://hagiasophiatrkey.com/virgin-child-mosaic/>.

<sup>26</sup> Steven Zucker and Beth Harris, “Theotokos Mosaic, Hagia Sophia, Istanbul,” *Smarthistory*, last modified December 15, 2015, accessed August 15, 2020, <https://smarthistory.org/theotokos-mosaic-hagia-sophia-istanbul/>.

<sup>27</sup> James, “Matters of Materiality,” 7.

This connection is uplifted by the four seraphim on each of the pendentives supporting the main dome (fig.7). Seraphim are the highest-ranked angels in Christianity with their six wings representing the six days God took to create the earth.<sup>28</sup> With flame-like appearances, they are the ministers of God that are closest to his throne<sup>29</sup> according to Isaiah 6:2: “Attending him were mighty seraphim, each having six wings. With two wings they covered their faces, with two they covered their feet, and with two they flew.”<sup>30</sup> Being God’s entourage, the Seraphim around the pendentives in Hagia Sophia upheld the significance of the Pantokrator in the center of the dome (now covered by Arabic scripts). This further consolidates Hagia Sophia’s sublimation into divinity.

A different set of Theotokos’ mosaics displays the Byzantine emperors’ direct relationship with the divine (fig. 8). The image on the southwestern vestibule depicts the Virgin and Child flanked by the Emperor Justinian I offering a model of Hagia Sophia on the left and Constantine I holding a model of Constantinople on the right. The Greek inscription next to Constantine I states, “Constantine, the great emperor among the saints,” and the one next to Justinian I states, “Justinian, emperor of illustrious memory.” The inscriptions were likely made in the late 10th century, far beyond the reign of both emperors.<sup>31</sup> These captions highlight the achievements of Constantine I and Justinian I. Constantine I was compared with the saints while Justinian I was given the title “emperor of illustrious memory,” with the term “illustrious memory” likely referring to his reconstruction of Hagia Sophia as a manifestation of the Christian heritage.

### 3.1.2 The Islamic Era

Despite the lack of figural elements due to Islam’s aniconic doctrines, the Islamization efforts made during Ottomans’ centuries-long occupation were powerful and long-lasting. Some key Islamic features found at Hagia Sophia today are the roundels of Arabic scripts, the mihrab, the minbar, and the sultan’s lounge, in addition to the four minarets on the exterior. Similar to the Christian images on the interior that connect believers with the higher realm, the Islamic symbols lead their beholders to savor the essence of Islam.

The mihrab, minbar, and the sultan’s lounge (fig. 9) become immediately visible as one enters Hagia Sophia because they are placed on the platform directly under the main apse. A traditional mihrab is an ornamented niche in the qibla wall of a mosque facing Mecca; however, Hagia Sophia was not originally intended to be oriented towards that direction. Thus, the mihrab stands at an angle to the wall of the main apse together with the minbar and sultan’s lounge. The pattern of the flooring on the platform is placed in the same direction as these three Islamic insertions, which helps to create the illusion that this part of the building is oriented towards Mecca. Both the minbar and the sultan’s lounge are elevated to showcase the imams’

---

<sup>28</sup> AR Editors, “Hagia Sofia Restoration Reveals Face of a Seraph Angel Mosaic,” *Architectural Review*, last modified July 16, 2020, accessed August 7, 2020, <https://www.architectural-review.com/essays/hagia-sofia-restoration-reveals-face-of-a-seraph-angel-mosaic>.

<sup>29</sup> “The Divinely-Revealed Appearance of Angels in Icons,” *A Reader’s Guide to Orthodox Icons*, last modified January 12, 2012, accessed August 7, 2020, <https://iconreader.wordpress.com/2012/01/12/the-divinely-revealed-appearance-of-angels-in-icons/>.

<sup>30</sup> Bible Hub. “Isaiah 6:2.” Accessed August 15, 2020. <https://biblehub.com/isaiah/6-2.htm>.

<sup>31</sup> Whittemore, Thomas. “The Mosaics of St. Sophia at Istanbul.” *American Journal of Archaeology* 42, no. 2 (1938): 219–26. <https://doi.org/10.2307/499667>.

and the sultans' relationships with the divine, as the sultans were seen as divinely mandated figures and the imams preached the holy words of Allah.<sup>32</sup>

The four pencil minarets (fig.10) contain the elevated spaces on the exterior as an architectural attempt to reach the heavenly world. Since they represent the royal power of sultans, these pencil minarets have become a distinguishing feature of the Turkish skyline for their unique style and their prominent role in Ottoman architecture.<sup>33</sup>

Another signifier of Hagia Sophia's Islamic identity is the eight calligraphic medallions (fig. 3) that were added during the Fossati restoration in the 18th century to the surrounding columns of the nave, each with a diameter of almost 8 m.<sup>34</sup> The golden Arabic scripts inscribe the names of Allah, Mohammed, and his grandsons, and the four Islam Caliphates.<sup>35</sup> Aside from broadcasting Islam's core figures as a proclamation of Islamic values in Hagia Sophia, the Arabic calligraphy itself is of great significance, as Blankenship claims in journal *Cultural Considerations: Arabic Calligraphy and Latin Typography*:

Arabic calligraphy expresses the significance of timelessness with a marked sense of rhythm, and with endless repetitions and decorative patterns... The abstract nature of the designs is more significant than the material aspect. Complexity is created to hold the reader's attention. Contemplation is encouraged; the reader delights in the intricacy...The surface of a page can have a transcendent quality which achieves a desired spiritual essence. The beauty of the world is ephemeral; and ornamentation is used to define and emphasize the functional which links the inside and outside.<sup>36</sup>

Displayed alongside the Byzantine mosaics, these roundels are a declaration of a different form of beauty. While both types of religious aesthetics have significant connotations, this conflict of religious ideology cries out for a different symbolic reading. Though the two distinct styles seem to compete for the more dominant role, they underline each other's presence at the same time, thus facilitating Hagia Sophia's religious coexistence.

### 3.2 The Intertwined identity

The Ottoman Empire's conquest of Constantinople simply meant a continuation of Hagia Sophia's Byzantine Christian past, though in a different form. The Ottoman Emperor Mehmed II immediately recognized Hagia Sophia's prestige and was eager to inherit the imperial power established by the Byzantine emperors. Holding the admiration of Byzantium and this magnificent piece of architecture, Mehmed II strived to justify his identity as the legitimate heir of the Byzantine Empire. Rather than constructing a superior-inferior complex, he chose to succeed the Byzantine heritage.

<sup>32</sup> Steven Zucker and Beth Harris, "Hagia Sophia as a Mosque (Video)," *Khan Academy* (Khan Academy, n.d.), accessed August 21, 2020, <https://www.khanacademy.org/humanities/ap-art-history/early-europe-and-colonial-americas/medieval-europe-islamic-world/v/hagia-sophia-mosque>.

<sup>33</sup> Zucker and Harris; "Hagia Sophia as a Mosque."

<sup>34</sup> Necipoğlu, "The Life of an Imperial Monument," 223.

<sup>35</sup> Hagia Sophia Research Team. "Calligraphic Roundels." Hagia Sophia Turkey. Accessed August 21, 2020. <https://hagiasophiatrkey.com/calligraphic-roundels/>.

<sup>36</sup> Sherry Blankenship, Sherry. "Cultural Considerations: Arabic Calligraphy and Latin Typography." *Design Issues* 19, no. 2 (March 2003): 60–63.



Though Hagia Sophia had become an emblem of Islam's victory over Christianity, Mehmed II still strove to legitimize his direct relationship with Hagia Sophia. To accomplish this, he fabricated a history that constructed a collective memory for the Byzantine and Ottoman Empires. To justify the keeping of the name Hagia Sophia, one of the myths states that Ayasofya (Hagia Sophia in Arabic) was divinely suggested to the Emperor Justinian in a dream by the messenger Hizir. To preserve the building and its original decorations, the myths utilized the stories of the prophet and a Muslim martyr to argue that Hagia Sophia was built consecrated to Islam; thus it shouldn't be subjected to radical changes. Though these myths seemed to only cater to Hagia Sophia's Islamic connection, the theme of the tales continued to imply that Hagia Sophia was a shared religious sanctuary for both Christianity and Islam.<sup>37</sup> In addition, Mehmed II appropriated the Christian legends linked to Hagia Sophia to expand the series of myths, further integrating the Christian and Islamic past through paralleled associations with the building.<sup>38</sup>

The significant changes made to Hagia Sophia in the generations after Mehmed II also mirror the original Byzantine model. For instance, the verse from the Quran which reads "Allah is the light of the heavens and the earth" (fig. 11) that used to replace the Pantokrator was consistent with the covered Greek inscription on the dome "Peace be with you; I am the light of the world." Both caption the light pouring through the ring of windows on the dome. The resemblance implies that the Islamization was conducted based on a thorough comprehension of the Christian aspects displayed on Hagia Sophia's interior.<sup>39</sup>

Though Hagia Sophia and other Islamic mosques previously shared architectural vocabulary concerning domed structures, the grandeur of Hagia Sophia has had a profound influence on Islamic architecture. The architectural style of Fatih Camii (fig. 12), the funerary mosque for Sultan Mehmed II that replaced the Church of the Holy Apostles, referenced Hagia Sophia. Fatih Camii, meaning Conqueror's Mosque, includes one central pendentive dome with surrounding complementary domes that resemble those of Hagia Sophia.<sup>40</sup> The domed exterior of Hagia Sophia continued to inspire generations of mosques, such as the Blue Mosque and the Suleymaniye Mosque (fig. 13&14). Thus, what was considered the Byzantine past has become a part of the Islamic present.

#### 4. The Restoration of Hagia Sophia as a Functional Mosque

In July 2020, the timeless dialogue cultivated by the Hagia Sophia Museum was terminated by a Turkish court decision that annulled Ataturk's 1934 secularization. The annulment provided Turkish president Recep Tayyip Erdogan with room for radical changes. Measures were immediately taken, preparing Hagia Sophia to host the five daily prayers as a functioning mosque.<sup>41</sup> To create a decidedly Islamic space for worship, the delicate Christian mosaics were veiled. Turquoise carpets were then used to cover the marble flooring to provide a suitable surface for the requisite prostrations

---

<sup>37</sup> Necipoğlu, "The Life of an Imperial Monument," 200-201.

<sup>38</sup> Necipoğlu, "The Life of an Imperial Monument," 201.

<sup>39</sup> Zucker and Harris, "Theotokos mosaic, Hagia Sophia, Istanbul."

<sup>40</sup> Ousterhout, Robert. "The East, the West, and the Appropriation of the Past in Early Ottoman Architecture." *Gesta* 43, no. 2 (January 1, 2004): 170. <https://doi.org/10.2307/25067103>.

<sup>41</sup> Aljazeera, "Turkey Reconverts Istanbul's Hagia Sophia Museum into a Mosque."

during prayers.<sup>42</sup> Though there weren't any physical changes, the masking of the Byzantine elements was a means of obscuring the Byzantine past and the scene of religious co-existence in Hagia Sophia. Similar to how Hagia Sophia's conversion into a museum was a symbolic move for Kemal Ataturk's establishment of his secular regime, this decree signaled the expansion of Erdogan's Turkish Justice and Development Party's power. As Erdogan's political power has waned over the years, suggested by his recent defeat in the municipal election, this act answers his desperate need to advance his campaign as he earned the favor of the pan-Islamic voters.<sup>43</sup> Indeed, this decision fulfilled the dreams of the Turkish Conservative Islamists who identified the secularization of Hagia Sophia as "a seal on the country's spirit." Returning Hagia Sophia to its previous Islamic identity was a declaration against the secular government established by Ataturk which was considered a surrender to the Western powers.

Hagia Sophia's restoration into a mosque was an expression of Erdogan's attempt to reassert the Ottoman conquest over Byzantium, and Islam's conquest over Christianity, as Erdogan claimed in his presidential speech addressing his decision, "The conquest of Istanbul and the conversion of the Hagia Sophia into a mosque are among the most glorious chapters in Turkish history."<sup>44</sup> In the same speech, Erdogan drew parallels between himself and Fatih Mehmed II, the sultan who conquered Constantinople and converted Hagia Sophia to a mosque. He branded this reconversion of Hagia Sophia as a "second conquest, a new resurrection."<sup>45</sup> Hagia Sophia's return to a mosque requires more than a simple reading; it represents Erdogan's effort to reclaim the centuries-old conquest narrative. This was his way of institutionalizing Islam's victory over Christianity, which further marginalizes other religious groups in Turkey.

This event put Hagia Sophia once again at the peak of global controversy. On July 10th, UNESCO expressed its mournful sentiment toward this sudden reconversion of a world heritage site. With its recent religious affiliation, there have been doubts regarding whether the reconversion is going to deplete Hagia Sophia of its universal value. Ecumenical Patriarch Bartholomew, the spiritual leader of the Eastern Orthodox Church, worried that this decision was going to "fracture the two worlds."<sup>46</sup> Elizabeth Prodromou, professor of religion and geopolitics at Tufts University, agreed that Erdogan's decision "has been a lightning rod for a synthesis of religionationalism and instrumentalized as a symbol by the Erdogan government."<sup>47</sup>

The secularization of Hagia Sophia was Turkey's chance to achieve redemption, after being deeply wounded by its past acts of religious purification.<sup>48</sup> However,

---

<sup>42</sup> Aljazeera. "Muslim Prayers in Hagia Sophia for First Time in 86 Years," July 24, 2020. <https://www.aljazeera.com/news/2020/07/turkey-muslim-prayers-hagia-sophia-time-86-years-200724063303261.html>.

<sup>43</sup> Herrin, "Opinion | Converting Hagia Sophia into a Mosque Is an Act of Cultural Cleansing."

<sup>44</sup> TRT World. "President Erdogan's public address on Hagia Sophia", *Youtube* video, 33:30, July 13, 2020. <https://www.youtube.com/watch?v=ox-ROvVnPXU>

<sup>45</sup> *Ibid.*

<sup>46</sup> Merrit Kennedy and Peter Kenyon, "Turkey Converts Istanbul's Iconic Hagia Sophia Back Into A Mosque," *NPR* (NPR, July 10, 2020), last modified July 10, 2020, accessed August 22, 2020, <https://www.npr.org/2020/07/10/889691777/turkish-court-ruling-clears-way-for-hagia-sophia-to-be-converted-to-a-mosque>.

<sup>47</sup> *Ibid.*

<sup>48</sup> Esra Akcan, "Erasing History at the Hagia Sophia," *Berkley Center Fo Religion, Peace and World Affairs*, last modified July 27, 2020, accessed August 23, 2020, <https://berkley-center.georgetown.edu/responses/erasing-history-at-the-hagia-sophia>.



---

Erdogan's revisitatioin to this national scar in the name of Turkish sovereignty could only exacerbate the cultural antagonism.<sup>49</sup>

## 5. Conclusion

From the Byzantine Era to the Ottoman Empire to the present-day Turkish Republic, Hagia Sophia has been served as the agent of different political and religious ideologies. These ideologies, reflected by the aesthetic designs of Hagia Sophia, have complex relationships. A large part of the Ottoman history as well as the history of Islam was constructed around the existence of Hagia Sophia, previously a cathedral built by non-Muslims, while the structure of Hagia Sophia has been a lingering influence in later Ottoman architecture. In the 20th century, Ataturk's secularization of Hagia Sophia had exceeded its political implications by creating the religious harmony which acknowledges all of the voices that used to narrate the identity of this spectacular monument.

The recent reconversion reveals Erdogan's intentions of reviving Turkey's Ottoman past. His intentions behind the masking of Hagia Sophia's Byzantine features, however, fail to imitate those of his Ottoman predecessors, which generally stemmed from religious tolerance and the respect dedicated to the prestige of the erstwhile Byzantine Empire. Since the 20th century, Hagia Sophia has become the manifestation of religious and cultural diversity. It symbolizes a timeless collaboration of past cultures whose profound implications are best understood in a museum context. To reconvert it is to put a stop to the collective memory of both Christianity and Islam that leads to the audiences' meditation on Hagia Sophia's complex history.

---

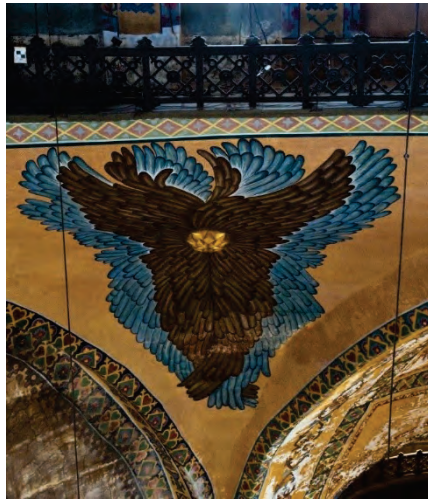
<sup>49</sup> TRT World, "President Erdogan's public address on Hagia Sophia."

## Images



**Figure 1. Hagia Sophia exterior**

532-537, Photographed: 1954. Hagia Sophia. [https://library.artstor.org/asset/ABRMAWR\\_MELLINKIG\\_10310735547](https://library.artstor.org/asset/ABRMAWR_MELLINKIG_10310735547).



**Figure 2. Seraphim on the Pendentive with Star Medallion**

532-537, Image: 25-Jun-09. Hagia Sophia. [https://library.artstor.org/asset/AEKI-NIG\\_10313455639](https://library.artstor.org/asset/AEKI-NIG_10313455639).



**Figure 3. Calligraphic Roundels on the interior of Hagia Sophia**  
Image: April, 2010. Interior, Nave, Medallions with Inscriptions in Arabic.

[https://library.artstor.org/asset/ASITESPHOTOIG\\_10313839717](https://library.artstor.org/asset/ASITESPHOTOIG_10313839717).



**Figure 4. Nave floor**

Barry, Fabio. "Walking on Water: Cosmic Floors in Antiquity and the Middle Ages." *The Art Bulletin* 89, no. 4 (2007): 627-56. Accessed August 26, 2020.

<http://www.jstor.org/stable/25067354>.

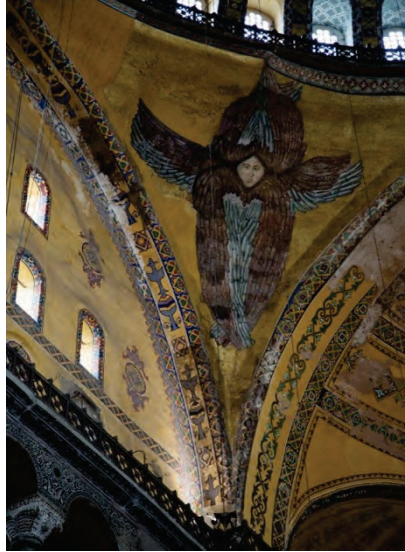


**Figure 5. Virgin and Child Mosaics on the apse**  
 Shortly after 834, dedicated 867. Constantinople, Hagia Sophia, Virgin and Child, Constantinople: Hagia Sophia: apse mosaic: Theotokos, Constantinople: Hagia Sophia: apse mosaic: Theotokos: det.: Head of Virgin Mary, Constantinople: Hagia Sophia: apse mosaic: Theotokos, Constantinople: Hagia. mosaic, mosaic. [https://library.artstor.org/asset/LESSING\\_ART\\_10311442169](https://library.artstor.org/asset/LESSING_ART_10311442169).



**Figure 6. Archangel Gabriel on the bema arch**  
 Current Church was ordered by Emperor Justinian; designed by Greek scientists Isidore of Miletus and Anthemius of Tralles. Current church built 532 - 537 AD; Mosaic probably dates to the 9th century, set against the original 6th-century Golden Background, Image: April, 2010. Interior, Nave, Apse, Archangel Gabriel Mosaic. [https://library.artstor.org/asset/ASITESPHO-TOIG\\_10313839655](https://library.artstor.org/asset/ASITESPHO-TOIG_10313839655).





**Figure 7. Seraphim on pendetive with face revealed**

Isidore of Miletus, Anthemius of Tralles. beginning date 532, completed in 537, Image: 5/1/2012. Hagia Sophia; view of the dome with four seraphim painted on the pendatives. [https://library.artstor.org/asset/ARTON-FILE\\_DB\\_10313670712](https://library.artstor.org/asset/ARTON-FILE_DB_10313670712).



**Figure 8. Virgin and Child flanked by Constantine I and Justinian on the Southwestern Vestibule**

Late 10th C. Constantinople, Hagia Sophia, Virgin and Child Enthroned between Emperors Justinian I, with the Model of Hagia Sophia [left], and Constantine I, with the Model of Byzantium, Constantinople: Hagia Sophia: outer narthex mosaic: enthroned Madonna between empe. mosaic. [https://library.artstor.org/asset/LES-SING\\_ART\\_10311441194](https://library.artstor.org/asset/LES-SING_ART_10311441194).



**Figure 9. Sultan's Lounge (left); Mihrab (under the apse); Minbar (right)**

*Sebah & Joaillier. Interieur de la Mosquée de Ste. Sophie. Vue generale - [Interior of Hagia Sophia, general view]. still image. Place: Special Collections and College Archives, Musselman Library, Gettysburg College. [https://library.artstor.org/asset/SS7731291\\_7731291\\_10908773](https://library.artstor.org/asset/SS7731291_7731291_10908773).*



**Figure 10. Hagia Sophia's four pencil minarets on the exterior.**  
Anthemius of Tralles. 532-537 AD, Photographed: 1956. *Hagia Sophia*. [https://library.artstor.org/asset/ABRMAWR\\_MELLINKIG\\_10310736414](https://library.artstor.org/asset/ABRMAWR_MELLINKIG_10310736414).



**Figure 11. The Arabic scripts on the central dome**

Constantinople: Hagia Sophia Int: central dome. [https://library.artstor.org/asset/ARTSTOR\\_103\\_41822003426101](https://library.artstor.org/asset/ARTSTOR_103_41822003426101).



**Figure 12. Fatih Camii in Istanbul**

Elbaksawy, Ezzeldin. *Sultan Muhammet Fatih camii*. Photograph. Istanbul, December 15, 2019.





**Figure 13. Blue Mosque in Istanbul.**

*Built by Sultan Ahmed I; designed by royal architect Mimar Sedefkar Mehmet Aga; tiles designed and supervised by master potter Kasap Haci and Baris Efendi from Avanos; Decorative Qur'anic verses mostly made by the calligrapher Seyyid Kasim Gubari. Built circa 1609-1616, Image: April, 2010. Sultan Ahmed Mosque / Sultanahmet Camii / "The Blue Mosque". [https://library.artstor.org/asset/ASITESPHOTOIG\\_10313840166](https://library.artstor.org/asset/ASITESPHOTOIG_10313840166).*



**Figure 14. Süleymaniye Mosque in Istanbul**

*Sebah & Joaillier. Vue panoramique de la Mosquée Suleymanie - [Exterior view of Sulieman's Mosque]. still image. Place: Special Collections and College Archives, Musselman Library, Gettysburg College. [https://library.artstor.org/asset/SS7731291\\_7731291\\_10908785](https://library.artstor.org/asset/SS7731291_7731291_10908785).*



## Bibliography

- Aljazeera. "Muslim Prayers in Hagia Sophia for First Time in 86 Years," July 24, 2020. <https://www.aljazeera.com/news/2020/07/turkey-muslim-prayers-hagia-sophia-time-86-years-200724063303261.html>.
- "Anthemius of Tralles ." Complete Dictionary of Scientific Biography. Encyclopedia.com. (August 11, 2020). <https://www.encyclopedia.com/science/dictionaries-thesauruses-pictures-and-press-releases/anthemius-tralles>.
- Barry, Fabio. "Walking on Water: Cosmic Floors in Antiquity and the Middle Ages." *Art Bulletin* 89, no. 4 (December 2007): 627–56.
- Barry, Fabio. "Walking on Water: Cosmic Floors in Antiquity and the Middle Ages." *The Art Bulletin* 89, no. 4 (2007): 627-56. Accessed August 26, 2020. <http://www.jstor.org/stable/25067354>.
- Bible Hub. "Job 37:10." Accessed August 7, 2020. <https://biblehub.com/job/37-10.htm>.
- Bible Hub. "Isaiah 6:2." Accessed August 15, 2020. <https://biblehub.com/isaiah/6-2.htm>.
- Bigelow, Anna. "Hagia Sophia's Tears and Smiles:" In *Istanbul*, edited by Nora Fisher-Onar, Susan C. Pearce, and E. Fuat Keyman, 112–28. Living with Difference in a Global City. Rutgers University Press, 2018. <https://doi.org/10.2307/j.ctt20vx-q27.12>.
- Bilsel Candaş, Ceren Katipoglu, and Cagla Caner-Yuksel. "Hagia Sophia 'Museum': A Humanist Project of the Turkish Republic." Essay. In *Constructing Cultural Identity, Representing Social Power*. Pisa: Plus, 2010.
- Bordewich, Fergus M. "A Monumental Struggle to Preserve Hagia Sophia." Last modified December 1, 2008. Accessed September 3, 2020. <https://www.smithsonianmag.com/travel/a-monumental-struggle-to-preserve-hagia-sophia-92038218/>.
- Butler, Lawrence E. "Whittemore, Thomas." Grove Art Online. 2003; Accessed 4 Sep. 2020. <https://www.oxfordartonline.com/groveart/view/10.1093/gao/9781884446054.001.0001/oao-9781884446054-e-7000091446>.
- Constantinople: Hagia Sophia Int: Central Dome. Artstor. Accessed 2020. [https://library.artstor.org/#/asset/ARTSTOR\\_103\\_41822003426101](https://library.artstor.org/#/asset/ARTSTOR_103_41822003426101).
- Editors, AR. "Hagia Sofia Restoration Reveals Face of a Seraph Angel Mosaic." *Architectural Review*. Last modified July 16, 2020. Accessed August 7, 2020. <https://www.architectural-review.com/essays/hagia-sofia-restoration-reveals-face-of-a-seraph-angel-mosaic>.
- Ekin, A. Cemal. *Hagia Sophia*. July 25, 2009. Artstor. [https://library.artstor.org/#/asset/AEKINIG\\_10313455639](https://library.artstor.org/#/asset/AEKINIG_10313455639).
- Esra Akcan, "Erasing History at the Hagia Sophia," *Berkley Center Fo Religion, Peace and World Affairs*, last modified July 27, 2020, accessed August 23, 2020, <https://berkeleycenter.georgetown.edu/responses/erasing-history-at-the-hagia-sophia>.
- Hagia Sophia Research Team. "The Virgin and Child Mosaic." Accessed August 7, 2020. <https://hagiasophiatrkey.com/virgin-child-mosaic/>.
- Hagia Sophia Research Team. "Calligraphic Roundels." Hagia Sophia Turkey. Accessed August 21, 2020. <https://hagia-sophiatrkey.com/calligraphic-roundels/>.

- Hagia Sophia; View of the Dome with Four Seraphim Painted on the Pendentives.* May 1, 2012. Artstor. [https://library.artstor.org/#/asset/ARTONFILE\\_DB\\_10313670712](https://library.artstor.org/#/asset/ARTONFILE_DB_10313670712) .
- Herrin, Judith. "Opinion | Converting Hagia Sophia into a Mosque Is an Act of Cultural Cleansing." The Washington Post. WP Company, July 15, 2020. Last modified July 15, 2020. Accessed August 16, 2020. <https://www.washingtonpost.com/opinions/2020/07/15/converting-hagia-sophia-into-mosque-is-an-act-cultural-cleansing/>.
- Interior, Nave, Medallions with Inscriptions in Arabic.* April 2010. [https://library.artstor.org/#/asset/ASITESPHOTOIG\\_10313839717](https://library.artstor.org/#/asset/ASITESPHOTOIG_10313839717) .
- "Isidorus of Miletus." Complete Dictionary of Scientific Biography. . Encyclopedia.com. (August 11, 2020). <https://www.encyclopedia.com/science/dictionaries-thesauruses-pictures-and-press-releases/isidorus-miletus>.
- James, Liz. "Matters of Materiality in Byzantium. The Archangel Gabriel in Hagia Sophia, Constantinople." *Konsthistorisk Tidskrift/Journal of Art History* 86, no. 3 (July 3, 2017): 145–57. <https://doi.org/10.1080/00233609.2017.1330279>.
- Lessing, Erich. *Constantinople, Hagia Sophia, Virgin and Child.* Artstor. Accessed 2020. [https://library.artstor.org/#/asset/LESSING\\_ART\\_10311442169](https://library.artstor.org/#/asset/LESSING_ART_10311442169) .
- Lessing, Erich. *Constantinople, Hagia Sophia, Virgin and Child Enthroned between Emperors Justinian I, with the Model of Hagia Sophia [Left], and Constantine I, with the Model of Byzantium.* Artstor. Accessed 2020. [https://library.artstor.org/#/asset/LESSING\\_ART\\_10311441194](https://library.artstor.org/#/asset/LESSING_ART_10311441194) .
- Magal, Shmuel. *Interior, Nave, Apse, Archangel Gabriel Mosaic.* Artstor. Accessed 2020. [https://library.artstor.org/#/asset/ASITESPHOTOIG\\_10313839655](https://library.artstor.org/#/asset/ASITESPHOTOIG_10313839655).
- Magal, Shmuel. *Sultan Ahmed Mosque / Sultanahmet Camii / "The Blue Mosque"*. April 2010. Artstor. [https://library.artstor.org/#/asset/ASITESPHOTOIG\\_10313840166](https://library.artstor.org/#/asset/ASITESPHOTOIG_10313840166).
- Mango, Cyril, and Ernest J. W. Hawkins. "The Apse Mosaics of St. Sophia at Istanbul. Report on Work Carried out in 1964." *Dumbarton Oaks Papers* 19 (1965): 113–51. <https://doi.org/10.2307/1291228>.
- Mark, Robert, A. S. Cakmak, and Gülru Necipoğlu. "The Life of an Imperial Monument: Hagia Sophia after the Byzantium." Essay. In *Hagia Sophia from the Age of Justinian to the Present*, 195–225. Cambridge: Cambridge University Press, 1992.
- "Mustafa Kemal Atatürk." Biography.com. A&E Networks Television, June 22, 2020. Last modified June 22, 2020. Accessed August 14, 2020. <https://www.biography.com/political-figure/mustafa-kemal-ataturk>.
- Mellink, Machteld Johanna. *Hagia Sophia.* 1954. [https://library.artstor.org/#/asset/ABRMAWR\\_MELLINKIG\\_10310735547](https://library.artstor.org/#/asset/ABRMAWR_MELLINKIG_10310735547) .
- Mellink, Machteld Johanna. *Hagia Sophia.* 1956. Artstor. [https://library.artstor.org/#/asset/ABRMAWR\\_MELLINKIG\\_10310736414](https://library.artstor.org/#/asset/ABRMAWR_MELLINKIG_10310736414) .
- Merrit Kennedy and Peter Kenyon, "Turkey Converts Istanbul's Iconic Hagia Sophia Back Into A Mosque," *NPR* (NPR, July 10, 2020), last modified July 10, 2020, accessed August 22, 2020, <https://www.npr.org/2020/07/10/889691777/turkish-court-ruling-clears-way-for-hagia-sophia-to-be-converted-to-a-mosque>.
- Nelson, Robert S, Margaret Olin, and Robert S Nelson. "Tourists, Terrorists, and Metaphysical Theatre at Hagia Sophia." Essay. In *Monuments and Memory, Made and Unmade.* Chicago: The University of Chicago Press, 2003.

- Ousterhout, Robert. "The East, the West, and the Appropriation of the Past in Early Ottoman Architecture." *Gesta* 43, no. 2 (January 1, 2004): 165–76. <https://doi.org/10.2307/25067103>.
- Sebah, and Joaillier. *Interieur De La Mosquée De Ste. Sophie. Vue Generale - [Interior of Hagia Sophia, General View]*. Artstor. Accessed 2020. [https://library.artstor.org/public/SS7731291\\_7731291\\_10908773](https://library.artstor.org/public/SS7731291_7731291_10908773).
- Sebah, and Joaillier. *Vue Panoramique De La Mosquée Suleymanie - [Exterior View of Sulieman's Mosque]*. Artstor. Accessed 2020. [https://library.artstor.org/public/SS7731291\\_7731291\\_10908785](https://library.artstor.org/public/SS7731291_7731291_10908785).
- Selim Koru. "Turkey's Islamist Dream Finally Becomes a Reality." *The New York Times*, July 14, 2020. <https://www.nytimes.com/2020/07/14/opinion/hagia-sophia-turkey-mosque.html>.
- Sherry Blankenship, Sherry. "Cultural Considerations: Arabic Calligraphy and Latin Typography." *Design Issues* 19, no. 2 (March 2003): 60–63.
- Steven Zucker and Beth Harris, "Hagia Sophia as a Mosque (Video)." *Khan Academy*. Khan Academy, n.d. Accessed August 21, 2020. <https://www.khan-academy.org/humanities/ap-art-history/early-europe-and-colonial-america/medieval-europe-islamic-world/v/hagia-sophia-mosque>.
- Tanyeri-Erdemir, Tuğba. "Archaeology as a Source of National Pride in the Early Years of the Turkish Republic." *Journal of Field Archaeology* 31, no. 4 (2006): 381–93. Accessed September 4, 2020. <http://www.jstor.org/stable/40025296>.
- "The Divinely-Revealed Appearance of Angels in Icons." *A Reader's Guide to Orthodox Icons*. Last modified January 12, 2012. Accessed August 7, 2020. <https://iconreader.wordpress.com/2012/01/12/the-divinely-revealed-appearance-of-angels-in-icons/>.
- The Editors of Encyclopaedia Britannica. "Hagia Sophia." *Encyclopædia Britannica*. Encyclopædia Britannica, inc., August 5, 2020. Last modified August 5, 2020. Accessed August 9, 2020. <https://www.britannica.com/topic/Hagia-Sophia>.
- The Editors of Encyclopaedia Britannica, "Theotokos," *Encyclopædia Britannica* (Encyclopædia Britannica, inc., December 1, 2017), last modified December 1, 2017, accessed August 16, 2020, <https://www.britannica.com/topic/Theotokos>.
- TRT World. "President Erdogan's public address on Hagia Sophia", *Youtube* video, 33:30, July 13, 2020. <https://www.youtube.com/watch?v=ox-ROvVnPXU>.
- "Turkey Reconverts Istanbul's Hagia Sophia Museum into a Mosque." *News | Al Jazeera*. Al Jazeera, July 10, 2020. Last modified July 10, 2020. Accessed August 18, 2020. <https://www.aljazeera.com/news/2020/07/turkey-court-revokes-hagia-sophia-museum-status-200710131419431.html>.
- Wegner, Emma. "Hagia Sophia, 532–37," *The MET*, October 2004, [https://www.metmuseum.org/toah/hd/haso/hd\\_haso.htm](https://www.metmuseum.org/toah/hd/haso/hd_haso.htm).
- Whittemore, Thomas. "The Mosaics of St. Sophia at Istanbul." *American Journal of Archaeology* 42, no. 2 (1938): 219–26. <https://doi.org/10.2307/499667>.
- Zucker, Steven, and Beth Harris. "Theotokos Mosaic, Hagia Sophia, Istanbul." *Smarthistory*. Last modified December 15, 2015. Accessed August 15, 2020. <https://smarthistory.org/theotokos-mosaic-hagia-sophia-istanbul/>.



# **Exit, Voice, and Violence: Contemporary Responses of Non-Hindu Groups to the BJP's Hindu Nationalism**

---

*Rushil Roy*

**Author Background:** *Rushil Roy grew up in the United States and currently attends Monte Vista High School, in Danville, California in the United States. His Pioneer research concentration was in the field of political science and titled "Big Questions in Comparative Politics: Nationalism, Modernity, the Social Question."*

## **Abstract**

Since rising to power in the 1998 Indian General Election, the Bharatiya Janata Party (BJP) has been most prominently defined by their strict adherence to Hindu nationalist (Hindutva — or literally translated, "Hinduness"<sup>1</sup>) values and their attempts to spread such values through their public rhetoric and policy formation. This ruling ideology has been dubbed "Hindu nationalism" by the global community and has been consistently recognized as an effective strategy for winning over the Hindu majority population in India's elections. Policies generally associated with Hindu nationalist principles include the abrogation of Article 370 of the Indian Constitution (repealing Jammu and Kashmir's right to autonomy), restriction of Muslim immigration into the country, and a proactive "Zero Tolerance"<sup>2</sup> approach to religious terrorism which in practice materializes as a determination to frame the Muslim population of India as terrorists. The party has maintained political influence in India since the late 1970s and has been the dominant political power since the 1998 general election. These circumstances logically beg the question of how all variety of non-Hindu groups, including Muslims, Buddhists, Christians, and various folk religions have responded to the BJP's Hindu nationalism. While some scholars of the subject have used Hirschman's 1970 formulation of "exit,

---

<sup>1</sup> Brown, Garrett Wallace., Alistair McMillan, and Iain McLean. *The Concise Oxford Dictionary of Politics and International Relations*. Oxford: Oxford University Press, 2018, pp. 381–386.

<sup>2</sup> Staff, *Financial Express*. 2019. "PM Modi Calls For 'Zero Tolerance To Terrorism' In Meeting With EU Delegation Ahead Of Kashmir Visit". *The Financial Express*.

voice, and loyalty" to analyze minority group responses to oppression, I argue that the most prominent contemporary response to Hindu nationalism has emerged in a distinct form, political violence. Although political violence in India has existed for centuries (predating the BJP), the findings of this paper indicate the substantial rise of movements predicated on political violence as the dominant response to persecution was uniquely provoked by the BJP's actions. Essentially, as conventional responses to oppression lost their utility or simply became inaccessible under BJP rule, violence filled the vacuum of traditional liberal politics' foreclosure for minority religious groups. This paper will both investigate the root factors inherent to the BJP practice of Hindu nationalism which have forced this transition towards non-Hindu political violence as well as evaluate the empirical and historical efficacy of these non-conventional methods. Using data gathered from the Indian National Census, the Reserve Bank of India, and the Electoral Commission of India, I prove a quantitative link between the BJP's policy strategy and the foreclosure of traditional avenues of political expression for marginalized religious groups. Based on this causal link, this paper then further examines how this policy strategy has directly provoked non-conventional, violent responses from these groups and the utility such responses have served.

## 1. Introduction

Since its inception, the BJP has been a controversial player in Indian politics. Often defined by ideological adherence to principles of right-wing nationalism, their rise to power over the past three decades was largely the product of a party platform based explicitly on Hindutva preservation. Hindutva, a form of Hindu nationalism that has grown substantially in popularity over the past century, has its ideological roots in Vinayak Damodar Savarkar, an Indian independence activist who developed the philosophical framework and praxis strategy for Hindutva.<sup>3</sup> Savarkar's envisioning of religious relations between Hindus and non-Hindu groups under Hindutva consisted of a system similar in domestic ruling strategy to Nazism in 1930s Germany, as he argued that "Nazism provided undeniably the savior of Germany."<sup>4</sup> Hindutva philosophy was founded on the belief that 1930s India found itself in an analogous position to 1930s Germany, as Savarkar furthered, "to keep up the purity of the nation and its culture... Germany shocked the world by her purging the country of the semitic races - the Jews... a good lesson for us in Hindustan to learn and profit by."<sup>5</sup> This early modeling of a philosophy prioritizing cultural homogeneity, religious purity, and a marginalizing ruling structure came more than a half-century prior to the growth of the BJP from the late 1980s onward. However, the party's ideological origins grant some necessary insight into the motivations fueling BJP political leaders' current

---

<sup>3</sup> Savarkar, Vinayak Damodar. *Essentials of Hindutva* 1923.

<sup>4</sup> Savarkar, Vinayak Damodar. *Hindu Rashtra Darshan*. Poona: Maharashtra Prantik Hindusabha, 1949, p. 81.

<sup>5</sup> Marzia Casolari. "Hindutva's Foreign Tie-Up in the 1930s: Archival Evidence." *Economic and Political Weekly* 35, no. 4 (2000), pp. 218-28.

interactions with non-Hindu groups. In fact, the BJP in recent years has made a concerted effort to strengthen their perceived ties with Savarkar's ideas in the public arena, often outright defending his legacy as a "fight[er] for freedom"<sup>6</sup> against claims to the counter from political rivals. Despite this association, the BJP began very far from this position and moved to where it stands now, less from ideological commitment to Savarkar's ideas and admiration of Nazism, than in an effort to reach a constituency that supported these ideas.

The BJP was not initially founded on Hindutvist, theocratic or even nationalist principles. It was only following an immense defeat in the 1984 Indian General election where the party only succeeded in winning two of the contested 533 seats in the Lok Sabha Legislature, that the party "decided to abandon the ideals of 'positive secularism and Gandhian socialism' that it had adopted in its inception."<sup>7</sup> This electoral defeat prompted a substantial overhaul of the party platform, effectively transforming the BJP from a moderate alternative to the Indian National Congress' traditional secular nationalism and socialism to a more religiously extreme embodiment of those same principles.<sup>8</sup> This strategy proved politically successful as the next Indian general election following these reforms (1989) saw the BJP's influence in the Lok Sabha Legislature increase more than forty-fold and their support amongst the Indian populace increase by 46.77%.<sup>9</sup> Such political successes essentially set the course of the modern BJP, with common policy positions of the party platform thereafter notably consisting of rejection to other parties' "pseudo-secularism" and alleged "appeasement" of Muslim demands.<sup>10</sup> In order to best understand the impact of the BJP's evolution, I employ a quantitative approach to election policy and fiscal policy analysis which provides insight into how this transition has substantively altered the material conditions of non-Hindu life in India. This quantitative analysis is backed by an abundance of primary empirical evidence including Indian National Census population data, voter turnout records from the Electoral Commission of India, and social welfare spending reports.

## **2. Theory Framing: Hirschman's Theory of Valid Responses Under Oppression**

Albert Hirschman's theory of valid responses under oppression<sup>11</sup> plays a foundational role to this paper by providing insight into existing theory on common group responses to state declination (in the instance of India, state declination is perceived by marginalized non-Hindu populations). Hirschman's ideas have attracted substantial attention from scholars of India, specifically in the context of religious

---

<sup>6</sup> Correspondent, "Congress Says Savarkar Did Not Fight for Freedom; BJP Demands Apology," *The Hindu*, October 19, 2019.

<sup>7</sup> Misra, Satish. "Understanding the Rise of the Bharatiya Janata Party." Observer Research Foundation, September 21, 2018.

<sup>8</sup> Engineer, Asghar Ali. "The BJP & Its 20 Years." *The Hindu*. The Hindu, June 1, 2000.

<sup>9</sup> Election Commission Of India New Delhi. Statistical Report On General Elections, 1989 To The Ninth Lok Sabha. Vol. 1. 1990.

<sup>10</sup> Guha, Ramachandra. *India after Gandhi: the History of the World's Largest Democracy*. London: Picador, 2007. p. 635

<sup>11</sup> Hirschman, Albert O. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, Mass: Harvard University Press, 1970.

governing structures based on Hindu principles.<sup>12</sup> His model argues that there are three primary responses a population can pose upon their recognition of their state's decline. They can "exit:" emigrate away from the country in question; "voice:" self-advocate for their own interests through both conventional (e.g. local government autonomy) and non-conventional (e.g. protests) channels; or practice "loyalty:" capitulation to the demands of the state, typically after having their allegiance manufactured through means like welfare dependency or authoritarian indoctrination.

Hirschman's three-pronged framework for organized group response relies on the assumptions of group uniformity and autonomy, where oppressed groups have the cohesion to strategize and operate freely under the rule of their oppressor. In the context of minority responses to Hindu nationalism, the accuracy of these assumptions varies on a case-by-case basis, entirely depending on the group at-hand, their historical uniformity, and the degree to which (if at all) their liberties are being restricted by Hindu nationalist governance. Hirschman's conclusions will also prove incredibly useful in evaluating the comparative success of various groups' responses to the rise of Hindu nationalism, as he explains, "in practice upward mobility for a disadvantaged or hitherto oppressed group is likely to require a mixture of the individual and the group process, that is, a mixture of exit and voice."<sup>13</sup> Understanding the inner-workings of both the most chosen and most effective responses pursued under this framework will be a core facet of this paper's examination of non-Hindu minority responses to Hindu nationalism.

### **3. History of the BJP: Strategic Capitalization on Religious Conflict**

Reactive responses to religious tensions have been a defining trait of Indian politics for centuries. While it is inaccurate to assert that all of the nation's contemporary religious conflicts were birthed from the BJP's policy platform from the late 20th century, the increased prevalence of violent religious conflict between Hindus and non-Hindus during the rise and rule of the party is historically significant. The BJP's capitalization on these instances of religious conflict and their integration of reactive policies related to these events into their party platform indicates a causal link between the growth of violent religious dissension and the expansion of the BJP's Hindu nationalist policies.

The most significant instance of violent religious discord linked directly to the BJP occurred during the 1992 Ayodhya mosque attack. The city of Ayodhya, Uttar Pradesh is somewhat unique to India as it is one of the few global territories with stated religious significance to both Muslims and Hindus. Hindus believe it to be the birthplace of the God, Rama<sup>14</sup> while Muslims contend its cultural significance stems from the existence of the Babri Masjid Mosque (built circa 1528). The Babri Masjid Mosque laid at the center of the 1992 Hindu-Muslim conflict over the territory, as Ayodhya Hindus defended that the 16th-century mosque was built following the demolition of a Hindu temple and thus they retained the sole right to religious

---

<sup>12</sup> Spinner-Halev, "Hinduism, Christianity, and Liberal Religious Toleration," *Political Theory* 33, no. 1 (February 2005): pp. 35-57

<sup>13</sup> Cited in *ibid.* 13 p. 111

<sup>14</sup> Mehta, Deepak. "Naming the Deity, Naming the City: Rama and Ayodhya." *South Asia Multidisciplinary Academic Journal*, no. 12 (2015). p. 7

construction on the land. There is currently no historical or empirical evidence to support this argument, but it is still widely accepted by the Indian Hindu population.

<sup>15</sup>

The mosque was demolished by a civilian mob composed of an unknown portion of the 150,000 Hindu attendees at a BJP political rally.<sup>16</sup> For years after this event, the BJP has repeatedly attempted to distance themselves from any association with provoking the act while simultaneously defending the merits of the mob demolition. As recently as 2019, the BJP party manifesto re-defended the notion that the land is rightfully of Hindu origin and restated their goal of initiating construction of the Hindu Ram Mandir Temple in place of the now-demolished Babri Masjid Mosque.<sup>17</sup> The Indian government, which at the time of the demolition was controlled by the Indian National Congress party (INC), made a stated commitment to determining the parties at fault for this illegal mob demolition and established the Liberhan commission to formally investigate the matter. The commission only reached a final conclusion a full 17 years after the demolition, in June of 2009, in a report which formally established culpability and responsibility. The commission identified principal senior members of the BJP as wholly culpable for both the demolition of the mosque, as well as the deaths of an estimated 4,000 stemming from the Hindu-Muslim riots which characterized the weeks following the events — concluding that BJP leadership had both prior knowledge of the attacks as well as the means to prevent them yet deliberately chose not to, as such incendiary religious conflict was politically advantageous for the party.<sup>18</sup>

This report articulates an accurate prediction for future BJP action: in order to amass as much influence as possible for the Hindu nationalist cause, they must actively stoke religious conflict between Hindus and religious minorities — a trend which continually emerges under BJP leadership. Such conditions of constant religious conflict are clearly antithetical to the democratic will of the largely non-extremist Indian Hindu populace, but as the report further notes, “[BJP] leaders have allowed their actions to be dictated not by the voters but by a small group of individuals who have used them to implement agendas unsanctioned by the will of the common person.”<sup>19</sup> This 1992 mosque assault marks the moment when the BJP first explicitly deviated from their stated populist appeals to the “common man” to become the party of Hindu elites.

#### **4. The BJP and Social Welfare Programs**

Often, during the rise of similar religious nationalist movements such as the contemporary Sri Lankan Sinhalese Buddhist nationalist movement and the 1980s Lebanese Christian nationalist movement, scholars have been forced to expand their

---

<sup>15</sup> Udayakumar, S. P. “Historicizing Myth and Mythologizing History: The ‘Ram Temple’ Drama.” *Social Scientist* 25, no. 7/8 (1997): p. 11

<sup>16</sup> Tully, Mark. “South Asia | Tearing down the Babri Masjid.” BBC News. BBC, December 5, 2002.

<sup>17</sup> Bharatiya Janata Party. *Bharatiya Janata Party: Sankalp Patra Lok Sabha 2019*. New Delhi: Bharatiya Janata Party, 2019. p. 36

<sup>18</sup> Liberhan Ayodhya Commission of Inquiry. *Report of the Liberhan Ayodhya Commission of Inquiry, 2009*. pp. 942-3

<sup>19</sup> *Ibid.*



analysis beyond the state's explicit policies to adequately understand the popularity and complexity of such ideologies. This is because such movements often amass popular support through non-traditional avenues; to focus solely on the material policy and legislative action taken to implement these theonationalist ideas would obscure the equally important role of ideologues' rhetoric, philosophy, and the perspectives they appealed to. The BJP's pursuit of Hindutva exemplifies an incredibly efficient integration of these two goals into a single party platform. The philosophical, ethos-based, popular appeals to Indian Hindus were well established by the religious conflicts which had characterized the past decades. Even before the BJP became the national hegemon it is today, the period immediately preceding their growth in popularity from 1990 to 1993 saw an average of nearly 10 instances of mass sectarian violence between Hindus and Muslims annually under the leadership on the Indian National Congress party, with death tolls of Muslims in such conflicts doubling the death tolls suffered by Hindus in some cases.<sup>20</sup> The foundation for a philosophy promoting cultural conflict had already been established; what set the BJP apart was their affinity for utilizing explicit legislative action to promote Hindutva ideas instead of solely relying on more implicit popular rhetorical appeals.

Nowhere is this execution of Hindutva proposals through legislative action more evident than in BJP welfare policy. Welfare programs in India have deep roots in the economic architecture of the country. The existence of the Indian caste system as a social class structure prior to its abolition by India's constitution made welfare programs a necessity to ensure that life in post-colonial India during its transition to an industrial economy would be sustainable for those coming from lower castes such as the Dalits (formally referred to as "Untouchables") and Shudras. Even as the lasting effects of the Indian caste system have gradually lessened over time, the importance of the social welfare programs established following the elimination of the caste system is the same, if not greater, than it was decades ago during its establishment. Current estimates put the Indian welfare's state coverage at over 100 million low-income households throughout the nation.<sup>21</sup> Indian citizens from all assortment of ethnic and religious backgrounds have a structural dependency on the welfare state, making it a prime tool for strategic political exploitation. If a party truly holds the end goal of creating a culturally homogeneous ethnostate, rid of the non "religiously pure,"<sup>22</sup> while simultaneously exercising complete authority over the distribution of resources to the nation's poor, then that party has the power to create the conditions for their ideal Hindutva society.

Of course, the BJP cannot simply establish a stated policy of solely granting welfare benefits to the "religiously pure" Hindu population. If they were to do so, they run the likely risk of alienating the plurality of their Western allies who have become increasingly critical of the BJP's promotion of religious division.<sup>23</sup> As a result, the party's use of social welfare systems for strategic political gain evolved to be more subtle, focusing largely on gradually depriving welfare benefits from geographic regions occupied disproportionately by non-Hindu groups, in order to maintain the veneer of non-religiously motivated welfare policy. The BJP has been able to successfully capitalize on their control of the welfare system to ensure that areas in

---

<sup>20</sup> Brass, Paul R. *Production of Hindu-Muslim Violence in Contemporary India*. Seattle; London: Univ Of Washington Press, 2015.

<sup>21</sup> Tikkanen, Roosa. "India," Commonwealth Fund (Commonwealth Fund, June 5, 2020).

<sup>22</sup> Roy, Abhik. "Anatomy of Fascism," *The Statesman*, November 16, 2019.

<sup>23</sup> PTI, "US Report on Global Religious Freedom Shows Bias against PM Modi: BJP," *The Economic Times* (Economic Times, June 22, 2019).

India with a high Muslim population suffer the conditions which make quality of life lower, economic declination accelerate, poverty more widespread, and crime more common. All of these factors culminate to reinforce the commonly proposed, yet factually untrue, BJP argument that the Muslim population in India is solely responsible for much of Indian society's ills such as rising crime rates.<sup>24</sup> By tactically propagating the conditions which lead to poverty and crime in areas with high Muslim populations, the BJP is essentially able to create truth to their own accusations while simultaneously giving India's Hindu population a reason to concur with Hindutva philosophy.

### **The Successes of the BJP's Social Welfare Strategy**

The BJP's weaponization of welfare policy to strategically marginalize non-Hindu groups was most clearly displayed during the 2019 Indian general election. The BJP had maintained a position of absolute political dominance for the five years following the 2014 general election, in which they achieved key victories including both the Prime Ministership, with candidate Narendra Modi, and India's lower house of parliament where they became the first party in over 30 years to achieve a majority.<sup>25</sup> However, by 2019, the party's near-universal public support was beginning to waver. The 2019 general election was the first instance following the party's 2014 victories in which the BJP seriously risked losing substantial amounts of political power. The culmination of several economic factors including steadily growing unemployment, an increased consumer tax rate, and the BJP "demonetization" policy in which 90% of currency notes were removed from circulation<sup>26</sup>, resulted in the party suffering a defeat in the 2018 general assembly elections just four months before the ever-crucial 2019 general election.<sup>27</sup> This looming threat to the BJP's power made it such that their weaponization of social welfare programs had to extend further than it had previously, with the hopes of either revitalizing the country's seemingly dulled Hindu nationalist sentiments or manufacturing the conditions which decreased the likelihood of non-Hindu groups turning out and voting in the first place.

Primarily, the BJP elected to focus on the latter option. Alterations in welfare policy immediately leading up to the 2019 election strongly indicate an interest in isolating non-Hindu groups away from the economic benefits associated with welfare programs. In order to more closely examine the effects of BJP policy, I use Indian National Census records, regional voter turnout data from the Electoral Commission of India, and spending records from the Reserve Bank of India in order to establish a quantitative link between BJP policies and adverse electoral effects on non-Hindus.

As of the most recent Census data, six of the 28 Indian states are composed of a non-Hindu majority.<sup>28</sup> Despite the Hindu population in these states ranging from

---

<sup>24</sup> PTI, "BJP MP Blames Muslims for Rising Crimes: India News - Times of India," The Times of India (TOI, July 27, 2018).

<sup>25</sup> Vaishnav, Milan, Jayaram Ravi, and Jamie Hinton. Publication. *Is The BJP India's New Hegemon?* Carnegie Endowment for International Peace, 2018, p. 2.

<sup>26</sup> Dhume, Sadanand. "Modi's Three Big Gambles: We'll Soon Know If Demonetisation, GST and Adityanath Damaged BJP's Re-Election Prospects." Times of India Blog, June 2, 2018.

<sup>27</sup> Ayres, Alyssa. "The BJP No Longer Looks Invincible in 2019 Indian Elections." Time. Time, December 12, 2018.

<sup>28</sup> Government of India, Office of the Registrar General & Census Commissioner, 2011 Census Records, "Population By Religious Community," Population By Religious Community § (2011).

as low as 2.78% to 41.79%, the months leading up to the 2019 election saw the BJP in full control of the state governments of all but one (Punjab<sup>29</sup>) allowing the party full authority to control access to and use of federal welfare funds. The BJP's control of both the federal government — tasked with allocating social welfare grants to individual states — as well as the state governments which controlled how those welfare grants were spent, allowed the party to coordinate a tactical approach to welfare planning in the fiscal year leading up to the 2019 election. Through their multi-level government control, the BJP-controlled federal government could allocate substantial amounts of welfare funding to the non-Hindu states so as to make their welfare policy appear non-discriminatory, while simultaneously coordinating with non-Hindu states' BJP governments to withhold or mispend the allocated federal aid, effectively ensuring non-Hindu regions suffered disproportionate economic challenges. Publicly released data quantifying each states' ratio of social spending expenditure to federal funding disbursement provides substantial empirical support for this claim. During the 2018-2019 fiscal year, the average percentage of federal funding disbursement which materialized as intended in spending for state welfare programs averaged 41.9% across all Indian states and territories (Hindu and non-Hindu).<sup>30</sup> In non-Hindu states and territories, welfare spending of government disbursement dipped as low as nearly half of this national average in certain instances and averaged 14.9% less overall than the national average. Continued comparative analysis between the six Hindu minority states and the 10 Hindu majority, non-BJP controlled states further supports this trend, as the social spending discrepancy widens to 15.5%.<sup>31</sup> This indicates that welfare fulfillment on this front transcends party lines and political affiliation; the only shared causal factor in these instances is religious majority. Even in regions in which the BJP retains non-controlling state or local-level influence, the social spending disbursement granted by the BJP federal government allows them to successfully fund social welfare initiatives more efficiently by several orders of magnitude compared to states with BJP influence and a non-Hindu majority.<sup>32</sup> The below graphs which compare the analyzed data from the Reserve Bank of India display the extent to which this discrepancy occurred.

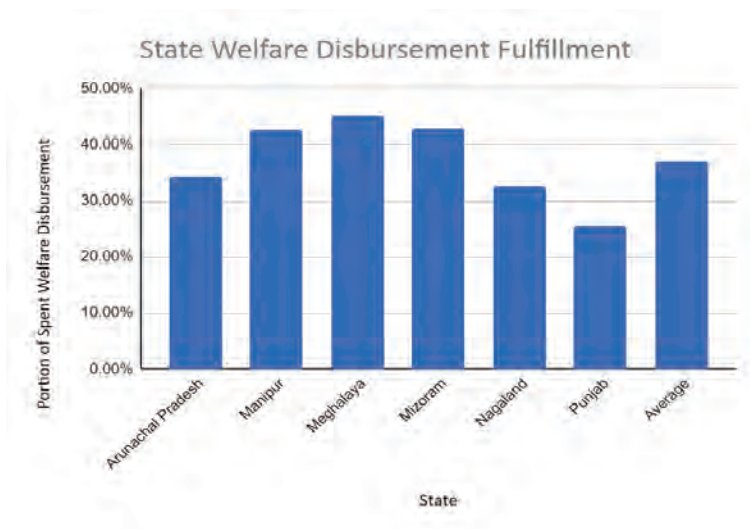
---

<sup>29</sup> IT Staff, "Maharashtra Setback: As Another State Slips out of BJP's Hand, Saffron Hues Shrink from India's Map," *India Today*, November 26, 2019.

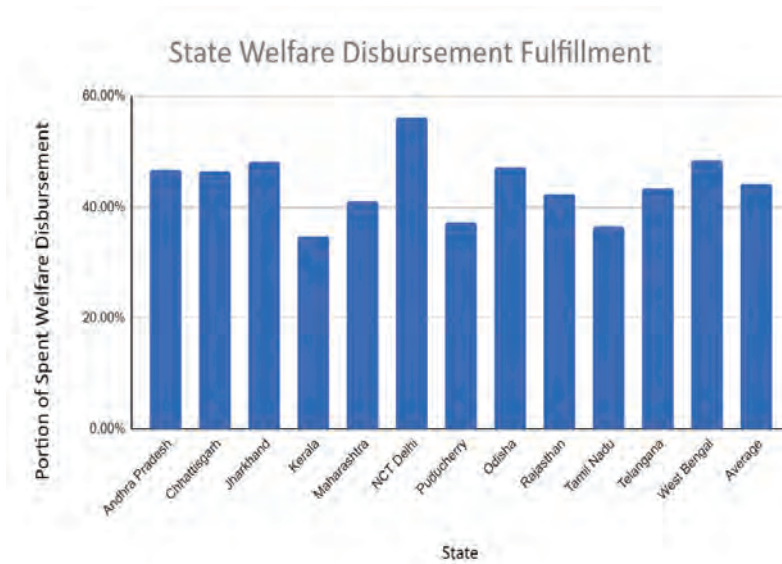
<sup>30</sup> Government of India, Reserve Bank of India, "State Finances : A Study of Budgets," *State Finances : A Study of Budgets* § (2019), p. 177.

<sup>31</sup> *Ibid.*

<sup>32</sup> *Ibid.*



**Figure 1.** Social spending fulfillment of non-Hindu majority states during the 2018-2019 FY<sup>33</sup>



**Figure 2.** Social spending fulfillment of Hindu majority non-BJP controlled states during the 2018-2019 FY<sup>34</sup>

This welfare discrepancy retains even more significance when juxtaposed with the strategy's impact on voter turnout in the 2019 general election. By selectively starving out the regions with religious demographics most likely to oppose the BJP's

<sup>33</sup> Ibid.

<sup>34</sup> Ibid.

Hindu nationalist philosophy and policies, the party was able to electorally disenfranchise these same regions. Despite an overall national average increase in voter turnout of 1.16 percentage points, non-Hindu states disproportionately suffered a decrease in voter turnout as a result of the successful BJP welfare strategy.<sup>35</sup> Similar analysis used to identify the BJP's welfare strategy itself can also be applied to isolate its effects on voter turnout. Comparing the difference in voter turnout between the previous 2014 Indian general election and the 2019 election which saw usage of the welfare strategy, a noteworthy 43.9% discrepancy emerges between the change in voter turnout in Hindu states which did not undergo substantial changes in their welfare budgets and the BJP controlled, non-Hindu states which did.<sup>36</sup> In essence, the BJP was able to engineer a policy-based strategy to implicitly suppress votes from the regions which represented majority non-Hindu demographics. Moreover, this plan proved incredibly successful in impacting the election as a whole as the regions which contained non-Hindu majorities saw less than half the average growth in voter base compared to Hindu majority states. This practice serves as conclusive indication of the multi-layered Hindutva practice scholars theorized would arise under BJP rule.<sup>37</sup> No longer was political Hindu nationalism's adherence to Hindutva ideology defined simply by vitriolic rhetoric and philosophy but now, as of the 2019 general election, by structural political violence against the "undesirable"<sup>38</sup> non-Hindu population.

Alternative explanations provided for this phenomenon have speculated that increased corruption or mismanagement of funds coincidentally befalling non-Hindu regions to a greater extent than their Hindu counterparts. This explanation seems relatively unlikely, most notably because these non-Hindu, non-BJP controlled states were subject to the exact same corruption regulations and spending audits as all other states, as the authority to conduct such investigations exists on the federal level, not the state level. Since all states were held to the same regulatory standard in this regard, the probability of a coincidental corruption disparity (as opposed to a strategic funding-based one) is incredibly low.

## 5. Non-Hindu Groups' Responses to the Rise of Hindu Nationalism

The BJP interpretation of Hindutva principles — a hybrid of the rhetoric and ideological beliefs supporting a culturally homogenous religious ethnostate coupled with the strategically prejudiced policy necessary to materialize it — has been dubbed "hyper-Hindutva."<sup>39</sup> Since the BJP is the first modern political party in India to successfully gain power operating under explicitly hyper-Hindutvist philosophy, their rise to power has forced non-Hindu groups to respond accordingly. New survival

---

<sup>35</sup> Government of India, National Informatics Centre, Parliamentary Constituency wise Turnout for General Election - 2014 § (2014).

<sup>36</sup> Ibid.

<sup>37</sup> Rauf, Taha A. "Violence Inflicted on Muslims: Direct, Cultural and Structural." *Economic and Political Weekly* 46, no. 23 (June 4, 2011): pp. 69–76.

<sup>38</sup> Mannathukkaren, Nissim. "Why the Bharatiya Janata Party Is Not Truly Bharatiya." *The Wire*, April 3, 2019.

<sup>39</sup> Ghose, Sagarika. "2017 Minus 2002: It's Now Hindutva plus in Gujarat." *Times of India*, December 3, 2017.

mechanisms had to be adopted by India's Muslims, Sikhs, Buddhists, and Christians in order to make life under the BJP's religiously oppressive rule sustainable. Whereas the "regular" Hindutva philosophy which characterized India's government prior to the late twentieth century was largely compatible with Hirschman's conventional theory of valid responses under oppression (exit, voice, and loyalty), the BJP specific hyper-Hindutva necessitated the creation of new, equally valid responses to oppression. While the general underlying assumptions of Hirschman's foundational theory remain valid, their basis in the politics of fifty years ago must evolve along with political trends. With this recognition, a new valid response under oppression emerges: political violence.

### **The Muslim Response: History and Empirical Evaluation**

The Muslim response to Hindu nationalism has evolved significantly in the past century. As the largest non-Hindu group in the nation, comprising 13.4% (138,188,240) of the total population,<sup>40</sup> the group has often found itself at the center of oppressive efforts by both contemporary and traditional Hindu nationalism. Worsening matters are the historically turbulent religious relations between Hindus and Muslims in the larger Eastern-Asian region as a whole leading to "substantial support from both conservative and moderate Hindus"<sup>41</sup> for extreme Hindu nationalist approaches to religious issues. As a result of this historical tumultuousness defining Hindu-Islamic relations, instances of widespread religious violence between Hindus and Muslims have been a constant, even during the transition between Hindutvist and hyper-Hindutvist times. In the 36-year time period from 1961 to 1997 which can be generalized as the "transitional period" to hyper-Hindutva as the BJP first gained widespread approval among the Indian populace, each year saw an average of 1.92 instances of mass communal religious violence between Hindus and Muslims.<sup>42</sup> The vast majority (95.6%) of these instances occurred with a Hindu majority heading the conflict against a Muslim minority. This historical imbalance disproportionately saw India's Muslim minority playing the role of the "victims of the mob"<sup>43</sup> with the Hindu majority as the aggressors. It is this historical constant which forced the deviation from conventional schools of political thought surrounding responses to oppression (such as Hirschmann's model) in favor of outright political violence.

Muslim political violence as a contemporary response to Hindu nationalism is most clearly displayed in the Union Territory of Jammu and Kashmir. The region has the largest Muslim population of any Indian territory with 66.97% of the 10,144,700 residing there being of Islamic faith compared to the just 29.63% Hindu population.<sup>44</sup> This region serves as an effective microcosm to understand militant Muslim mobilization throughout India. This is because the territory has birthed the

---

<sup>40</sup> Government of India, Office of the Registrar General & Census Commissioner, 2001 Census Records, Census And You / Religion § (2001).

<sup>41</sup> Aoun, Joy. Issue brief. Religious Movements, Militancy, and Conflict in South Asia Cases from India, Pakistan, and Afghanistan. A Report of the CSIS Program on Crisis, Conflict, and Cooperation. Center for Strategic and International Studies, 2012.

<sup>42</sup> Graff, Violette. "Hindu-Muslim Communal Riots in India I (1947-1986)." *Sciences Po*, August 20, 2013.

<sup>43</sup> Bagger, Matthew. "Review: Religion and Democracy in India." *Journal of International Affairs* 61, no. 1 (2007): p. 249

<sup>44</sup> State Government of Jammu & Kashmir, Demographics § (2012).

most prominent and active Muslim militancy movement in the country, granting scholars reasonable insight as to how various other developing Muslim militancy movements will progress towards their end goal of freeing India's Muslim population from Hindutva oppression. The Kashmiri Muslim militant movement has its roots in guerilla war tactics, often operating outside the bounds of the rules and norms which dictate traditional conflicts. Attacks extending beyond the battlefield, often targeting both civilians and political officials<sup>45</sup> are common and are posed by the militant movement as a "valid response to the structural and targeted violence perpetrated by the state against the communities that the insurgency seeks to represent."<sup>46</sup> Examples of these guerilla strategies yielding success on the civilian front date back to the 1990 exodus of 250,000 Hindus from the territory of Jammu and Kashmir, just one year after the BJP began their rise to prominence on the national level. This exodus came at the end of a carefully planned and structured, aggressive campaign of unrestricted criminal violence against the Kashmiri Hindu population. The campaign often relied on extrajudicially cruel measures as a means of forcing the displacement of the Hindu population including murder, torture, and rape. This marks one of the first periods in contemporary Indian history where the Muslim population overtook the role of the aggressor. The fact that it yielded the intended material societal impact of lessening the Hindu presence in the region — specifically via the forced displacement of a quarter-million Hindu civilians — caused many members of the insurgency movement to perceive the campaign as a success despite the reprehensible moral implications of the means by which they achieved their intended end.

Just as the BJP's brand of hyper-Hindutva religious nationalism has evolved a great deal since its origins in the early 1990s, the Kashmir militancy movement has done the same. The movement has formed a more concrete identity for itself which has resultantly inspired several other similar Muslim insurgencies in India dedicated to the overthrowal of Hindutva.<sup>47</sup> Part of this formation of a concrete identity has involved the insurgency's establishment of alliances with several militant Islamic organizations including Al-Qaeda, the Taliban, and certain dedicated divisions of ISIL/ISIS. All of these allied entities share both recognition from the global community as Islamic fundamentalist terrorist organizations as well as a stated emphasis on *jihad* — an armed Islamic cultural and religious struggle for freedom<sup>48</sup> — in this context referring to the Muslim struggle for freedom from Hindutva rule. The Kashmiri insurgency seemingly shares this Jihadist philosophy in their pursuit of a free Kashmir absent of non-Muslim influences yet interestingly, these alliances with Islamic fundamentalist terrorist organizations have seemingly strayed from traditional conceptions of both Kashmiri and non-Kashmiri Jihad and their characteristic indiscriminate violence against non-Muslims. Following Al-Qaeda's 2014 establishment of a branch in India dedicated to militarily aiding the Kashmiri insurgency (Al-Qaeda in the Indian Subcontinent or AQIS), the strategies of cruel uninhibited violence against Hindus which had historically provoked desired action such as the 1990 Hindu exodus, were abandoned by those fighting for Kashmir's independence from Hindutva rule. The official AQIS code of conduct explicitly

---

<sup>45</sup> Ganguly, Šumit. "Explaining the Kashmir Insurgency: Political Mobilization and Institutional Decay." *International Security* 21, no. 2 (1996): pp. 76–107.

<sup>46</sup> Jain, Vamika. "Lal Salaam: Explaining the Violence and Longevity of the Naxalite-Maoist Insurgency in Three Frames." *The Journal of Contemporary Asian Studies*, August 5, 2020.

<sup>47</sup> Prasad, Hari. "The Salafi-Jihadist Reaction to Hindu Nationalism," *Hudson Institute* (Hudson Institute, July 11, 2019).

<sup>48</sup> Esposito, John L. Rep. *Jihad: Holy or Unholy War?* . United Nations Alliance of Civilizations, 2006.

prohibits violence targeted at Hindu, Christian, or Buddhist civilians and instead argues violent acts should be targeted at the Hindu nationalists who propagated “the policy of oppressing the Muslims in Kashmir and India, destroying their homes, weakening them in social status, and coercing them to convert to Hinduism.”<sup>49</sup> The Kashmiri militant movement’s increased organization and structure which came with the formal establishment of alliances between 2014 and 2019 has produced a more effective and more focused strategy which is proving itself increasingly efficient at responding to the state’s growing opposition to the movement. As with all valid responses to oppression, political violence adapts to combat the conditions of oppression it operates within. For the Kashmiri Muslims continuing their armed rejection of Hindu nationalist ruling and Hindutva, the indiscriminate political violence of the 1990s is no longer a sufficient means of achieving their end goal. Although an effective way of removing the terrorized civilian population, when faced with the armed power of the Indian state, indiscriminate violence does nothing. In order to effectively achieve the militancy’s new goal of ridding the region of the state’s Hindutva oppression, new means using the same medium – political violence – must be adopted.

Political violence, despite its evidently devastating humanitarian impact, was seen as a necessary measure to achieve Muslim autonomy at the dawn of the 1990s. Hirschmann’s model maintains relevance in explaining the Muslim pursuit of “voice” in Jammu and Kashmir for decades preceding their violent insurgency. Throughout the 1980s, Jammu and Kashmir were consistently defined by their rejection of mainstream national political parties in favor of regionalist parties which purported to solely represent the interests of the Kashmiri population. For example, the 1983 Jammu and Kashmir Legislative Assembly election had a voter turnout rate of 73%, more than 1.6 times higher than the most recent general assembly election, with 47 of the territory’s 75 parliamentary seats (62.67%) going to either the Jammu & Kashmir National Conference or the Jammu & Kashmir Peoples Conference.<sup>50</sup> However, despite the fact that it seems that Jammu and Kashmir aptly followed Hirschmann’s theory of “voice” by expressing political autonomy through regional self-advocacy in government, this was evidently not sustainable. The 1987 Jammu and Kashmir Legislative Assembly election was plagued with allegations of corruption, ballot manipulation, and election rigging by the Indian federal government.<sup>51</sup> The election saw the continuation of the region’s trend of high voter turnout with the region of Kashmir topping 80% voter participation.<sup>52</sup> However, despite the increased voter turnout and the targeted mobilization of Muslim voters by the Muslim United Front, the elections resulted in a significant decrease in representation of the Kashmiri regionalist parties. Regionalist parties only attained 40 of the territory’s 76 seats while the Indian National Congress, BJP, and various non-regionalist independents secured the remaining seats. Immediately, this birthed theories of election rigging on the part of the Indian National Congress, which at the time, was the nation’s dominant political party. Although these allegations of electoral

---

<sup>49</sup> Rep. “Al-Qaeda in the Indian Subcontinent (AQIS)” (Counter Extremism Project, 2020).

<sup>50</sup> Government of India, Election Commission Of India New Delhi. Statistical Report On Jammu & Kashmir Parliamentary Elections. Vol. 1. 1983.

<sup>51</sup> Akhtar, Shaheen. “Reviving the Elections Farce in Jammu and Kashmir.” *Strategic Studies* 18, no. 2/3 (1996): pp. 39–56.

<sup>52</sup> Bose, Sumantra. *Kashmir: Roots of Conflict, Paths to Peace*. Cambridge, MA: Harvard University Press, 2005. p. 49



malpractice weren't confirmed until 15 years later in 2002 when INC party leader, Khem Lata Wukhloo, admitted "there was a massive rigging in 1987 elections,"<sup>53</sup> the Kashmiri people immediately recognized they had been stripped of their democratic representation — their voice. Adherence to Hirschman's model could no longer assist India's Muslim community; they were soon forced to raise the question of what comes next when the people are robbed of their voice and violence becomes the logical answer. If attempts to operate within the system are rejected and suppressed by the state, then the only remaining option is rejection of the state's authority itself.

### **The Influence of Muslim Violence as a Response to Oppression on Other Non-Hindu Groups**

The most important indication of political violence's significance has been its acceptance by persecuted groups as a valid response to oppression. Although Islamic groups in India were some of the first to prove the response's efficacy and use it as a consistent means to materialize a political goal, many other groups have since recognized this and adopted the strategy in their own rejection of Hindu nationalism. Each of these non-Hindu groups' usage of political violence has served as a more contemporary evolution away from conventional means of combating persecution after they gradually became less viable.

Sikhs in India have undergone a paradoxical social development over the past decades. Although the Indian government has officially recognized them as a Hindu subgroup (a label the community rejects) they have still been granted religious minority "protected status" by federal law.<sup>54</sup> Many scholars of religious political violence have argued that Sikhs were the first group to respond to perceived Hindu nationalism with political violence considering the 1984 assassination of Prime Minister Indira Gandhi by two of her Sikh bodyguards for religiously motivated political reasons.<sup>55</sup> Although this is certainly one of the earliest instances of political violence in India as a response to oppression under Hindu nationalism, it would be inaccurate to characterize this as a larger overarching group response instead of as an independent action taken on by a distinct minority of the Sikh groups. This initial aversion to political violence as a means of countering persecution by the Sikh community was displayed in the immediate aftermath of PM Gandhi's assassination. During the weeks following the assassination, riots organized by Hindu nationalists resulted in over 17,000 Sikhs being "murdered, burned, raped, and tortured... all over India" yet there exists no evidence of Sikh's ever countering the riots with their own violent response.<sup>56</sup> Political violence only emerged as a prevalent Sikh strategy following the suppression of their "voice" (similar to what occurred following Jammu and Kashmir's 1987 election), which occurred when the regional party's representation in Punjab Parliament dropped by 95.9%, from 73 seats to 3, in a single

---

<sup>53</sup> Hussain, Altaf. "South Asia | Kashmir's Flawed Elections." BBC News. BBC, September 14, 2002.

<sup>54</sup> Bureau of Democracy, Human Rights, and Labor, International Religious Freedom Report for 2018 § (2019).

<sup>55</sup> Mahmood, Cynthia. "Sikh Rebellion and the Hindu Concept of Order." *Asian Survey* 29, no. 3 (March 1999): pp. 326–40.

<sup>56</sup> Rana, Maanushi. "1984 Sikh Genocide: A Forgotten Chapter?" *Warpaint Journal*, July 2020.

election in 1992.<sup>57</sup> Similar to the response of Muslims in Kashmir after being robbed of their voice, the natural evolution beyond voice was violence. In the decades following the 1992 election, the Khalistan Liberation Force (the primary Sikh militant group) resorted to several instances of targeted extrajudicial violence against Hindu nationalist political leaders including members of the BJP-affiliated RSS and the BJP itself.<sup>58</sup> As noted, the contemporary Sikh militancy movement is still in its infancy, meaning many of its most impactful acts of political violence have been incredibly recent, with the majority of the most prominent examples being concentrated in the past 5 years. This means it's difficult to foresee the material effects such violence will have on the political conditions of Sikhs in India. However, considering that the movement is based on similar motives and strategies to the notably successful Muslim political violence campaign, it is likely the movement will find at least some degree of success off its existing momentum.

Sikh and Muslim political violence movements are by no means the only ones to emerge under the BJP's rule of hyper-Hindutva, but they are undoubtedly the most well-developed and prevalent ones. All variety of non-Hindu groups in the region, notably Christians, Buddhists, and various folk religions have endured substantial persecution which makes conventional responses to oppression, such as "exit" and "voice," non-viable options. However, each of these groups evolve over a unique timeframe. Although Muslim militants in Kashmir mobilized almost immediately following the deprivation of their voice in the 1987 parliamentary elections, Sikhs took much longer since they were a significantly smaller community who suffered a remarkable amount of public disdain from the Indian population. Other non-Hindu groups who are yet to mobilize are operating under similar conditions to the Sikhs of the 1980s and 1990s, with public rejection to these groups at all-time highs under the Hindu nationalism of the BJP and active political efforts being made by the party to sink the prevalence and impact of their "voice". However, if the examples of Muslim and Sikh militancy indicate anything for the future of these groups, it's that political violence may become a necessary, inevitable, survival strategy.

## 6. Conclusion

The BJP's trademark brand of Hindu nationalism can be aptly summarized as a new spin on an old tale. With Hindutva's philosophical origins dating back to the early 20th century, decades before the BJP's creation, non-Hindu groups have been forced to adapt throughout India's history to the seemingly inherent dominance of Hindu nationalist ideology in the political sphere. Under the BJP, this dominance of religious nationalism became even more prevalent and oppressive with the advent of hyper-Hindutva which consists of a strict adherence to both the long-standing rhetorical and philosophical elements of traditional Hindutva coupled with a contemporary policy strategy formulated with the explicit intent of marginalizing non-Hindus. The BJP expertly utilized the hyper-Hindutva strategy on programs such as social welfare to ensure that the party's power, as derived from direct appeals

---

<sup>57</sup> "Punjab Assembly Election 2017 Results," Elections in India (Elections in India Archive, 2017)

<sup>58</sup> Hindustan Times, "Unidentified Men Fire Shots at RSS Shakha in Ludhiana," Hindustan Times, January 18, 2016.;

to the Hindu voter base's desire for a culturally homogenous society, was never questioned by the non-Hindus who would suffer under such a state. As a result, even social welfare programs became a tool in the BJP's arsenal of methods to make non-Hindu life in India unsustainable by making any conventional avenues to promote change or respond to persecution completely inaccessible. These instances of increased state aggression against non-Hindu groups under the BJP has forced non-Hindu groups to adapt their responses to persecution accordingly. Thus, India has witnessed the contemporary emergence of political violence as an effective counter to oppression.

The findings of this research are significant on two key fronts. Primarily it provides an empirically and historically based framework for non-conventional responses to oppression. Hirschmann theorized half a century ago that the three "conventional" responses of exit, voice, and loyalty each had their circumstantial merits dependent on factors like group cohesion, autonomy, and ability. These factors evidently matter a great deal in the context of normal politics where democratic participation and self-advocacy matter and are not obscured by electoral fraud and politically motivated corruption. However, today the question has arisen of what comes next upon the failure of conventional responses. When the state deprives a group of the voice it uses to self-advocate for minority regional interests, they clearly can no longer operate within the bounds of the conventional response framing and must instead opt to deviate to means the state cannot suppress as easily. The findings of this paper strongly indicate that the most accessible and effective non-conventional response to oppression following the deprivation of voice is political violence. Political violence inherently operates outside of the state's authority by rejecting the state's claim to a monopoly on the legitimate use of force and is thus much more difficult for the state to completely suppress. Unlike the voice that Kashmiri Muslims and Punjabi Sikhs were stripped of prior to crucial regional elections, an armed militancy movement engages in a different medium where the state is forced to confront the movement head-on rather than simply silencing them. Secondly, this contemporary framework can be used to paint a tangible picture of how future action on the part of non-Hindu groups in India will materialize. Although as of now, the evolution to political violence has only undergone sizable development by Sikhs and Muslims, Hindu nationalism has made it such that greater conditions of oppression are being suffered across every non-Hindu group indiscriminately. The provided analysis regarding welfare and its strategic use to manipulate elections under the BJP suggests Hindu nationalist aggression is no longer confined to individual regional or cultural disputes but now attempts to subjugate all non-Hindus universally. If this remains true under BJP rule going forward, the vast majority of existing empirical and historical evidence suggests that the future will see Indian non-Hindu groups of all varieties — Christians, Buddhists, and followers of folk religions alike — following in the footsteps of Muslims and Sikhs in a revived, now-armed, attempt to survive the BJP's campaign of oppression.

### Works Cited

- Akhtar, Shaheen. "Reviving the Elections Farce in Jammu and Kashmir." *Strategic Studies* 18, no. 2/3 (1996): 39–56.
- Aoun, Joy. Issue brief. *Religious Movements, Militancy, and Conflict in South Asia Cases from India, Pakistan, and Afghanistan*. A Report of the CSIS Program on Crisis, Conflict, and Cooperation. Center for Strategic and International Studies, 2012.

- Ayres, Alyssa. "The BJP No Longer Looks Invincible in 2019 Indian Elections." *Time*. Time, December 12, 2018. <https://time.com/5477783/bjp-state-loss-elections-narendra-modi/>.
- Bagger, Matthew. "Review: Religion and Democracy in India." *Journal of International Affairs* 61, no. 1 (2007): 249.
- Bharatiya Janata Party. *Bharatiya Janata Party: Sankalp Patra Lok Sabha 2019*. New Delhi: Bharatiya Janata Party, 2019.
- Bose, Sumantra. *Kashmir: Roots of Conflict, Paths to Peace*. Cambridge, MA: Harvard University Press, 2005. 49.
- Brass, Paul R. *Production of Hindu-Muslim Violence in Contemporary India*. Seattle; London: Univ Of Washington Press, 2015.
- Brown, Garrett Wallace., Alistair McMillan, and Iain McLean. *The Concise Oxford Dictionary of Politics and International Relations*. Oxford: Oxford University Press, 2018.
- Bureau of Democracy, Human Rights, and Labor, International Religious Freedom Report for 2018 § (2019).
- Correspondent, Special. "Congress Says Savarkar Did Not Fight for Freedom; BJP Demands Apology." *The Hindu*, October 19, 2019. <https://www.thehindu.com/news/national/karnataka/congress-says-savarkar-did-not-fight-for-freedom-bjp-demands-apology/article29742197.ece>.
- Deogan, Tarsem. "RSS Leader Ravinder Gosai Shot Dead in Ludhiana." *Hindustan Times*, October 17, 2017. <https://www.hindustantimes.com/punjab/rss-leader-ravinder-gosai-shot-dead-in-ludhiana/story-doDAwBmalUseUoyMGhZzmJ.html>.
- Dhume, Sadanand. "Modi's Three Big Gambles: We'll Soon Know If Demonetisation, GST and Adityanath Damaged BJP's Re-Election Prospects." *Times of India Blog*, June 2, 2018. <https://timesofindia.indiatimes.com/blogs/toi-edit-page/modis-three-big-gambles-well-soon-know-if-demonetisation-gst-and-adityanath-damaged-bjps-re-election-prospects/>.
- Engineer, Asghar Ali. "The BJP & Its 20 Years." *The Hindu*. *The Hindu*, May 31, 2000. <https://www.thehindu.com/todays-paper/tp-miscellaneous/tp-others/the-bjp-its-20-years/article28018819.ece>.
- Esposito, John L. *Rep. Jihad: Holy or Unholy War?*. United Nations Alliance of Civilizations, 2006.
- Ganguly, Šumit. "Explaining the Kashmir Insurgency: Political Mobilization and Institutional Decay." *International Security* 21, no. 2 (1996): 76–107.
- Ghose, Sagarika. "2017 Minus 2002: It's Now Hindutva plus in Gujarat." *Times of India*, December 3, 2017. <https://timesofindia.indiatimes.com/blogs/bloody-mary/2017-minus-2002-its-now-hindutva-plus-in-gujarat/>.
- Government of India, Election Commission of India New Delhi. *Statistical Report on General Elections, 1989 to the Ninth Lok Sabha. Vol. 1. Statistical Report on General Elections, 1990*.
- Government of India, National Informatics Centre, Parliamentary Constituency wise Turnout for General Election - 2014 § (2014).
- Government of India, National Informatics Centre, Parliamentary Constituency wise Turnout for General Election - 2019 § (2019).
- Government of India, Office of the Registrar General & Census Commissioner, 2001 Census Records, *Census And You / Religion* § (2001).
- Government of India, Office of the Registrar General & Census Commissioner, 2011 Census Records, *Population By Religious Community* § (2011).

- Government of India, Reserve Bank of India, STATE FINANCES : A STUDY OF BUDGETS § (2019).
- Graff, Violette. "Hindu-Muslim Communal Riots in India II (1947-1986)." Sciences Po, July 15, 2013. <https://www.sciencespo.fr/mass-violence-war-massacre-resistance/en/document/hindu-muslim-communal-riots-india-i-1947-1986.html>
- Graff, Violette. "Hindu-Muslim Communal Riots in India II (1986-2011)." Sciences Po, August 20, 2013. <https://www.sciencespo.fr/mass-violence-war-massacre-resistance/en/document/hindu-muslim-communal-riots-india-ii-1986-2011.html>.
- Guha, Ramachandra. *India after Gandhi: The History of the World's Largest Democracy*. London: Picador, 2007.
- Hindustan Times, Correspondent. "Unidentified Men Fire Shots at RSS Shakha in Ludhiana." Hindustan Times, January 18, 2016. <https://www.hindustantimes.com/punjab/2-unidentified-men-open-fire-at-rss-shaka-in-ludhiana/story-Alnk8qG1z2CnGQeDmNe34H.html>.
- Hirschman, Albert O. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, Mass: Harvard University Press, 1970.
- Hussain, Altaf. "South Asia | Kashmir's Flawed Elections." BBC News. BBC, September 14, 2002. [http://news.bbc.co.uk/2/hi/south\\_asia/2223364.stm](http://news.bbc.co.uk/2/hi/south_asia/2223364.stm).
- IT Staff. "Maharashtra Setback: As Another State Slips out of BJP's Hand, Saffron Hues Shrink from India's Map." India Today, November 26, 2019. <https://www.indiatoday.in/india/story/bjp-india-saffron-map-2019-after-maharashtra-turmoil-1622837-2019-11-26>.
- Jain, Vamika. "Lal Salaam: Explaining the Violence and Longevity of the Naxalite-Maoist Insurgency in Three Frames." *The Journal of Contemporary Asian Studies*, August 5, 2020.
- Liberhan Ayodhya Commission of Inquiry. *Report of the Liberhan Ayodhya Commission of Inquiry*, 2009.
- Mahmood, Cynthia. "Sikh Rebellion and the Hindu Concept of Order." *Asian Survey* 29, no. 3 (March 1999): 326–40.
- Mannathukkaren, Nissim. "Why the Bharatiya Janata Party Is Not Truly Bharatiya." *The Wire*, April 3, 2019. <https://thewire.in/politics/why-the-bharatiya-janata-party-is-not-truly-bharatiya>.
- Marzia Casolari. "Hindutva's Foreign Tie-Up in the 1930s: Archival Evidence." *Economic and Political Weekly* 35, no. 4 (2000), 218–28.
- Mehta, Deepak. "Naming the Deity, Naming the City: Rama and Ayodhya." *South Asia Multidisciplinary Academic Journal*, no. 12 (2015). <https://doi.org/10.4000/samaj.4053>.
- Misra, Satish. "Understanding the Rise of the Bharatiya Janata Party." Observer Research Foundation, September 21, 2018. <https://www.orfonline.org/research/44401-understanding-the-rise-of-the-bharatiya-janata-party/>.
- Prasad, Hari. "The Salafi-Jihadist Reaction to Hindu Nationalism." Hudson Institute. Hudson Institute, July 11, 2019. <https://www.hudson.org/research/15137-the-salafi-jihadist-reaction-to-hindu-nationalism>.
- PTI. "BJP MP Blames Muslims for Rising Crimes: India News - Times of India." *The Times of India*. TOI, July 27, 2018. <https://timesofindia.indiatimes.com/india/bjp-mp-blames-muslims-for-rising-crimes/articleshow/65169444.cms>.

- PTI. "Little-Known Sikh Outfit Claims to Have Killed Senior RSS Leader Gagneja." *Deccan Chronicle*, September 29, 2016. <https://www.deccanchronicle.com/nation/crime/290916/little-known-sikh-outfit-claims-to-have-killed-senior-rss-leader-gagneja.html>.
- PTI. "US Report on Global Religious Freedom Shows Bias against PM Modi: BJP." *The Economic Times*. *Economic Times*, June 22, 2019. <https://economictimes.indiatimes.com/news/politics-and-nation/us-report-on-global-religious-freedom-shows-bias-against-pm-modi-bjp/articleshow/69904977.cms?from=mdr>.
- "Punjab Assembly Election 2017 Results." *Elections in India*. *Elections in India Archive*, 2017. <https://www.elections.in/punjab/assembly-constituencies/>.
- Rana, Maanushi. "1984 Sikh Genocide: A Forgotten Chapter?" *Warpaint Journal*, July 2020.
- Rauf, Taha A. "Violence Inflicted on Muslims: Direct, Cultural and Structural." *Economic and Political Weekly* 46, no. 23 (June 4, 2011): 69–76.
- Rep. *Al-Qaeda in the Indian Subcontinent (AQIS)*. Counter Extremism Project, 2020.
- Roy, Abhik. "Anatomy of Fascism." *The Statesman*, November 16, 2019. <https://www.thestatesman.com/opinion/anatomy-of-fascism-1502822639.html>.
- Savarkar, Vinayak Damodar. *Essentials of Hindutva* 1923.
- Savarkar, Vinayak Damodar. *Hindu Rashtra Darshan*. Poona: Maharashtra Prantik Hindusabha, 1949.
- Spinner-Halev, Jeff. "Hinduism, Christianity, and Liberal Religious Toleration." *Political Theory* 33, no. 1 (February 2005): 35–57.
- Staff, Financial Express. "PM Modi Calls for 'Zero Tolerance to Terrorism' in Meeting with EU Delegation Ahead of Kashmir Visit." *The Financial Express*. *The Financial Express*, October 28, 2019. <https://www.financialexpress.com/india-news/pm-modi-calls-for-zero-tolerance-to-terrorism-in-meeting-with-eu-delegation-ahead-of-kashmir-visit/1747794/>.
- State Government of Jammu & Kashmir, Demographics § (2012).
- Tikkanen, Roosa. "India." *Commonwealth Fund*. *Commonwealth Fund*, June 5, 2020. <https://www.commonwealthfund.org/international-health-policy-center/countries/india>.
- Tully, Mark. "South Asia | Tearing down the Babri Masjid." *BBC News*. *BBC*, December 5, 2002. [http://news.bbc.co.uk/2/hi/south\\_asia/2528025.stm](http://news.bbc.co.uk/2/hi/south_asia/2528025.stm).
- Udayakumar, S. P. "Historicizing Myth and Mythologizing History: The 'Ram Temple' Drama." *Social Scientist* 25, no. 7/8 (1997): 11. <https://doi.org/10.2307/3517601>.
- Vaishnav, Milan, Jayaram Ravi, and Jamie Hinton. Publication. *Is the BJP India's New Hegemon?* Carnegie Endowment for International Peace, 2018.



# **“Ya Me Cansé:” The Role of International NGOs in the 2014 Ayotzinapa Massacre**

---

*Seo Yoon (Yoonie) Yang*

**Author Background:** *Seo Yoon (Yoonie) Yang grew up in the United States and currently attends Signal Mountain Middle High School in Signal Mountain, Tennessee in the United States. Her Pioneer research concentration was in the field of history and titled “Political Protest in Latin America.”*

## **Introduction**

On September 26, 2014, 43 students from the Ayotzinapa Normal School suspiciously “disappeared” in Iguala, Guerrero after a series of attacks from masked men and municipal police. Catalyzing overdue frustration from Mexican civilians who had previously been witnesses to hundreds of disappearances, kidnappings, and homicides, the tragedy, which would later be coined the Ayotzinapa Massacre of 2014, invigorated the country to demand thorough investigation and public reports from their regional and federal governments. The Ayotzinapa Massacre of 2014 was particularly significant because it inspired international momentum; the entire world turned their eyes to the thousands of Mexican civilians crowding the streets yelling “Ya me cansé” (“I’ve had enough”), and international non-governmental organizations (NGOs) began to take a closer look at the developing situation of passionate societal turmoil.

José Manuel Vivanco, the Americas director of the international NGO the Human Rights Watch, observed the response of both the Mexican government and the civilians and categorized Ayotzinapa as one of the “the worst atrocities we’ve seen in Mexico in years, but...hardly isolated incidents” (Goldman 3). Instead, he advocated that Ayotzinapa was reflective of the “longstanding failure of Mexican authorities,” focusing his statement on the historically lackluster response of the Mexican officials to large-scale tragedies (Goldman 3). The statement released by Vivanco shifted a large portion of the discussion from the immediate aftermath of Ayotzinapa to the greater problem undermining Mexican society and the safety of its citizens.

This paper uses an investigation of Ayotzinapa to investigate how external, international influences affected the internal, deeply rooted problems of Mexico. Initially, the paper will recognize the progression of events on September 26 before referring to previous Mexican NGO research conducted by various scholars. Through observation of the unique roles of Mexican NGOs, international investigative organizations, and international NGOs in the 2014 Ayotzinapa Massacre, the paper will ultimately recognize the pivotal role of these organizations in highlighting Mexico’s historic pattern of human rights violations to the global community.

## Context of the Ayotzinapa Massacre of 2014

According to Francisco Goldman, a *New Yorker* journalist who followed the progression of events of Ayotzinapa, in the early 21st century, the Ayotzinapa Normal School aimed to teach students how to become educators. Often selecting the brightest students from marginalized, indigenous, rural communities, the school provided a clear path for individuals to climb the social ladder while giving back to the community. The students were often given free time to visit their families; however, in late September of 2014, they were told by the school administration that family trips would be postponed in order to hold an “actividad de lucha” or a public action (Goldman 2). The students planned to block a highway to collect funds to participate in the October 2nd March, which commemorated the 1968 Tlatelolco Plaza Massacre, a government-led attack on student protestors.

On Friday, September 26, 2014, the students boarded their respective buses to travel to the destination of their organized protest; however, due to an insufficient number of busses, the group took a quick detour to Iguala, Guerrero. That night, masked men and municipal police attacked the students, abducting 43. The attack was seemingly organized by the wife of the Iguala mayor. Her brother, known as “El Molón,” was the leader of an illicit drug trafficking group called Guerreros Unidos, which held a significant amount of power over the city. The kidnapped students were eventually handed over to Guerreros Unidos, and although the exact details of the homicides are not known, it is widely believed that the violent gang brutally murdered the students.

Since the attacks, the Iguala mayor and his wife have been on the run. Thus, the public can only speculate as to why these rural, aspiring-teachers’ lives were brutally taken. The media speculates that their appearance upset the wife, who was scheduled to have a speech that night; furthermore, many believe that the mayor felt threatened by the student activists and their protests. However, there is minimal debate that the government was not involved in this incident; Guerreros Unidos has even threatened to release the names of the government officials who supported their violent attacks.

As the victims’ families demand investigations, authorities have discovered an even greater issue: the Ayotzinapa Massacre was only one incident out of thousands of equally violent crimes. The discovery of more and more mass graves affirmed the suspicion held for decades: since the militarization of the battle against drug cartels in 2006, “seventy thousand Mexicans have been killed and some twenty-seven thousand disappeared, effectively turning the whole country into a ‘narco grave’” (Goldman 3). This stark realization sparked national and international interest, bringing international NGOs to investigate. Through the work of these groups, the Ayotzinapa Massacre of 2014 has illustrated the tip of the iceberg, revealing thousands of tragedies hiding under the facade of the government.

## Literature Review

Previous studies have illustrated the rising tide of NGO influence through the examination of these working bodies at various levels of government. They have determined that NGOs employ various methodologies in their advocacy, impacting political reformation and the transformation of government ideologies. Douglas A. Chalmers and Kerianne Piester published research in 1996 to explore the contribution of NGOs in organizing civil society, hypothesizing the political



ramification of increased association between NGOs, the public, and the state. Although their research studying the role of NGO networks was inconclusive, Chalmers and Piester suggested (based on their preliminary analysis) that NGOs catalyzed social change (including increased self-determination in organizations leading to effective state relations), an enhanced presence of political pluralism to prevent political consolidation, and the “internationalization of politics” (Chalmers and Piester 3-7). The effect of the “internationalization of politics” directly relates to the international NGOs of Ayotzinapa as Mexican NGOs have set a precedent allowing foreign organizations to intervene in national affairs. This precedent introduced a mutual relationship between foreign individuals bringing external resources to Mexico and international organizations being given the opportunity to develop state-NGO relationships (Chalmers and Piester 7). Although the study argues that the increased association between NGOs and the state will result in “a more plural, specialized, non-partisan and internationalized politics,” the research does not illustrate the measures NGOs take in order to hold the Mexican government accountable; the study rather focuses on the long-term political impacts of the NGO phenomenon (Chalmers and Piester 9).

In June and July of 1997, as a witness to the first democratic election for the mayor of Mexico City, Katherine M. Bailey interviewed representatives from 23 Mexico City NGOs and recognized how NGOs stabilized Mexico as a democratic institution. Compared to the research of Chalmers and Piester, Bailey further delineates subcategories of NGOs, focusing her research on civic NGOs, informally defining the subcategory. Her research has attempted to categorize types of NGOs, but in the context of the 2014 Ayotzinapa Massacre investigation, the definitions of NGOs will take a flexible form, including various organizations with different purposes. Bailey sets the foundation for her argument on the role of civic NGOs, stating that NGOs are integral in holding government officials accountable within democratic systems. To elaborate, following evident PRI electoral fraud in the 1988 presidential election<sup>1</sup>, civic NGOs served as the public’s government and election watchdog (Bailey 8). NGOs intentionally empowered the public by creating flyers, pamphlets, and workshops that reinvigorated civic engagement. The public appearance of NGOs regulating the government-organized elections were “at least partially responsible for rejuvenating and legitimizing the corrupted institution of elections, promoting voting and voter awareness” (Bailey 17). Complementing the argument of Chalmers and Piester, Bailey demonstrates that civic NGOs not only shape the political atmosphere of Mexico but also assist in moving the political needle to favor democratic government ideologies.

However, NGO research, and its relationship to government structures, is not limited to analysis of civic NGOs. Krista M. Brumley conducted a case study about NGOs in Monterrey, Mexico between 2000 and 2003 to study NGOs on a purely local level and challenge the previously held notion that NGOs shift advocacy strategies in response to changing socio-political environments. She conducted her research through qualitative data derived from interviews with NGO leaders, observation of the inner workings of the NGOs, and archival research. Brumley aimed to conduct a case study that would fill many of the gaps in the Mexican NGO research, as previous NGO studies have only explored specific types of NGOs, strategies of NGOs under distinct circumstances, or NGOs on a national or international level. Through cross referencing the goals of NGO leaders with specific low-risk, medium-risk, and high-

---

<sup>1</sup> To learn more about the 1988 Presidential Election: <https://www.latimes.com/archives/la-xpm-1988-07-15-mn-7206-story.html>

risk<sup>2</sup> strategies used, Brumley argues that an NGO's strategy depends on its specific goals and their risk evaluation of the situation to meet agenda items within the constraints of their environment; for example, in Monterrey, NGOs evaluated their strategies and goals within the context of the conservative local government, the significant presence of Catholicism and machismo, and the lack of activism sympathy from media outlets (Brumley 398). Brumley's research fails to relate NGOs directly to the political dimension but supplements the findings of Bailey, Chalmers, and Piester as it offers a specific, localized interpretation of NGOs in Mexico; the local scope of Brumley's study reflects the intent to explore NGO influence in the 2014 Ayotzinapa Massacre, closely following the role of NGOs in a specific national tragedy.

This paper builds on existing NGO research by studying how international NGOs, often in cooperation with national NGOs, attempt to ensure the accountability of the Mexican government through foreign intervention in large-scale human rights crises. Detailed exploration of Ayotzinapa, through the lens of international NGO organizations, provides an opportunity to understand the influence of foreign bodies within a state's internal affairs. This exploration will help construct the argument that NGOs were a significant reason why Ayotzinapa became a stepping stone for international criticism of Mexico's failure to enforce human rights.

### **Effects of Mexican Non-Governmental Organization (NGO) Responses**

In the wake of national outrage, Mexican NGOs served as the primary actors consolidating investigative information and providing support for the victims' families. In particular, Tlachinollan: Center of Human Rights of the Mountain, a local NGO serving as a human rights legal consultant within the region of Guerrero, was significant in shifting the narrative of the everchanging Ayotzinapa situation. Tlachinollan Human Rights Center, in cooperation with other NGOs including the Miguel Agustín Pro Juárez Human Rights Center (Centro Prodh) and José María Morelos y Pavón Regional Center for Human Rights, collaborated to eventually publish a 270-page report titled "From the Trenches of Ayotzinapa: Defense for the education and life of the children of the people," highlighting the testimonies of the impacted rural community and families ("Desde las trincheras de Ayotzinapa" 53). The Tlachinollan Human Rights Center's report provided an opportunity for a historically marginalized community to share their stories in a period of national turmoil. Micro-level reports such as these have critical historic value because they ultimately serve as a primary source. The afterlife of the report will impact the way Ayotzinapa is remembered in history, driving the narrative for future Mexican civilians.

In contrast to the internal network of other Mexican NGOs, the work of the Tlachinollan Human Rights Center extends beyond Mexico's borders. On April 28, 2015, Jorge Luis Clemente Balbuena, a survivor of the 2014 Ayotzinapa attacks, and Hilda Legideño Vargas, the mother of disappeared Jose Antonio, accompanied by Isidoro Vicario Aguilar, a lawyer with Tlachinollan Human Rights Center, testified before the Parliamentary Subcommittee for International Human Rights in Ottawa, Canada. They came to demand Canada's recognition of Mexico's long-standing record

---

<sup>2</sup> For an explanation and breakdown of risk strategies employed by NGOs in Monterrey, refer to page 399 of Brumley's report.

of violating human rights (“Witnesses from Mexico will Testify” 1). Demanding the removal of Mexico from Canada’s 2012 Designated Countries of Origin list (a list outlining countries the Canadian government deems to be safe), Legideño Vargas highlighted that “in Mexico no one is really safe. We have been beaten. We have been threatened” (qtd. in Smith 3-4). The listing of Mexico on Canada’s Designated Countries of Origin complicates the ability of Mexican citizens to seek asylum in Canada because Canada recognizes the listed countries to safeguard human rights and offer national protection, deeming international asylum unnecessary. Libby Davies, a member of Parliament from the New Democratic Party, and Philip Stewart, President of the Union of B.C. Indian Chiefs, called upon the conservative Canadian government to hold Mexico accountable by cutting off relations until thorough investigations of Ayotzinapa were underway (Smith 4-5). The testimonies of Vargas, Antonio, and Aguilar shed light on the moral obligation of Canada to publicly condemn the actions of Mexico, an economic partner of the Canadian state. With international coverage of the committee meeting, the work of Tlachinollan Human Rights Center shifted the dynamic of the Canadian parliament, and the public sentiments of Davies and Stewart pressured the central government of Canada to recognize, rather than deny, the reality of Mexico in its inability to safeguard its citizens. While there was limited information regarding Canada’s continued discussion about the legitimacy of Mexico on the list, regardless, the list was eventually dissolved in 2019; many worried that publicly categorizing a state as safe would prevent desperate asylum seekers of those deemed “safe” countries from receiving Canadian aid in the future (Canada ends the Designated Country 1). It is unclear how much the Ayotzinapa testimonies influenced this decision.

### **Effects of International Investigative Organizations in Ayotzinapa**

Compared to previous smaller-scale disappearances in Mexico, the Ayotzinapa Massacre led Mexican civilians to seek the support from international organizations. In October 2014, the Mexican Federal Attorney General opened an investigation of Ayotzinapa in conjunction with an investigation conducted independently by Guerrero State authorities. However, to appease the victims’ families, the federal government allowed the Argentinean Forensic Anthropology Team (EAAF)<sup>3</sup> to begin its own investigation<sup>4</sup> (“Mexico: Ayotzinapa student’s enforced disappearances” 2).

On January 27, 2015, the Attorney General Murillo Karam released the results of the federal investigation: the 43 students had been kidnapped by the municipal police force, killed by the Guerreros Unidos group, burned in a garbage dump, then later dumped into the San Juan River (Meyer 2). However, the following EAAF report stated that there was not sufficient forensic evidence to support the government’s conclusion because, among other aspects, the soil sample surrounding the dump did not reveal evidence of a large-scale fire; furthermore, there was insufficient evidence connecting the discovered bone fragments in the garbage dump

---

<sup>3</sup> To learn more about the organization and its past investigative cases, refer to the link (in Spanish): <https://eaf.org/quienes-somos/>

<sup>4</sup> The intervention of EAAF into the Ayotzinapa investigations did not come without objections to its validity and reliability. Organizations claimed that government officials selected recovered human remains that could be analyzed by EAAF; furthermore, results would be kept from the public. To learn more about the criticisms against EAAF, refer to [www.laht.com/article.asp?ArticleId=2409696&CategoryId=14091](http://www.laht.com/article.asp?ArticleId=2409696&CategoryId=14091)

to the 43 students<sup>5</sup> (Meyer 2). The EAAF, previously having worked with the United Nations International Criminal Tribunal for former Yugoslavia and the International Criminal Court (“Quiénes somos” 2), held a degree of international credibility. Serving as a reputable and objective third party, the EAAF highlighted the refusal of the Mexican government and Attorney General to conduct accurate investigations following the Ayotzinapa Massacre. The EAAF revealed to the victims’ families, Mexican citizens, and the international community that the Mexican government was not equipped to hold forensic investigations, owing to faulty conclusions and the government’s general inability to respond to organized crime. Additionally, because the EAAF publicly discounted the investigations of the Mexican government, they set a precedent for third-party organizations to remain objective in interventionist investigations.

In addition to the forensics investigation by the EAAF, in an agreement between the victims’ families and officials from the Mexican government, a group of experts, the Interdisciplinary Group of Independent Expert (GIEI), was appointed by the Inter-American Commission on Human Rights to advise the investigations and provide third-party insight<sup>6</sup> (“Mexico: Ayotzinapa student’s enforced disappearances” 3). The GIEI later released a summary on the agreed facts of the incident. The report discussed various aspects of September 26, 2014 from the time frame of the attacks to the presence of military and security force personnel. At the end of the report, the GIEI provided recommendations for the federal government concerning the future of the investigation that were divided into four subgroups: investigation, responsibilities, search, and care for victims. Recommendations included encouraging the holistic consideration of human rights violations in Ayotzinapa, extending the scope of the investigation from murder to include excessive use of force and abuse of authority, and promoting increased dialogue between victims’ families, communities, legal representatives, and state officials (“Ayotzinapa Report” 29).

After the group of experts published their preliminary report, Enrique Peña Nieto, the president of Mexico, tweeted a response in gratitude, stating that he valued the recommendations (Leveille and Campbell 5). Twitter was not a particularly new platform for the president; he often used tweets to open dialogue between himself and the President of the United States, even notifying the public about meeting cancellations through Twitter (Campoy 1). Although Nieto hid behind the social media platform with his tweet of gratitude, the authentic response of the government unknown, the public tweets demonstrate that the international investigation led to national consideration by authority figures. It demonstrates that international interventions result in governmental acknowledgement, a success considering the federal government’s history of turning a blind eye to the hundreds of disappearances and homicides within its borders. Although the contract between the GIEI and the Mexican government expired in April 2016 and was not renewed, the tweets of Nieto provided an opportunity for the Mexican public to hold their authority accountable and demonstrated how the Nieto administration was addressing the reports released

---

<sup>5</sup> To read the full explanation as to why the EAAF did not support the government’s findings: <https://www.wola.org/analysis/mexicos-government-claims-the-43-ayotzinapa-students-were-incinerated-in-a-trash-dump-heres-how-theyre-wrong/>

<sup>6</sup> Civil society groups are raising questions and criticizing the legitimacy of the reports, claiming that the IACHR is working in conjunction with the Mexican government. Refer to <https://panampost.com/elena-toledo/2016/01/05/mexican-ngo-foreign-investigators-bury-facts-in-ayotzinapa-case/?cn-reloaded=1>

by third parties (Leveille and Campbell 3-5). Pressuring Nieto to release a statement after the release of the report, the GIEI assisted in empowering Mexican citizens by instigating a tangible government response to the inconsistencies lining the report.

## **Effects of Prominent International Non-Governmental Organization (NGO) Responses**

Although the EAAF and GIEI served as international investigative organizations, focusing primarily on forensic study and deducing the events that transpired on September 26, 2014, one prominent NGO, the Human Rights Watch (HRW), contributed to the aftermath of Ayotzinapa, filling the more traditional role of international NGOs. On November 3, 2014, the Human Rights Watch traveled to Guerrero and Mexico City to speak with victims' families, witnesses of both the Ayotzinapa Massacre and another similar incident of 22 individuals killed in Tlatlaya, and government officials ("Mexico: Delays, Cover-Up" 2-3.) Through these actions, the Human Rights Watch followed the typical procedures of international NGOs, often traveling to the affected countries or regions to provide support for victims and discuss policies and recommendations. Yet the Human Rights Watch was pivotal in bringing to light Mexico's continued human rights violations because the NGO did not seek justice only for the Ayotzinapa Massacre; rather, the NGO highlighted a larger-scale tragedy as well, emphasizing the fact that homicides and violence are a trend within the country.

In a similar manner, the Human Rights Watch used Ayotzinapa, an incident that had sparked national and international outrage, as an opportunity to bring justice and open dialogue about the thousands of previous disappearances in Mexico. For example, José Miguel Vivanco, the Americas director of the Human Rights Watch, published an open letter to the Attorney General Arely Gómez González. The letter discussed the lack of initiative on the part of the government in investigating large-scale atrocities. For example, it took the government 10 days to begin investigating the Ayotzinapa Massacre, and three months to investigate the tragedy in Tlatlaya (Miguel Vivanco 2-3). The open letter further highlighted the 250 disappearances plaguing former president Felipe Calderón's administration; it provided information that the Human Rights Watch had found research connecting those disappearances to security and authority figures and evidence of a routine disregard for accurate and timely investigations (Miguel Vivanco 1). Although these 250 disappearances were not directly related to the Ayotzinapa Massacre, the statistic served to exemplify how the tragedy of Ayotzinapa could have been prevented if the Mexican government had worked to enforce human rights in the past. By highlighting in a single document how other organizations, including the United Nations Committee of Enforced Disappearances, the National Human Rights Commission, and the International Red Cross, had previously sent recommendations to the Mexican government, the Human Rights Watch effectively illustrated how Mexico's history of human rights violations had been continuously brushed under the rug.

The Human Rights Watch used the 2014 Ayotzinapa Massacre as a catalyst to demonstrate Mexico's historic disregard for universal human rights and its responsibilities as a federal government to ensure the rights of millions of citizens. In the letter, the Americas Director of HRW specifically recommended that Attorney General, Gómez González "Release the names of the 30,000 people who were missing and have been found, as well as the 22,000 who remain missing," as catalogued and

registered in government databases. Through this specific request, and through publishing the open letter on an internationally accessible Human Rights Watch platform, Miguel Vivanco succeeded in bringing justice to the thousands of unsolved deaths and kidnappings in Mexico; furthermore, he emphasized the international concern that thousands of prematurely-closed investigative cases must be reopened in order to guarantee the legitimate governing of the Mexican state.

The intersection of international and national NGOs in Ayotzinapa and its effects are difficult to quantitatively measure. Additionally, although there was no intention of erasing the significant work conducted by Mexican NGOs, it is important to address the disproportionate emphasis of international NGOs to national organizations in the paper. Given the rarity of international intervention, it is plausible that international NGOs overshadowed national NGOs in the narrative of Ayotzinapa. However, the effects of international NGOs heavily relied upon the preliminary research and published reports of Mexican NGOs; without the pre-established relationships between the Mexican NGOs and the public, it is unlikely that the international NGOs would have effectively made an impact.

Thus, rather than drawing direct causes and effects of an individual NGO, there is greater meaning in recognizing how the NGOs served as fluid, constant actors shaping post-Ayotzinapa Mexico. Other than the significant investigations conducted primarily by the EAAF and GIEI, it is difficult to discern a particular NGO or a particular campaign that catalyzed international change. Many factors were at play during post-Ayotzinapa Mexico including international NGOs, national protests, and statements from government authorities, and it can be best understood that it was this combination of players that resulted in global recognition of the internal affairs of Mexico. Collaboration and cooperation were key. These NGOs collectively succeeded in pressuring the government authorities to begin investigations on the incident of Ayotzinapa and to hire third-party investigative organizations to conduct independent research. Fighting a continuous battle of government denial and corruption, however, the NGOs did not succeed in any passage of any specific policy or ensure full justice for the 43 victims of Ayotzinapa.

## **Global Implications of the 2014 Ayotzinapa Massacre**

The consequences of the 2014 Ayotzinapa Massacre extend far beyond the borders of Iguala, Guerrero, directly impacting the relations between Mexico and other states and setting an international precedent for state accountability. Initially, the global outrage over the faulty investigations of the Mexican government and Attorney General resulted in direct action from the United Nations Committee on Enforced Disappearances. After receiving submissions and reports from various NGOs, the Committee investigated Mexico's history concerning disappearances, comparing the government actions against the Rome Statute of the International Criminal Court and the Inter-American Convention on Forced Disappearances of Persons.

The Committee released a report on February 13, 2015, and it commended Mexico's ratification of all core UN human rights treaties and Mexico's decisions to have an open-door visitation policy for the Human Rights Council ("Concluding observations on the report submitted by Mexico" 1-2). However, the report stated that the force of the legislation, its implementation, and its use in legitimate situations were not in full compliance with the two aforementioned conventions. The report showed Mexico's repetitive violations, from the inappropriate handling of disappearances or homicides to the "serious challenges the State party faces in the

prevention, investigation and punishment of enforced disappearances and the search for disappeared persons” (“Concluding observations on the report submitted by Mexico” 2). The report specifically mentioned the Ayotzinapa Massacre, an incident that helped highlight Mexico’s lack of compliance with the UN conventions.

The United Nations is the primary inter-governmental organization developing international treaties in the modern global community. Serving as an objective third party assessing federal governments and their compliance to international legislation, the United Nations holds considerable weight in the international community. Similar to the pressures within the Canadian government resulting from the testimonies of Ayotzinapa discussed earlier in the paper, the United Nations highlighting the lack of compliance of Mexico will affect the reputation of the country. Other countries, especially states that find national pride in their enforcement of human rights, will be more hesitant to economically and politically support Mexico after recognizing their continuous lack of regard for human rights. Thus, the Ayotzinapa Massacre and the involvement of international NGOs affected the international footing of Mexico by decreasing the willingness of other countries to engage in diplomatic, economic, and governmental relations.

## Conclusion

The Ayotzinapa Massacre of 2014 was never completely resolved. Although Mexican officials argue that the forensics investigations have concluded, revealing the remnants of students in the trash dump and river, reports by the EAAF and GIEI have repetitively determined these claims to be built on insufficient evidence. Additionally, many international NGOs consider the investigations to be inconclusive; on June 13, 2016, twelve NGOs published a letter to President Enrique Peña Nieto, expressing their strong concern that further investigations had not been conducted two months after the report released by the GIEI (Letter to Lic. Enrique Peña Nieto 1). Given the reigning principle of state sovereignty in national affairs, it is unlikely that further investigation into the Ayotzinapa Massacre will be led by the Mexican government unless coerced by the civilians or international community.

Yet the legacy of Ayotzinapa continues. The national and international scrutiny of Mexican officials throughout the forensic investigations have informed global citizens around the world of Mexico’s continuous human rights violations. For example, November 15 was named “International Day of Global Action” and countries including Spain, France, and Croatia came together in solidarity to protest the silence that has surrounded the disappearances of Mexico for decades (“Ayotzinapa's 43 Disappeared Students” 1). These international protests illustrate how the 2014 Ayotzinapa Massacre, and the collaborative work of national and international NGOs, has a widespread global impact, inspiring everyday civilians to stand up against Mexico’s systemic negligence of enforcing human rights. Through understanding of the Ayotzinapa Massacre, review of relevant NGO research, investigation of international and national NGOs, and exploration of the global impacts of the massacre, it is clear that international NGOs shaped the narrative of the Ayotzinapa Massacre of 2014, addressing authority figures, highlighting the shortcomings of the government investigations, and publishing comprehensive reports for the public.

## Bibliography

- Amnistía Internacional, et al. Received by Lic. Enrique Peña Nieto, 13 July 2016. "Ayotzinapa's 43 Disappeared Students." *Ayotzinapas 43 Disappeared Students*, blogs.haverford.edu/disappeared-students/international/.
- Bailey, Katherine M. "NGOs Take to Politics: The Role of Non-Governmental Organizations in Mexico's Democratization Effort." *Penn State University*, citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.577.5275&rep=rep1&type=pdf.
- Brumley, Krista. "Understanding Mexican NGOs: Goals, Strategies, and the Local Context." *Research Gate*, Sept. 2010, www.researchgate.net/publication/226702057\_Understanding\_Mexican\_NGOs\_Goals\_Strategies\_and\_the\_Local\_Context.
- Campoy, Ana. "Mexico's President Canceled His Meeting with Trump after Being Baited on Twitter." *Quartz*, Quartz, 26 Jan. 2017, qz.com/895684/mexican-president-enrique-pena-nieto-cancelled-his-jan-31-meeting-with-us-president-donald-trump-over-a-tweet/.
- Chalmers, Douglas A, and Kerianne Piester. "NGOs and the Changing Structure of Mexican Politics." *Columbia University*, www.columbia.edu/~chalmers/NGO.html.
- Committee on Enforced Disappearances. International Convention for the Protection of All Persons from Enforced Disappearance, 2015, pp. 1–10, *Concluding Observations on the Report Submitted by Mexico under Article 29, Paragraph 1, of the Convention*.\*.
- EAAF, Contenidos. "Quiénes Somos." *EAAF*, eaaf.org/quienes-somos/.
- Goldman, Francisco. "Crisis in Mexico: The Protests for the Missing Forty-Three." *The New Yorker*, 12 Nov. 2014, www.newyorker.com/news/news-desk/crisis-mexico-protests-missing-forty-three.
- Immigration, Refugees and Citizenship Canada. "Canada Ends the Designated Country of Origin Practice." *Canada.ca*, Government of Canada, 17 May 2019, www.canada.ca/en/immigration-refugees-citizenship/news/2019/05/canada-ends-the-designated-country-of-origin-practice.html.
- Leveille, David, and Monica Campbell. "Investigators Release a Damning Report on Mexico's 43 Missing Students." *The World from PRX*, 25 Apr. 2016, www.pri.org/stories/2016-04-25/investigators-release-damning-report-mexicos-missing-43-students.
- "Mexico: Ayotzinapa Student's Enforced Disappearance – Timeline." *Amnesty International*, 23 Sept. 2015, www.amnesty.org/en/latest/news/2015/09/mexico-ayotzinapa-student-s-enforced-disappearance-timeline/.
- "Mexico: Delays, Cover-Up Mar Atrocities Response." *Human Rights Watch*, 24 Aug. 2020, www.hrw.org/news/2014/11/07/mexico-delays-cover-mar-atrocities-response#.
- Meyer, Maureen. "Mexico's Government Claims the 43 Ayotzinapa Students Were Incinerated in a Trash Dump. Here's How They're Wrong." *WOLA*, 23 Apr. 2016, www.wola.org/analysis/mexicos-government-claims-the-43-ayotzinapa-students-were-incinerated-in-a-trash-dump-heres-how-theyre-wrong/.
- Miguel Vivanco, José. "Letter to Mexico's Attorney General on Human Rights Crisis." *Human Rights Watch*, 24 Aug. 2020, www.hrw.org/news/2015/04/28/letter-mexicos-attorney-general-human-rights-crisis.



- Smith, Julia. "Canada Called upon to Recognize Human Rights Violations in Mexico." *Rabble.ca*, 14 Apr. 2015, [rabble.ca/news/2015/04/canada-called-upon-to-recognize-human-rights-violations-mexico](http://rabble.ca/news/2015/04/canada-called-upon-to-recognize-human-rights-violations-mexico).
- Tlachinollan Centro De Derechos Humanos De La Montana, 2014, *Desde Las Trincheras De Ayotzinapa: La Defensa Por La Educación y La Vida De Los Hijos Del Pueblo*, [www.tlachinollan.org/wp-content/uploads/2015/08/Informe-XXI-Tlachinollan.pdf](http://www.tlachinollan.org/wp-content/uploads/2015/08/Informe-XXI-Tlachinollan.pdf).
- Tlachinollan, por. "XXI INFORME: Desde Las Trincheras De Ayotzinapa." *Tlachinollan*, 29 Aug. 2015, [www.tlachinollan.org/xxi-informe-desde-las-trincheras-de-ayotzinapa-la-defensa-por-la-educacion-y-la-vida-de-los-hijos-del-pueblo/](http://www.tlachinollan.org/xxi-informe-desde-las-trincheras-de-ayotzinapa-la-defensa-por-la-educacion-y-la-vida-de-los-hijos-del-pueblo/).
- Villa, Alejandro Valencia, et al. Interdisciplinary Group of Independent Experts, *Ayotzinapa Report: Research and Initial Conclusions of the Disappearances and Homicides of the Normalistas from Ayotzinapa*, [media.wix.com/ugd/3a9f6f\\_e1df5a84680a4a8a969bd45453da1e31.pdf](http://media.wix.com/ugd/3a9f6f_e1df5a84680a4a8a969bd45453da1e31.pdf).
- "Witnesses from Mexico Will Testify Today to Parliament's International Human Rights Subcommittee About the Missing Students of Ayotzinapa." *No More Blood in Mexico*, Amnesty International , 28 Apr. 2015, [nomorebloodinmexico.org/tag/amnesty-international/](http://nomorebloodinmexico.org/tag/amnesty-international/).



# **Intangible Assets: The Cause of Veblen Goods A Breakdown of the Intangible Factors that Drive Pricing Premiums**

---

*Vardaan Tekriwal*

**Author Background:** *Vardaan Tekriwal grew up in Indonesia and currently attends the Jakarta Intercultural School in Jakarta, Indonesia. His Pioneer research concentration was in the field of business and titled “Financial Accounting & Analysis, Capital Markets, Entrepreneurship, and Decision Making.”*

## **Abstract**

Veblen goods are goods that go against traditional financial and economic theories: instead of their quantity demand decreasing with an increase in price, the opposite occurs. Intangible assets are the portion of a firm's value that lacks physical substance. By dissecting intangible assets into two factors – brand recognition and intellectual & creative rights – this paper investigates their correlation with Veblen goods. This is accomplished through the comparison of advertising expenses (a direct cause of customer brand recognition) and research & development expenses (a direct cause of the company's intellectual & creative rights) to the gross margin (depicting the firm's ability to mark-up goods). The sample of the study includes various public firms in two distinct sectors: the consumer electronics industry and the consumer apparel industry. The tension lies in the fact that unique companies have unique business models and strategies, creating a possibility of no such correlation (and causation). Nonetheless, this paper finds a clear causal relationship between intangible assets and Veblen goods.

**Keywords:** Intangible assets; Veblen goods; brand recognition; intellectual & creative rights; advertising expense; research & development; behavioral finance

## **1. Introduction**

Most traditional finance and economics studies rely on the assumption that quantity supply and quantity demand curves indicate that the higher the product's price, the

less quantity demanded (Marshall 1890). Goods with the same utility – operational, performance, and intimacy – should yield the same response from the market assuming the product's price is the same (Black and Scholes 1973). Although this holds true for the majority of the market, Veblen goods defy this theory. Veblen goods recognize a higher quantity demanded for an increase in price whereby the utility of the product does not necessarily play a role in the trade (e.g. jewelry, luxury cars). As a result, Veblen goods are more productive assets than conventional products: they induce more quantity demand at a higher price.

By traditional means, goods in a market are not predetermined to a standard or Veblen status; the market reaction determines their outcome. That being said, a company owns both tangible and intangible assets. The tangible assets vary as a result of the management of the company: decisions the management make include areas of investment, candidates to employ, sources/suppliers of materials, etc. Rarely does this impact the consumer perception of a brand and their products (Routledge et al. 2007). When compared to the market price without premiums, consumers who continue to endorse products that have a higher perceived value are the main catalysts for Veblen goods (Veblen 1934). As such, the effects that tangible assets have on a company's ability to make Veblen goods are considered to be negligible. On the other hand, while the tangible assets are based on the management of the company, intangible assets are valued by the appeal to the consumer base (i.e. behavioral finance). This facet is the primary stimulant for Veblen goods. Hence, while traditional finance theorizes the effectiveness of tangible assets, Veblen goods are one of the various effects of the intangible assets.

To prove that this idea holds true, this paper intends to define a correlation between the two topics: intangible assets and Veblen goods.

### **1.1 Breakdown of Intangible Assets**

The intangible assets value varies due to various factors, out of which the two defining ones are:

- brand recognition
- intellectual & creative rights

Brand recognition includes any element that directly impacts the means whereby the consumer base perceives a unique brand: a brand's status symbol, brand awareness, perceived quality, etc. In general, one can expect an increase in sales if a brand is well known and positively associated with the customer base. Attempting to model such a claim is one of the objectives of this paper.

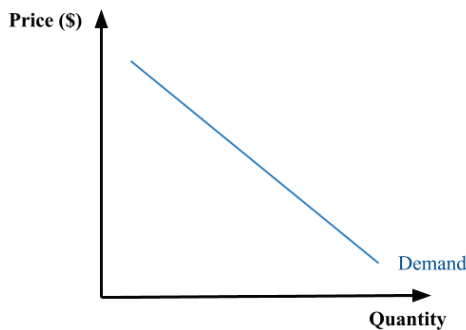
Intellectual & creative rights include patents, copyrights, trademarks, licenses, contracts, and the like. Another way to trigger customer engagement is to offer more advantageous products. Products with distinct points of differences may appeal to the public, increasing their quantity demanded, thus increasing sales and profitability. This is the second aspect that this paper attempts to determine: whether the correlation between the intellectual & creative rights and Veblen goods suggests causation.

Consumer biases cause different expectations for this intangible value in different industries and products, and thus alter the effects that these components have. The impact that the brand value has on the ability to mark up a product (without sacrificing sales) of a company in one industry may differ from a company in another

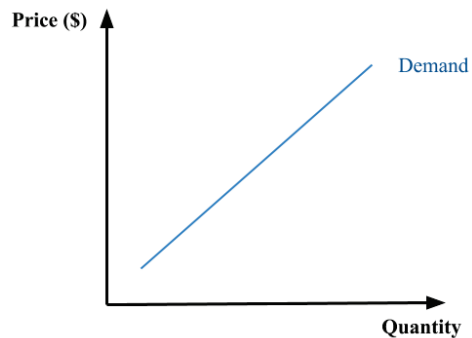
industry. The two unique industries that are analyzed in this report are the consumer electronics industry and the consumer apparel industry.

### 1.2 The Veblenian Spectrum

It is possible to classify goods in a black-or-white sense — simply as a Veblen good or not. A good whose demand curve follows the curve shown in Figure 1 can be classified as a normal good; accordingly, a good whose demand schedule follows one similar to that shown in Figure 2 can be classified as a Veblen good.



**Figure 1**



**Figure 2**

The slopes, or curves, of the lines depict the type of good. If the curve of the graph is negative (Figure 1), then it is a normal good (it follows traditional economics), but if it is positive (Figure 2), then it might be a Veblen<sup>[1]</sup>. The gradients of the curve can be of any value; the closer the value is to a positive number, the higher its Veblenian 'value,' and vice versa.

This, in turn, indicates that any Veblenian value of a good falls into a spectrum: there are an infinite number of values for these slopes. Having infinite possible values makes the Veblenian value a continuous variable.

### 1.3 Breakdown of the Industries Analyzed

The electronics industry hosts competition for various business models. Firms may appeal to the cheap-tech niche, or they may position themselves in the luxury-tech market. Smartphones, laptops, and tablets, amongst others, have seen considerable growth over the past years. Due to prices and profitability of the products varying immensely, analyzing the effects of intangibles that have caused this growth should yield results and conclusions that help firms enhance their business models.

The apparel industry encompasses a wide range of products varying in prices and styles, which implies that the business models and budgets of firms should vary as well. Companies may sell their products as, first, affordable and unbranded; second, premium and branded; third, superior athletically; and fourth, an amalgamation of the previous three listed. Deducing a correlation regardless of these considerable differences – a correlation that holds for all four models – can suggest consumer behavior in the industry as a whole.

## 1.4 Contribution to Literature

Prior research has been conducted on the financialization of the pharmaceutical industry through the Veblenian lens. Recent literature shows that intangible assets are the main driving forces behind the profits earned in this industry (Baranes 2017) and that the general consumer base (in all industries) is the cause of the Veblenization of products (McCormick 2002). This paper extends this field by proving the relationship between two additional industries (the consumer electronics and consumer apparel industries), while also attributing the Veblenization of products to another factor – brand recognition.

The paper also demonstrates the importance, to firms, of analyzing consumer behavior and psychology; these areas, if exploited, have the potential to cause a sizable uptick in the brand's profitability. The fact that each industry needs to target a different factor of intangible assets remains, but the overall impact of consumer behavior should not be overlooked. Profits, a corporation's primary goal, are the result of cutting costs and increasing revenue. Targeting the right audience (in the most effective manner) serves both of these aspects: it increases revenues by attracting the optimal audience, while simultaneously reducing the unnecessary expenses of probing into a less beneficial niche (Marx 1867).

Lastly, another facet of the market behavior that this paper advances is a brand's pricing decisions: factors that help companies can understand what turns their products into Veblen. With this in mind, the course of action following a successful advertising campaign or a breakthrough in innovation suggests that a corporation reevaluates its current pricing schemes due to the increase of their goods' Veblenian value.

## 2. Hypotheses

Given that intellectual & creative rights support technological advancements, and that consumers who obtain unique technology should be more willing to pay a premium, it is likely that the technological industry witnesses a larger impact of intellectual & creative rights on its Veblen goods as compared to the advertising expense.

On the other hand, brand reputation, status, and perceived value should be the driving forces of sales in the clothing industry. Apart from the athletic aspect of the industry, there are not many channels or methods that exemplify the effect of intellectual & creative rights. Trademarks, patents, and copyrights may yield revenue from royalties; however, it is unlikely that this value is greater than that previously mentioned.

Hence:

**H1:** Intellectual & creative rights impact a firm's ability to produce Veblen goods in the consumer electronics industry more than the firm's brand recognition does.

**H2:** Intellectual & creative rights impact a firm's ability to produce Veblen goods in the consumer *clothing* industry *less* than the firm's brand recognition does.

In terms of the regressions, the graphs with logarithmic functions should entail more reliable correlations, as the nature of the log explains a gradual negative incremental change in the *y* axis. As competition in industries suppresses firms' gross margins, the occurrence of extremely high margins is rare. There is also a limit on the proxy whereby the intangible assets are calculated. A corporation typically does not

extensively spend for such accounts, but even if they do, the slowing growth in the  $y$  axis acts as a cut-off. The coefficients of natural logs can determine the significance of the regression as well: an aspect that is imperative for the data analysis. The next-most reliable regression is the linear regression. As all data is concentrated in Quadrant 1 of the cartesian plane (because there are no possible negative values for either variable), the linear regression offers a constant change in the  $y$  axis. Its coefficient, the gradient, also allows for deeper analysis of the regression. Any other type of function displays a gradual positive incremental change for  $y$ , something that goes against the intuition of the causal relationship.

### 3. Method

The precise value of intangible assets is nearly impossible to discern as they are qualitative variables. However, there are quantitative methods that can provide a reasonable estimate for the value of such intangible assets.

Also, there is not simply a single factor that wholly impacts the branch of the intangible asset. The sale of a product is not only due to one factor: utility factors, operational factors, and attributional factors all play a role in the market reception.

#### 3.1 Brand Recognition

Brand recognition should be the effect of advertising. The role of advertising is to create brand awareness, customer engagement and loyalty, and overall, the desire to purchase a good. As explained in the mere-exposure effect theory, constant exposure to the same product increases its perceived value (Grimes et al. 2007). Similarly, advertising instills the brand name, or product, in the customer. In some instances, advertising is apprehended within a specific schema, which may engage the customer when another facet of the same schema is tapped. This holds especially true for when a consumer is exposed to another product from the same firm; they prefer products from firms that they have experience with, as compared to a new one (Stephan 2013). These effects depict a causal relationship between the advertisement of a firm and the value of their brand recognition. Therefore, their total advertising expense can be used as a reasonable proxy for its brand recognition. Granted, various corporations may circumvent the reliability of this proxy, with some already having a strong and loyal consumer base, or others not needing to spend on advertising as impacts from prior campaigns are still sustained. Yet, for the majority of corporations, the advertising expense succeeds as a proxy within fiat.

#### 3.2 Intellectual & Creative Rights

Intellectual & creative rights should be the byproduct of a firm's research & development (R&D) sector. Patents, copyrights, trademarks, and licenses fall under the umbrella of inherent brand uniqueness. They are the unique selling points (USP) of companies. In theory, more resources allocated to a project should mean a higher success rate (Greenhalgh et al. 2006). In this scenario, where the success rate is measured by the extent of USPs alongside the allocated resources expended in the sector, this total R&D expense should be a reliable alternative to the value of all the intangible rights of the firm. Henceforth, the total R&D expense is used as a substitute for a firm's intangible rights.

### 3.3 Veblen Goods

Similar to the value of intangibles, not all Veblen goods are the same: they fall into a spectrum. The increase in quantity demanded for a dollar increase in price varies amongst goods. Since it is classified as a qualitative variable, one way to get a discrete value for each firm is through its gross margin. The gross depicts the profitability of the firm<sup>[2]</sup>. The higher on the Veblen spectrum the good falls, the more profitable it is – the firm charges a greater price for the same product without sacrificing sales – so the profitability also increases. By using the gross margin instead of simply the qualitative label of a Veblen good, a measurable value is obtained. This measurable value will be the quantitative substitute for the qualitative Veblenian value variable.

### 3.4 Data Collection

In order to analyze the aforementioned variables, the companies need to have shared their financial and accounting statements. Since private companies are not obligated to share this information, this paper analyzes the data provided by 20-30 public companies in either of the studied industries<sup>[3]</sup>. Such companies are selected at random, without any biases towards the choices. Companies that uphold unique business models are more effective in this paper, as they represent all the firms in said industry.

The data collection process includes the following variables for each:

- sales (total revenue)
- cost of goods sold (COGS)
- gross margin
- advertising expense
- research & development expense

Each of the aforementioned data points will be collected from annual financial statements for each firm over the past five years. Corporations are required to disclose four out of the five criteria: total revenue from sales, costs of goods sold, gross margin, and their research & development expenses. These accounts can either be found within the balance sheets or the income statements that a corporation releases. However, the advertising expense is not a value that a corporation is mandated to disclose, although many nonetheless share it in the analyses of their financial summaries.

### 3.5 Data Processing

The data over the five years for each individual firm from each category (sales, COGS, gross margin, advertisement, and R&D) is processed to acquire the arithmetic mean (average). This is done to mitigate the effect of anomalies in the set. Another measure taken to increase accuracy is the method of crosschecking. The calculated gross margin value is crosschecked with the value given by the firm. Any discrepancies found warrant further research into the firm to identify the valid value.

Each category is taken as a percentage of the sales: common sizing. Since these companies are listed in a myriad of stock-exchanges, they convey their values in different currencies, allowing leeway for uncertainty. Common sizing these values ensures that external factors like inflation, interest rates, etc. are lessened. All in all,

this method uniforms the entirety of the data, while minimizing the effects of random errors.

This processed data is then converted into a scatterplot, comparing the gross margin against both proxies — the advertising expenses and the research & development. The independent variable (on the  $x$  axis) is set as either of the proxies, while the dependent variable (on the  $y$  axis) represents the gross margin. This is because the paper strives to find a causation between these factors: the proxies are the explanatory variables, while the gross margin is the response. The least squared minimum residual/correlation coefficient and descriptive statistics (i.e. central and variable tendencies) are the means whereby the calculation and the analyses are derived, and thence, the conclusions between the two variables.

#### 4. Data

Over 500 data points have been processed from various corporations over the past five years between FY 2014/15 and FY 2018/19. These points have been recorded randomly from corporations with unique business models and pricing schemes: a simplified random sample (SRS).

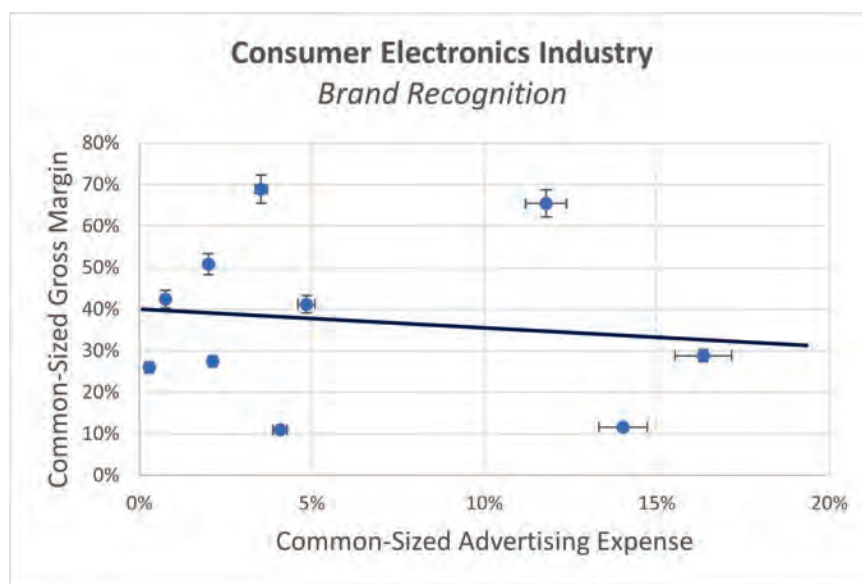
The error bars extending from each coordinate in the graphs refer to the percentage variance within the minimum and maximum of the variables they contest. Seeing as companies may have altered their business models and strategies between the time period in which the data was collected and when it was graphed, the error bars are due to both random and systematic errors. However, as each firm is independent of one another, systematic errors for the data as a whole are improbable.

The strength of each correlation depicts the impact that each proxy might have on the gross margin. However, some anomalies may exist as well, causing a skew. Also, the regressions are not analyzed with a biased lens of linearity. The regressions are based simply on means that are best suited for representing the data. While the logarithmic regression is optimal, as stated in Section 2: Hypotheses, regressions might not follow this trend. It should also be noted that not all axes begin at zero. This is for representative purposes, to emphasize the area in question.

Each table represents the bivariate descriptive statistics of the subsequent graph. Hence, the number on the graph and the number on the table are the same. The regressions are represented in a decimal format and not a percentage, because either axis is a percentage: the percentages for the  $x$  and the  $y$  axes cancel one another out during this calculation.



### 4.1.1 Consumer Electronics Industry, Brand Recognition



**Graph 1**

**Table 1**

Description	Value
Regression	$- 0.458x + 0.401$
Correlation Coefficient ( $r$ )	$- 0.134$
Variance ( $r^2$ )	0.018
$S_x$ R&D expense <sup>[4]</sup>	5.84%
$\bar{x}$ R&D expense <sup>[5]</sup>	5.98%
$S_x$ gross margin	20.19%
$\bar{x}$ gross margin	37.35%

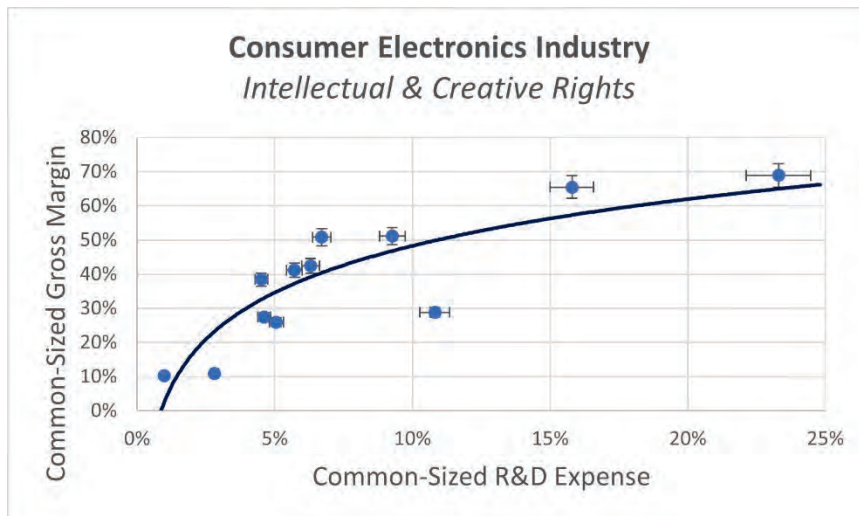
The randomness depicted in Graph 1 shows that any correlation between the two variables is negligible. To begin with, the correlation coefficient is  $-0.134$ . This value, as it is close to 0, is one method that describes the very weak correlation. This negative correlation is complemented with the variance: 0.018. This implies that 1.8% of the changes in a firm's gross margin can be attributed to its advertising expense in the consumer electronics industry. In sum, the negligible correlation between the two

variables indicates that brand recognition does not have an impact on the Veblenization of a product.

The center and spread statistics illustrate the randomness of the corporations at hand. The mean of the advertising expense is 5.98%, while its standard deviation is 5.84%. The standard deviation, as a large relative portion of the mean, shows the data's variability. One standard deviation away from the mean encompasses an area of 68%, so 32% of the data lies below 0.14% and above 11.82%. The gross margin is also extensively volatile; however, this is less variable than the advertising expense due to market dynamics and competition. This shows the unbiasedness in the data collection process, an objective that increases its accuracy.

The coordinates in the scatter plot do not follow any particular trend (either logarithmic or linear) yet they have a positive skew in the advertising expense. The error bars vary greatly for some coordinate points while they do not for others; some points have large vertical error bars yet maintain lower horizontal ones, and vice versa. This suggests that one variable does not overly impact the other; if this were the case, the error bars in either of the axes should fluctuate systematically. The slope of the regression is  $-0.458$  and is nearly completely horizontal; this emphasizes the random volatility of the data.

#### 4.1.2 Consumer Electronics Industry, Intellectual & Creative Rights



**Graph 2**

**Table 2**

Description	Value
Regression	$0.197(\ln(x)) + 0.938$
Correlation Coefficient ( $r$ )	0.857
Variance ( $r^2$ )	0.735
$S_x$ R&D expense	6.20%
$\bar{x}$ R&D expense	7.99%
$S_x$ gross margin	18.86%
$\bar{x}$ gross margin	38.48%

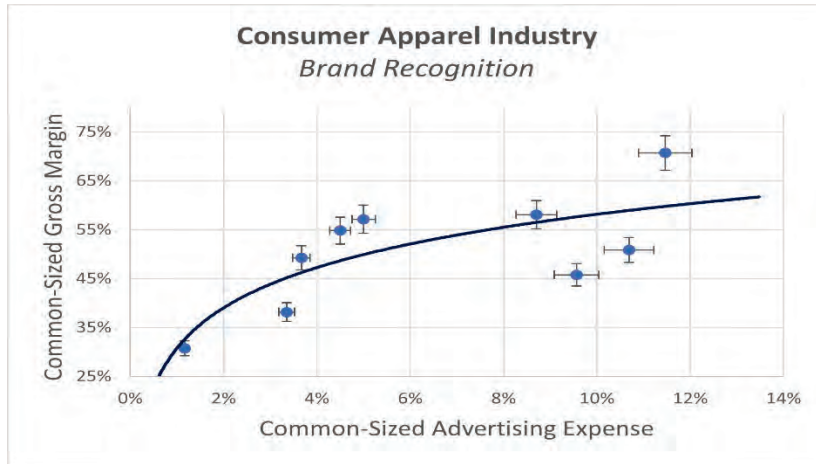
The correlation of the intellectual & creative rights to the Veblen ‘value’ in the electronics industry is axiomatic. Shown by the correlation coefficient of bivariate Graph 2, 0.857, the regression is strong and positive. The variance is 0.735, attesting that 73.5% of the change in the gross margin can be attributed to the research & development sector for the electronics industry as a whole. This number is the single most influential factor in the industry, as it accounts for more than 50%. All in all, there is a clear correlation between these two variables.

The regression of best fit is a natural log function. Compared to the linear function, the difference in the variance is +0.067: the logarithmic function is over 9% more accurate at representing the data as compared to the linear model. The coefficient of this natural logarithmic regression is 0.197: with every 1% change in the research & development of a firm, its gross margin changes by roughly 19.7%. Complemented by the strength of the regression, it can be inferred within reason that the aforementioned statement that the natural log regression is the best fit holds true.

The center and spread of the data represent its randomness and reliability. Again, the standard deviation for the research & development values makes up 77.60% of the mean. Due to the empirical rule of statistics, together, 32% of the dataset lies above 1.78 and below 0.22 times the mean. Discussions on the randomness of the gross margins are relatively constant between Graph 1 and Graph 2, as the companies in the same industry should have this same trend. The majority of the error bars have a proportional length for their horizontal and vertical components. This shows that there are more systematic errors in the data collection than there are random errors; the maximum and minimum values for either axis might have occurred during the same year, which would increase the reliability as it further supports the regression.

The  $x$ -intercept of the graph lies at 0.009. Although estimating the value for when the gross margin is zero is counterproductive (as a company should not be operating if they are selling their goods at a lower price than it cost them to acquire them), it can yield insights into industry dynamics. In this case, the  $x$ -intercept of the regression implies that the minimum percentage that companies spend on research & development in the consumer electronics industry, regardless of their gross margin, is 0.9% of their total sales revenue.

#### 4.2.1 Consumer Apparel Industry, Brand Recognition



**Graph 3**

**Table 3**

Description	Value
Regression	$0.112(\ln(x)) + 0.856$
Correlation Coefficient ( $r$ )	0.750
Variance ( $r^2$ )	0.562
$S_x$ R&D expense	3.69%
$\bar{x}$ R&D expense	6.46%
$S_x$ gross margin	11.70%
$\bar{x}$ gross margin	50.61%

Having a value of 0.75, the correlation coefficient of the scatter plot is a strong positive one, evincing a definite correlation between the brand recognition and the Veblenization of a good in the consumer apparel industry. The variance, 0.562, attributes 56.2% of the change in the gross margin as a factor of the advertising expense. This value, as it is greater than 50%, is likely the factor that influences the volatility in the gross margin the most.

The best fit regression for this dataset is a function of the natural logarithm. As compared to the linear model, the difference in variance is +0.106, denoting that the strength of the regression of the log function is nearly 19% better at representing this data as compared to a linear one. Its coefficient is 0.112, which means that for

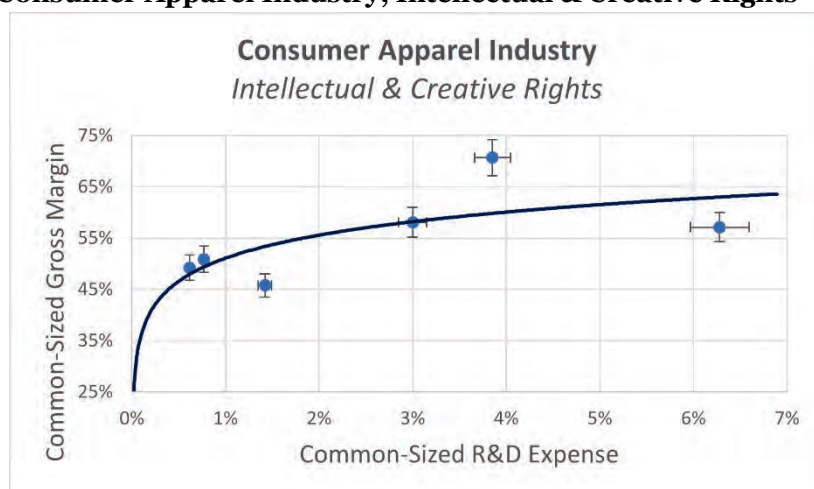
every percent change in the advertising expense of a company in the apparel industry, their gross margin fluctuates by 11.2%. Bolstered by the fact that the brand recognition of a corporation is the major influencing factor in its Veblenian value, the prior statements hold.

Although the extent to which the data is spread from its center is less than that of the electronics industry, it still describes a considerable randomness. The standard deviation for the advertising expense makes up 57% of the mean; the values are more centralized compared to *Graph 2*. For the gross margin, the same is true: this value is 23%. This is also seen in the graph itself, as the majority of the coordinates have a  $y$  value of approximately 55%. This analysis indicates that either the data collection process is at fault, and more varied data should have been collected, or that the market for this unique industry behaves in this manner in general.

The error bars in this scenario demonstrate both systematic and random errors, showing reliability to a certain degree. Still, the systematic errors outweigh the random errors; the lengths for the horizontal and vertical limbs are the same to a certain extent. There are error bars that have an elongated vertical component, and others that have elongated horizontal ones. This volatility is attributed to random error: factors beyond the bivariate analysis.

In *Graph 3*, the single  $x$ -intercept occurs at 0.0005. This means that the corresponding gross margin is 0%, or that the company is not profitable whatsoever. While this means that the company is counterproductive and should be in production, it can still offer insights to the market as a whole. In the case shown above, the  $x$ -intercept of the regression depicts that the minimum percentage companies spend on advertising in the consumer apparel industry, regardless of their gross margin, is 0.05% of their total sales revenues – practically negligible. In theory, this shows a level of accuracy, as companies that are not profitable (or are not in operation) do not spend on developing their brand recognition.

#### 4.2.2 Consumer Apparel Industry, Intellectual & Creative Rights



**Graph 4**

**Table 4**

Description	Value
Regression	$0.065(\ln(x)) + 0.810$
Correlation Coefficient ( $r$ )	0.677
Variance ( $r^2$ )	0.459
$S_x$ R&D expense	2.19%
$\bar{x}$ R&D expense	2.66%
$S_x$ gross margin	8.89%
$\bar{x}$ gross margin	55.30%

The correlation between intellectual & creative rights and the Veblenian value of the goods in the clothing industry is a moderate and positive one. The regression of 0.677 shows that while there is some level of correlation between the two, it is not precise. The variance of this data set is 0.459, or a 45.9% explanation value, showing that the majority of the fluctuations in the graph are not a causal byproduct.

The regression of best fit is a natural logarithmic function, which is proven by the difference in the variance of best fitted linear relationship, 0.362, more than a 16% reduction in the explanation value. While the leading coefficient of the function is a relatively small number in comparison to those in previous graphs, 0.065, the average percentage spent on R&D in this industry is also a relatively small proportion, 2.66%. Nonetheless, approximately 6.5% of a change in the gross margin is caused with a 1% change in the R&D expenditure.

The spread of the R&D spending is significant. As shown in Table 4, the standard deviation is nearly as significant as the mean. This is also shown by Graph 4, which has a positive skew (as the anomalies lie towards the right); apart from the cluster around  $x = 1\%$ , the data is widely spread. The  $y$  values, on the other hand, maintain a smaller spread, as previously demonstrated in Graph 3. In fact, the spread in this graph, Graph 4, is smaller, as the standard deviation makes up only 16% of the mean instead of the 23% in Graph 3. This implies that the sample for the data collection was not entirely random, thus impacting the accuracy of the data. The error bars support this idea, as they have varying degrees of limb extensions – an observation that also entails random errors and unreliability of the data.

However, the  $x$ -intercept does seem to yield an accurate value. The  $x$  intercept occurs when the research & development value is 0.000004, practically negligible. In theory, this shows accuracy, as firms that do not make money should not spend on innovation: the regression should pass through the coordinate (0,0). However, as the reliability of the data is uncertain, this detail does not ensure that this notion is correct.

## 5. Insights & Analyses

The applications of the data are derived from the prior discussion on its accuracy and reliability. As such, the insights for the different industries differ, and empirical analyses on the data per se are extended.

### 5.1 Insights & Analyses, Consumer Electronics Industry

In the consumer electronics industry, the impact of intangible assets on a firm's position on the Veblen spectrum is primarily attributed to the intellectual & creative rights they own. The brand recognition does not have a sizable impact — as a matter of fact, the gradient depicted in Graph 1 implies that as the intangible value of brand recognition for a firm increases, their Veblen status slightly decreases; this is neither accurate nor reliable, as the strength of the correlation is very weak. In contrast, the intellectual & creative rights not only have a strong correlation, but also the regression is within reason, and it is justified by the hypothesis.

The sum of the variances in Table 1 and Table 2 does not equal 100%. This discrepancy is because there are more than two factors of intangible assets. In this scenario, the sum of the variances is 0.753, or 75.3%. This means that approximately a quarter of value of the intangible assets consist of errors in the data collection, other external factors pertaining to the firm, or external factors pertaining to the volatility of the market. Although this may seem to be a large portion, a causation value of over 50% shows that the impact that these variables have comprises the majority.

The extent of the impact of the intellectual & creative rights demonstrates their importance for a firm in the electronics industry. Companies in this industry wanting to charge a premium for their products should look to innovate and advance their technology and obtain new USPs. Consumers value attributional factors over operational and utility ones. On the other hand, corporations that overspend on improving the status of their brand should consider a budget restructuring. Understanding the main factors that impact their ability to charge premiums without losing sales is imperative for a successful firm.

### 5.2 Insights & Analyses, Consumer Apparel Industry

The apparel industry sees a more balanced impact between the factors of brand recognition and intellectual & creative rights. Both curves follow a positive logarithmic function and are somewhat reliable. Nonetheless, due to the higher level of reliability and accuracy, the effect of the Veblenian value of the firm is mainly attributed to its brand.

Although the sum of the variances in Table 3 and Table 4 does not equate to 100%, they show a smaller deviation. The sum of the variances in these tables equates to 1.02, or 102%. This value exceeding the 100% benchmark not only demonstrates a degree of systematic and random errors, but also endorses some level of interdependence between the two variables. There might be influences that are common between the two factors. Adding on to the other factors that constitute the umbrella of intangible assets, the prior statements could account for a sizable segment of the error.

Despite the branding of a firm outweighing their innovation, companies should strike a balance between the two to optimize their pricing. While a consumer values the utility factors more in their purchase decisions, operational and



attributional factors still partly contribute. However, instead of engaging in a broad industry, reducing the sample size by focusing on a specific niche in which the company is positioned may yield more accurate results. This understanding of consumer behavior is pivotal for firms.

## 6. Conclusion

The results of this paper align with its hypotheses. Despite having a causal relationship, different industries present different ratios of the factors that intangible assets have on Veblen goods. Firms looking to increase revenue and profitability by converting their goods to Veblen goods should be vigilant of the intangible factors that allow the others in their industry to succeed. Seeing as the data proves this for three distinctive industries (two analyzed in this study, and one cited in the introduction) — ones without significant interdependence — the theory should be applicable to every industry and niche.

Hence, by conducting market research and interpreting the patterns that such intangible assets produce on consumer behavior, companies should make decisions that alter their business models. The bottom line of a business is to earn profits, and the higher up the Veblen spectrum a company attains, the larger the response of the domino effect the business oversees: price increases alongside quantity demanded, revenues from sales increase and accrued expenses remain relatively constant and thus, profits increase. In sum, intangible assets are an intrinsic part of Veblen goods, but the influential factors within the canopy of these assets are divergent from company to company, from industry to industry.

## 7. References

### 7.1 Notes

[1] Note that not all goods that follow this curve are Veblen goods. There are other classifications of goods that have an inverse demand curve (i.e. Giffen goods).

[2] The value of a firm's *Gross Margin* equates to their *Total Revenues* minus their *Cost of Goods Sold (COGS)*.

[3] Many brands are the subsidiaries of larger corporations who own multiple brands in the same industry. If one such selected brand falls into this category, the data collected is of the parent firm, since financial statements of individual brands are unavailable.

[4]  $S_x$  is the sample standard deviation of the variable.

[5]  $\bar{x}$  is the sample arithmetic mean of the variable.

### 7.2 Scholarly References

Alina, Jurca Maria, and Plăiaș Ioan. "Schema Congruity – A Basis for Evaluating Ambient Advertising." *International Journal of Market Research*, vol. 49, no. 2007, pp. 191–209., doi:10.1177/147078530704900205.

Baranes, Avraham Izhar. "Financialization in the American Pharmaceutical Industry:



- A Veblenian Approach.” *Journal of Economic Issues*, vol. 51, no. 2, 2017, pp. 351–358., doi:10.1080/00213624.2017.1320895.
- Black, Fischer, and Myron Scholes. “The Pricing of Options and Corporate Liabilities.” *Journal of Political Economy*, vol. 81, no. 3, 1973, pp. 637–654., doi:10.1086/260062.
- Chan, Tat Y., et al. “Measuring the Lifetime Value of Customers Acquired from Google Search Advertising.” *SSRN Electronic Journal*, vol. 30, no. 5, 2011, pp. 757–944., doi:10.2139/ssrn.1446829.
- Greenhalgh, Christine, and Mark Rogers. “Intellectual Property, Competition and the Value of UK Firms.” *Economic and Management Perspectives on Intellectual Property Rights*, 2006, pp. 40–57., doi:10.1057/9780230504745\_3.
- Grimes, Anthony, and Philip J. Kitchen. “Researching Mere Exposure Effects to Advertising – Theoretical Foundations and Methodological Implications.” *International Journal of Market Research*, vol. 49, no. 2, 2007, pp. 191–219., doi:10.1177/147078530704900205.
- Gupta, Sunil, et al. “Modeling Customer Lifetime Value.” *Journal of Service Research, Research*, vol. 9, no. 2, 2006, pp. 139–155., doi:10.1177/1094670506293810.
- Investopedia. “Who Discovered the Law of Supply and Demand?” *Investopedia*, Investopedia, 28 Aug. 2020.
- Koh, Yoon, et al. “Impact of Brand Recognition and Brand Reputation on Firm Performance: U.S.-Based Multinational Restaurant Companies’ Perspective.” *International Journal of Hospitality Management*, vol. 28, no. 4, 2009, pp. 620–630., doi:10.1016/j.ijhm.2009.04.002.
- Korda, Aleksandra Pismanik, et al. “Antecedents and Outcomes of Perceived Service Value: Evidence From a Transition Economy in Central Europe.” *Ekonomie a Management*, Jan. 2012, pp. 105–115.
- Lee, Huang Ning, et al. “An Empirical Analysis of Brand as Symbol, Perceived Transaction Value, Perceived Acquisition Value and Customer Loyalty Using Structural Equation Modeling.” *Sustainability*, vol. 11, no. 7, 2019, p. 2116., doi:10.3390/su11072116.
- Marshall, Alfred. “Equilibrium of Normal Demand and Supply.” *Principles of Economics*, vol. I, pp. 281–291, <https://doi.org/10.1057/9781137375261>.
- Oh, Haemoon. “The Effect of Brand Class, Brand Awareness, and Price on Customer Value and Behavioral Intentions.” *Journal of Hospitality & Tourism Research*, vol. 24, no. 2, May 2000, pp. 136–162., doi:10.1177/109634800002400202.
- Prychitko, David, et al. “Marxism.” *Econlib*, [www.econlib.org/library/Enc/Marxism.html](http://www.econlib.org/library/Enc/Marxism.html).
- Routledge. *Journal of Behavioral Finance*, vol.8, no. 3, 2007, doi:10.1207/jpbfm.2007.8.issue-3.
- Sandner, Philipp G., and Joern Block. “The Market Value of R&D, Patents, and Trademarks.” *Research Policy*, vol. 40, no. 7, Apr. 2011, pp. 969–985., doi:10.1016/j.respol.2011.04.004.
- Stephan, Paula “The Economics of Science,” *Journal of Economic Literature* vol. 34, no. 3, 2013, p. 1765 - 1774.
- Veblen, Thorstein, and Stuart Chase. *The Theory of the Leisure Class: An Economic Study of Institutions*. New York, N.Y: Modern library, 1934. Print.
- Villas-Boas, J. Miguel. “Consumer Learning, Brand Loyalty, and Competition.” *Marketing Science*, vol. 23, no. 1, 2004, pp. 134–145., doi:10.1287/mksc.1030.0044.

### 7.3 Data Collection References

- Acer, Inc. *Form 10-K 2014*, 2015. Web. 2015.  
Acer, Inc. *Form 10-K 2015*, 2016. Web. 2016.  
Acer, Inc. *Form 10-K 2016*, 2017. Web. 2017.  
Acer, Inc. *Form 10-K 2017*, 2018. Web. 2018.  
Acer, Inc. *Form 10-K 2018*, 2019. Web. 2019.  
Adidas AG. *Form 10-K 2014*, 2015. Web. 2015.  
Adidas AG. *Form 10-K 2015*, 2016. Web. 2016.  
Adidas AG. *Form 10-K 2016*, 2017. Web. 2017.  
Adidas AG. *Form 10-K 2017*, 2018. Web. 2018.  
Adidas AG. *Form 10-K 2018*, 2019. Web. 2019.  
Alphabet, Inc. *Form 10-K 2014*, 2015. Web. 2015.  
Alphabet, Inc. *Form 10-K 2015*, 2016. Web. 2016.  
Alphabet, Inc. *Form 10-K 2016*, 2017. Web. 2017.  
Alphabet, Inc. *Form 10-K 2017*, 2018. Web. 2018.  
Alphabet, Inc. *Form 10-K 2018*, 2019. Web. 2019.  
Apple, Inc. *Form 10-K 2014*, 2015. Web. 2015.  
Apple, Inc. *Form 10-K 2015*, 2016. Web. 2016.  
Apple, Inc. *Form 10-K 2016*, 2017. Web. 2017.  
Apple, Inc. *Form 10-K 2017*, 2018. Web. 2018.  
Apple, Inc. *Form 10-K 2018*, 2019. Web. 2019.  
Blackberry, Ltd. *Form 10-K 2014*, 2015. Web. 2015.  
Blackberry, Ltd. *Form 10-K 2015*, 2016. Web. 2016.  
Blackberry, Ltd. *Form 10-K 2016*, 2017. Web. 2017.  
Blackberry, Ltd. *Form 10-K 2017*, 2018. Web. 2018.  
Blackberry, Ltd. *Form 10-K 2018*, 2019. Web. 2019.  
Foot Locker Retail, Inc. *Form 10-K 2014*, 2015. Web. 2015.  
Foot Locker Retail, Inc. *Form 10-K 2015*, 2016. Web. 2016.  
Foot Locker Retail, Inc. *Form 10-K 2016*, 2017. Web. 2017.  
Foot Locker Retail, Inc. *Form 10-K 2017*, 2018. Web. 2018.  
Foot Locker Retail, Inc. *Form 10-K 2018*, 2019. Web. 2019.  
Huawei Technologies Co., Ltd. *Form 10-K 2014*, 2015. Web. 2015.  
Huawei Technologies Co., Ltd. *Form 10-K 2015*, 2016. Web. 2016.  
Huawei Technologies Co., Ltd. *Form 10-K 2016*, 2017. Web. 2017.  
Huawei Technologies Co., Ltd. *Form 10-K 2017*, 2018. Web. 2018.  
Huawei Technologies Co., Ltd. *Form 10-K 2018*, 2019. Web. 2019.  
HTC, Corp. *Form 10-K 2014*, 2015. Web. 2015.  
HTC, Corp. *Form 10-K 2015*, 2016. Web. 2016.  
HTC, Corp. *Form 10-K 2016*, 2017. Web. 2017.  
HTC, Corp. *Form 10-K 2017*, 2018. Web. 2018.  
HTC, Corp. *Form 10-K 2018*, 2019. Web. 2019.  
Kering. *Form 10-K 2014*, 2015. Web. 2015.  
Kering. *Form 10-K 2015*, 2016. Web. 2016.  
Kering. *Form 10-K 2016*, 2017. Web. 2017.  
Kering. *Form 10-K 2017*, 2018. Web. 2018.  
Kering. *Form 10-K 2018*, 2019. Web. 2019.  
LG Electronics, Inc. *Form 10-K 2014*, 2015. Web. 2015.  
LG Electronics, Inc. *Form 10-K 2015*, 2016. Web. 2016.  
LG Electronics, Inc. *Form 10-K 2016*, 2017. Web. 2017.  
LG Electronics, Inc. *Form 10-K 2017*, 2018. Web. 2018.

LG Electronics, Inc. *Form 10-K 2018*, 2019. Web. 2019.

LVMH Moët Hennessy Louis Vuitton. *Form 10-K 2014*, 2015. Web. 2015.

LVMH Moët Hennessy Louis Vuitton. *Form 10-K 2015*, 2016. Web. 2016.

LVMH Moët Hennessy Louis Vuitton. *Form 10-K 2016*, 2017. Web. 2017.

LVMH Moët Hennessy Louis Vuitton. *Form 10-K 2017*, 2018. Web. 2018.

LVMH Moët Hennessy Louis Vuitton. *Form 10-K 2018*, 2019. Web. 2019.

Motorola, Inc. *Form 10-K 2014*, 2015. Web. 2015.

Motorola, Inc. *Form 10-K 2015*, 2016. Web. 2016.

Motorola, Inc. *Form 10-K 2016*, 2017. Web. 2017.

Motorola, Inc. *Form 10-K 2017*, 2018. Web. 2018.

Motorola, Inc. *Form 10-K 2018*, 2019. Web. 2019.

Nike, Inc. *Form 10-K 2014*, 2015. Web. 2015.

Nike, Inc. *Form 10-K 2015*, 2016. Web. 2016.

Nike, Inc. *Form 10-K 2016*, 2017. Web. 2017.

Nike, Inc. *Form 10-K 2017*, 2018. Web. 2018.

Nike, Inc. *Form 10-K 2018*, 2019. Web. 2019.

PVH, Corp. *Form 10-K 2014*, 2015. Web. 2015.

PVH, Corp. *Form 10-K 2015*, 2016. Web. 2016.

PVH, Corp. *Form 10-K 2016*, 2017. Web. 2017.

PVH, Corp. *Form 10-K 2017*, 2018. Web. 2018.

PVH, Corp. *Form 10-K 2018*, 2019. Web. 2019.

Razer, Inc. *Form 10-K 2014*, 2015. Web. 2015.

Razer, Inc. *Form 10-K 2015*, 2016. Web. 2016.

Razer, Inc. *Form 10-K 2016*, 2017. Web. 2017.

Razer, Inc. *Form 10-K 2017*, 2018. Web. 2018.

Razer, Inc. *Form 10-K 2018*, 2019. Web. 2019.

Samsung Group. *Form 10-K 2014*, 2015. Web. 2015.

Samsung Group. *Form 10-K 2015*, 2016. Web. 2016.

Samsung Group. *Form 10-K 2016*, 2017. Web. 2017.

Samsung Group. *Form 10-K 2017*, 2018. Web. 2018.

Samsung Group. *Form 10-K 2018*, 2019. Web. 2019.

Sony, Corp. *Form 10-K 2014*, 2015. Web. 2015.

Sony, Corp. *Form 10-K 2015*, 2016. Web. 2016.

Sony, Corp. *Form 10-K 2016*, 2017. Web. 2017.

Sony, Corp. *Form 10-K 2017*, 2018. Web. 2018.

Sony, Corp. *Form 10-K 2018*, 2019. Web. 2019.

TJ Maxx. *Form 10-K 2014*, 2015. Web. 2015.

TJ Maxx. *Form 10-K 2015*, 2016. Web. 2016.

TJ Maxx. *Form 10-K 2016*, 2017. Web. 2017.

TJ Maxx. *Form 10-K 2017*, 2018. Web. 2018.

TJ Maxx. *Form 10-K 2018*, 2019. Web. 2019.

Toshiba, Corp. *Form 10-K 2014*, 2015. Web. 2015.

Toshiba, Corp. *Form 10-K 2015*, 2016. Web. 2016.

Toshiba, Corp. *Form 10-K 2016*, 2017. Web. 2017.

Toshiba, Corp. *Form 10-K 2017*, 2018. Web. 2018.

Toshiba, Corp. *Form 10-K 2018*, 2019. Web. 2019.

Urban Outfitters, Inc. *Form 10-K 2014*, 2015. Web. 2015.

Urban Outfitters, Inc. *Form 10-K 2015*, 2016. Web. 2016.

Urban Outfitters, Inc. *Form 10-K 2016*, 2017. Web. 2017.

Urban Outfitters, Inc. *Form 10-K 2017*, 2018. Web. 2018.

Urban Outfitters, Inc. *Form 10-K 2018*, 2019. Web. 2019.

VF, Corp. *Form 10-K 2014*, 2015. Web. 2015.  
VF, Corp. *Form 10-K 2015*, 2016. Web. 2016.  
VF, Corp. *Form 10-K 2016*, 2017. Web. 2017.  
VF, Corp. *Form 10-K 2017*, 2018. Web. 2018.  
VF, Corp. *Form 10-K 2018*, 2019. Web. 2019.  
Xiaomi, Corp. *Form 10-K 2014*, 2015. Web. 2015.  
Xiaomi, Corp. *Form 10-K 2015*, 2016. Web. 2016.  
Xiaomi, Corp. *Form 10-K 2016*, 2017. Web. 2017.  
Xiaomi, Corp. *Form 10-K 2017*, 2018. Web. 2018.  
Xiaomi, Corp. *Form 10-K 2018*, 2019. Web. 2019.



# The *Danmu* Register and the Establishment of its Social Existence in a Youth-centered Speech Community in China

Xinlin Yu

**Author Background:** Xinlin Yu grew up in China and currently attends the International Department of The Affiliated High School of South China Normal University, in Guangzhou, China. Her Pioneer research concentration was in the field of anthropology and titled “Communication and Culture.”

## 1. Introduction



**Figure 1.** An example of danmu. This figure is a screenshot taken by the author (Pandapia, 2019).

*Danmu* (弹幕), or “bullet comment,” is a new type of video comment functioning on Chinese online video websites (Figure 1). These comments are scrolled down the screen synchronically as the users watch a video, and the ones sent at a certain point within the video are seen by the viewers as the video plays back that specific time. Such movements and synchronizations of texts provide users with the feeling of co-viewing, and they quickly gained popularity among the Chinese audience (Chen et al., 2015). Starting as a feature in the Chinese ACG (animation, comic, and game) community, *danmu* now appears on nearly all mainstream video websites in China. Public events such as cinema movies, online music concerts, sports game live streams, and the annual Spring Festival Gala also allow participants to send *danmu* through their electronic devices (Liu, 2020; Qin, 2014).

While *danmu* is now incorporated into many traditional recreational events, its participants remain mostly young people. According to past analyses by Chen (2015), Wu (2016), and Youku (2019), since all websites that support *danmu* give

users the choice to turn off the “scrolling comments” while watching online videos, the senders and the viewers of *danmu* comments mainly consist of ages 15 to 30. Categorized as an emerging Chinese youth subculture by linguists, anthropologists, and sociologists, *danmu* has become a new generational marker of Chinese society.

This paper aims to examine how the *danmu* register establishes its social existence in a youth-centered speech community in China. I begin with a general background of the Chinese *danmu*. Then, I turn to an in-depth analysis of the characteristics of the *danmu* speech repertoire that I identify. Finally, I examine the metapragmatic evaluations of *danmu* in Chinese society, exploring how such practices exclude adults and unite the youths. Henceforward, “the adults” refers to individuals that are equal to or above 30 years old, “the youths” refers to individuals that are under 30 years old, “*danmu* viewers” refers to those who watch *danmu* but may not send it, “*danmu* senders” refers to those who view *danmu* and post *danmu* frequently, “*danmu* users” refers to *danmu* viewers and *danmu* senders as a whole, “*danmu*-non-viewers” refers to those who choose to turn off the *danmu* function while watching online videos, and the “*danmu* interface” refers to the space in which the *danmu* viewers and the *danmu* senders “contact” each other (J. Li, 2017).

While most analyses of *danmu* in this paper are based on bilibili.com (哔哩哔哩, Bilibili), the biggest *danmu* website in China, some examples from other websites will also be examined. It should be noted that, although Bilibili started as a website exclusively for the ACG community, it now covers most video content from traditional media, including movies, teleplays, and variety shows. Therefore, there are no significant differences between the *danmu* on Bilibili and those on other websites.

Besides making references to previous research and analyzing examples from various *danmu* websites, I also conducted an online survey of the general public (survey questions included in Appendix A), through which a total of 416 responses were received (Table 1). The results of the survey are incorporated into different parts of the paper when appropriate.

**Table 1.** *General information from the survey. The survey question asking whether one is a danmu sender (question 8) was designed to appear only when a respondent reported themselves as a danmu viewer. Therefore, in this table, the count for danmu senders is included in the “Danmu viewer” column.*

User Type	Age Groups		
	The Youths	The Adults	Total
<b>Danmu viewer</b>	101	57	158
<b>Danmu-non-viewer</b>	58	200	258
<b>Total</b>	159	257	416

## 2. Background of Chinese *Danmu*

Although *danmu* is now quite popular in China, it did not emerge from China. In 2006, the Japanese video website nicovideo.jp (ニコニコ, Niconico) launched a

video comment feature that allowed the users to send comments that scrolled down the screen synchronically as the video proceeded. This new function quickly prevailed on the website and was referred to as “コメント (English: comment)” when it appeared in a small amount. However, when there was a huge amount of such “scrolling comments” on the screen, looking like a barrage of bullets shooting across the sky, people started to call it *danmaku*, “弹幕 (だんまく),” which literally means “bullet curtain” in Japanese (Chen, 2015).

As Niconico initially focused on providing videos of the ACG culture, *danmaku* became popular among the ACG community in Japan. Soon, this trend spread its influence across the sea: in June 2007, the Chinese website acfun.cn (AcFun) was created, imitating the models of Niconico and introducing *danmaku* to the Chinese ACG community. Since *danmaku*, “弹幕,” is originally written in *kanji* (Japanese Chinese characters), its users in China translated it directly to *danmu*, “弹幕,” which also means “bullet curtain” in Chinese. However, while Japanese users exclusively use “*danmaku*” as a visual description for a huge amount of “scrolling comments” on the screen, Chinese users use “*danmu*” as an equivalent to “scrolling comments,” regardless of the amount of the text scrolling down the videos.

The fast development of *danmu* came two years after its first introduction to China. In 2009, due to conflicts between the management, AcFun went into a month-long shutdown, leading to the establishment of MikuFans, a new *danmu* website that changed its name to Bilibili in January 2010. Having similar functions and audiences, the competition between AcFun and Bilibili fast-forwarded the general development of *danmu* within the Chinese ACG community (Guo, 2015).

Along with the emergence of other ACG-related *danmu* websites, such a unique comment feature was soon brought to the general public. In 2012, tudou.com (土豆) became the first Chinese mainstream video website that supported *danmu*, and it was followed by three other mainstream websites in 2014 (Chen, 2015). Currently, *danmu* features on nearly all Chinese mainstream video websites and is combined with various traditional entertainment such as movies, concerts, and even novel reading. The spreading of *danmu* into mainstream platforms has also transformed the original *danmu* websites. Starting in 2013, Bilibili cooperates with mainstream video websites such as iqiyi.com (爱奇艺, Iqiyi), introducing diverse video content to the former ACG community (Guo, 2015). Today, Bilibili is no longer considered an ACG-exclusive website, but it covers video content ranging from daily vlogs to discovery documentaries. The mutual development of mainstream and *danmu* video websites greatly contributed to the growing familiarity of *danmu* among the Chinese public.

As more and more people recognize *danmu* as a new way of video commenting, many confuse it with Chinese youth slang or internet slang. Since it originates from the ACG community – a largely young community – and only exists on the Internet, the *danmu* register indeed resembles some characteristics of Chinese youth slang or internet slang, with overlapping terminologies and similar social domains. However, when examined closely, *danmu* is quite distinctive from both registers. According to Eble (1996), youth (college) slang takes place when there is a need to socialize, such as breaking the ice during awkward moments or responding to compliments, usually in conversations of more than two people. However, since *danmu* comments are anonymously shown and can only be read by later *danmu* viewers, the *danmu* senders are not in direct conversation with other users. As a result, there is no need for *danmu* to serve as an icebreaker or an immediate response, which deviates from Chinese youth slang. Moreover, *danmu* is more exclusive than

typical Chinese internet slang. In the list “2019 top 10 Chinese Internet Buzzwords” released by the National Center for Language Resources Monitoring and Research, six out of ten words are descriptive of actions and elements in daily life, such as safety rules, white-collar working mode, and housekeeping ideals (Huanqiu Web, 2019; Qiu, 2020). Since *danmu* comments are sent synchronically as the video proceeds, many of which refer exclusively to the video content, they are more evaluative than typical Chinese internet slang. Therefore, *danmu*'s unique features of time-lapsing communications and video-comment connections open up an alternative linguistic space for exploration. In this paper, I consider *danmu* as a third register that is different from both Chinese youth slang and internet slang.

Though facing failures, besides China, the Japanese *danmaku* also exists in multiple other countries – such as South Korea, America, Germany, and Spain – and should be considered a transnational phenomenon (Koyinu, 2017; Loveridge, 2012; Okada, 2008). However, for this paper, I intend to focus on the register of Chinese (Mandarin) *danmu*, in which all analyses given are based on the “scrolling comments” sent and received by Chinese users on Chinese *danmu* websites.

### 3. Characteristics of *Danmu* Speech Repertoire on Chinese *Danmu* Websites

As mentioned in Section 2, I see *danmu* as an independent register that is distinguishable from both Chinese youth slang and Chinese internet slang. In this section, I identify and categorize four common characteristics of the *danmu* speech repertoire on Chinese *danmu* websites: abbreviations, homophonic translations, interactive designs, and evaluative expressions. The *danmu* expressions analyzed below mainly came from the survey responses and the annual *danmu* billboards released by Chinese video websites.

#### 3.1 Abbreviations

Abbreviations are one of the most frequently used *danmu* comments. There are three subtypes of abbreviation: English acronyms, Chinese *pinyin* acronyms, and Chinese *hanzi* abbreviations. The English acronyms used in the *danmu* register are no different from the English acronyms used by English speakers, involving common expressions such as OMG (“oh my god”), CP (“character pairing”), and BGM (“background music”). Therefore, I will not expand further on this particular type of abbreviation and will concentrate on Chinese *pinyin* acronyms and Chinese *hanzi* abbreviations.

The Chinese *pinyin* acronyms are a unique case of abbreviations. In a linguistic sense, as a logographic language, Mandarin Chinese cannot be turned into acronyms. However, the Chinese romanization system, *pinyin*, transcribes Chinese characters into Latin letters and enables the production of acronyms (Chen, 2014). An example of a Chinese *pinyin* acronym is shown below (Bilibili, 2019):

“awsl”	a	w	s	l	“Ah, I’m dead.”
	a	wo	si	le	
	啊	我	死	了	
	Ah	I	die	-ed	



**Note.** When someone sees something really cute or attractive, they use the Chinese pinyin acronym of “awsl,” “ah, I’m dead,” to express their excitement.

Among *danmu* participants, the use of the Chinese *pinyin* acronym is very popular. In the 2019 *danmu* billboard released by Bilibili, “awsl” was entitled as “the *danmu* of the year,” sent 3, 296, 443 times by the website users (“2019 Bilibili,” 2019). Other Chinese *pinyin* acronyms include “xswl (xiao si wo le, 笑死我了, ‘ah, I laugh to death,’ describing something really funny),” “szd (shi zhen de, 是真的, ‘it’s true,’ used when someone sees intimate interactions between their imaginary celebrity couples),” and “gg (gan ga, 尴尬, ‘awkward,’ used when users feel awkward for the characters in the video or when watching an awkward scenario).” Generally, *danmu* senders use Chinese *pinyin* acronyms for efficiency, sometimes for the avoidance of sensitive politics or certain people’s full names.

Another type of abbreviation in the *danmu* register is Chinese *hanzi* (Chinese character) abbreviations. Usually, users select certain characters from a sentence to make a phrase that expresses the same meaning as the sentence. However, since the Chinese language is complicated in its part of speech tagging and the abbreviated sentences are largely colloquial, there is no general structure for such abbreviations. Below is an example of a Chinese *hanzi* abbreviation, selected from “2018 top 10 *Danmu*” released by Bilibili (Guanchazhe Web, 2018):

“颜表立”	颜色	表明	立场	“My color shows my stance”
	color	show	stance	

**Note.** On Bilibili, the senders are allowed to change the colors of their *danmu*, and there are several usages of “颜表立.” For one, when there is a conflict between characters in the video content, the fans of a certain character would send the phrase “颜表立,” “my color shows my stance,” in the character’s representative color to show support. Secondly, in Chinese color holds different connotations (e.g. green connects with cheating on a partner, yellow connects with sexual interactions). When the users send “颜表立” in these colors, they evaluate the behaviors in the video content.

*Danmu* senders create other Chinese *hanzi* abbreviations such as “注孤生 (注定孤独一生, ‘you are determined to be alone for your whole life,’ used when someone in the video content does something so unacceptable such as forgetting about their girl/boyfriend’s birthday)” and “气抖冷 (气得我浑身发抖, 大热天的全身冷汗手脚冰凉, ‘I am so angry that in hot summer my body shakes and feels cold,’ used when someone sees something that deviates from their moral values such as when a feminist saw a video about a girl that spends all of her salary on raising her younger brother).” Similar to Chinese *pinyin* acronyms, people use Chinese *hanzi* abbreviations for efficiency.

### 3.2 Homophonic Translations of Foreign Languages

Homophonic translations, by definition, refer to the practice of borrowing the sound of a certain text in one language and translating it directly into another form of text, usually in another language (Hilson, 2013). It is often the case that the new forms share the same sound as the original forms, but have deviated meanings. In the

*danmu* register, although there are homophonic translations from *hanzi* with similar *pinyin* forms and dialects, the dominant kind of homophonic translations are from foreign languages. Therefore, in this section, I will focus on the homophonic translations of foreign languages within the *danmu* register.

Among *danmu* users, there is a specific term for such homophonic translation of foreign languages, the so-called *konger* (空耳). Like the term *danmu*, *konger* also comes from a Japanese word. When describing misheard lyrics, Japanese people use the term “空耳 (そらみみ, *soramimi*).” As it is also written in Japanese *kanji*, Chinese *danmu* users translate directly to “空耳 (*konger*)” in Chinese. Initially, *konger* referred to the homophonic translation of a foreign language in lyrics; however, as *danmu* covered more and more diverse video content, it also now refers to translations of a character’s lines or certain nouns. Below, I examine two popular *kongers* in Chinese *danmu*.

The first came from the German movie *Der Untergang* (English: *Downfall*). In the movie, as Hitler gets angry for losing the war, he shouts, “Und doch habe ich allein, allein auf mich gestellt, ganz Europa erobert (‘And yet I alone, on my own, conquered all of Europe!’)” The line “Und doch habe ich allein” is translated homophonically by *danmu* senders (Aiya Liule, 2019):

Und	doch	habe	ich	allein
wo	dao	Hebei	sheng	lai
我	到	河北	省	来
I	come	Hebei	province	-ed

“I came to Hebei province.”

“I came to Hebei province;” *danmu* senders make the joke that this line is a remark by a tourist who is visiting the Hebei province in China.

The second example is from the Japanese animation *魔法少女まどか☆マギカ* (English: *Puella Magi Madoka Magica*). In the animation, the main character Kaname Madoka becomes a “魔法少女 (English: magical girl).” The talented *danmu* senders translated the term “魔法少女” by sound (*Mofa shaonv xiaoyuan*, n.d.):

魔	法	少	女
ま	ほう	しょう	じょ
ma	hou	shou	jo
ma	hou	shao	jiu
马	猴	烧	酒
horse	monkey	burn	wine

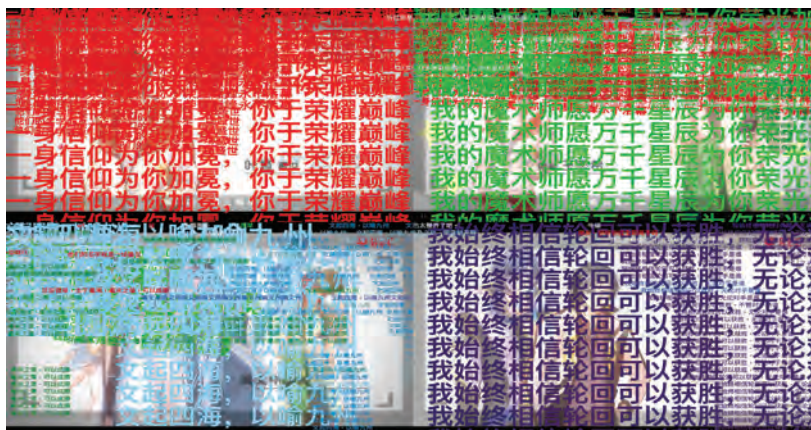
“Horse and monkey burn wine together”

“Horse and monkey burn wine together;” this line makes no sense in English, and it does not make sense in Chinese either. *Danmu* senders create this homophonic translation for the fun of its similar pronunciation to Japanese and its funky combination of a horse and a monkey.

As shown by the above examples, Chinese *kongers* do not always have semantic meanings. Even when they do, they have meanings that are completely off-track. However, these homophonic translations not only give the senders an opportunity to re-create what they hear, but also provide the viewers with entertainment and humor (S. Li, 2017). Unlike abbreviations that are used for efficiency, in the *danmu* register, homophonic transliterations of foreign languages are tools for enjoyment and relaxation.

### 3.3 Interactive Designs

As *danmu* viewers can only see “scrolling comments” sent by other users from a previous point in time, the *danmu* comment feature does not provide a platform for users to communicate directly and synchronically. However, certain *danmu* comments involve content that responds to an earlier *danmu* or that aims for responses from other *danmu*. Some *danmu* senders also enjoy sending the same comments multiple times and “speaking” to those who watch the video after them as an early-comer. In this section, I collectively identify such in-direct communications within the *danmu* register as “interactive designs.” I will discuss two important types of such interactive designs: spamming and conversation-like comments.



**Figure 2.** An example of fan spamming on Bilibili. This figure is a screenshot taken by the author (Fengxue Huanghun, 2015).

Spamming, according to Merriam-Webster, is defined as “unsolicited usually commercial messages sent to a large number of recipients or posted in a large number of places” (“Spamming,” n.d.). In the *danmu* register, however, “spamming” refers to the phenomenon in which a large amount of the same *danmu* appears at the same time in the same video. Generally speaking, spamming takes place when different *danmu* senders post the same *danmu* to show their similar feelings while watching the videos. For example, in Bilibili’s tenth-anniversary celebration video released in June 2019, the senders spammed “哔哩哔哩-( ° - ° )つ口乾杯~ (English: Bilibili-( ° - ° )つ口 cheers~),” the website’s headline, to show their excitement (Bilibili *Danmu* Website, 2019). Moreover, when their favorite character or idol appears in the video content, fans of a certain animation, TV drama, novel, or K-pop group also spam their star’s classic line in the so-called representative colors to show support (Figure 2). By sending the same *danmu* as others and seeing repetitive comments on the screen, the *danmu* senders and viewers indirectly interact with each other, sharing the same

feelings and gaining a sense of belonging to a larger community. I will further examine such feelings in Section 4.5.

A more common type of “interactive design” in the *danmu* register is the conversation-like comment. In such comments, the content itself “speaks” to an imaginary audience. Appearing in both Bilibili’s “2018 top 10 *Danmu*” and Iqiyi’s “Iqiyi 2018 *Danmu* Buzzwords,” an example of a conversation-like comment is shown below (Guanchazhe Web, 2018; Iqiyi, 2018).

---

“前方高能”	前方	高	能	“Be aware of the video content ahead”
	Before	High	Energy	

---

**Note.** *In a video, when there’s a funny joke or a scary scene ahead, some users who have already watched the content would return to the point in time before it happens and send “前方高能,” “be aware of the video content ahead,” to prepare the late comers for the surprising content ahead.*

As the senders use “be aware of the video content ahead” to alert others, the viewers get a sense of someone reminding them while watching the video. In this way, both the senders and the viewers imagine themselves as participants in a conversation, and the interactions between users are established indirectly. Other examples of such conversation-like comments include “欢迎回来 (‘welcome back,’ used when the video content is so interesting that people want to watch it again. The senders of such comments anticipate others’ actions and “welcome” them to watch the video for the second time)” and “前面的人等等我 (‘the person ahead, wait for me!’ Used when a viewer sees a *danmu* comment on the screen that represents their current thoughts, agreeing with a former *danmu*).”

Both spamming and conversation-like comments enable in-direct communications between *danmu* senders and viewers. Like their name, interactive designs build a bridge between *danmu* users and contribute to the creation of a speech community (further discussion in Section 4).

### 3.4 Evaluative Expressions

As *danmu* are fundamentally comments, it is quite common for users to express their feelings or thoughts through posting *danmu* onto the screen. However, what distinguishes *danmu* from traditional comments is that the “scrolling comments” are shown synchronically with the video content. In many cases, the *danmu* comments *evaluate* an element in the videos, a character’s line, or an uploader’s action, expressing praise or disgust. In this section, I further divide such “evaluative expressions” into three different types: video-concentrated, video-self connected, and ironic expressions. I will provide specific examples to explain each.

The video-concentrated evaluative expressions refer to expressions that focus solely on the evaluation of a video’s content. It can be illustrated through the commonly-seen *danmu* “名场面 (‘famous scene’).” According to the responses of the survey, when the users watch a classic scene, a well-known scene, or an impressive scene, they will send the term “famous scene” as an evaluation. For instance, in the classic movie *The Sound of Music*, before the song “Do Re Mi,” many people post “famous scene” on the screen (*Yinyuezhisheng*, n.d.). In some cases, however, when the senders evaluate the video content as memorable, they do not directly post “famous scene,” instead, they post “合影 (‘take photos together’),” as if they are taking

photos in front of a place of interest (Figure 3). Similar evaluative expressions are “打卡” (‘check,’ similar to checking out places on one’s bucket list) and “世界名画” (‘a masterpiece of painting,’ implying that if the video scene was painted, it would become a masterpiece).”



**Figure 3.** *Danmu* senders on Bilibili “taking photos” with a classic scene from *The Great Gatsby*. This figure is a screenshot taken by the author (Liaobuqi de Gaicibi, n.d.).

The second kind of evaluative expression is the video-self connected expression, in which the viewers evaluate the video content by connecting what they see in the videos to their personal experiences or feelings. An example can be drawn from Bilibili’s “2018 top 10 *Danmu*,” the top 1 *danmu* of 2018—“真实 (‘true’ or ‘that’s true’)” (Guanchazhe Web, 2018). As reflected from the survey results, many respondents reported that they send the *danmu* “true” when they see video scenes that “resonate with their daily lives” or that “touch their hearts.” For example, in the Bilibili video “家长会——几家欢喜几家愁 (‘Parent’s Meeting—Some Family End Up Happy and Some End Up Sad’),” many viewers send “true” to the screen when the uploader imitates an iconic parent reaction during the Chinese parents’ meetings (Xijing Mudan, 2018). Such connections are further supported by the explanation of “true” given by the Annual *Danmu* Committee from the Chinese Academy of Social Science:

“true” has three layers of semantic meanings. One refers to the alignment with the reality, the second refers to the consistent with cognition, the third refers to the connection with the heart. (Guanchazhe Web, 2018)

When the users evaluate the video content to be similar to their personal experiences, or to be in accordance with their thoughts, they would send “true” to express agreement. In some cases, even when the video content deviates from reality but aligns with the viewers’ self-consciousness, they also send “true” to show approval. As put by Chu, a researcher at the Institute of Linguistics, Chinese Academy of Social Sciences and an expert from the Annual *Danmu* Committee:

Even though many works have exaggerated elements, if it aligns with their imagination, and connects with their feelings at heart, the young audience is still willing to use “true” in *danmu* to show their identification with the content. (Guanchazhe Web, 2018)

Some other examples of such video-connected evaluative expressions are “泪目 (‘tears in eyes,’ used when someone watches a scene that touches their heart)” and “妙啊 (‘fantastic!’ Used when the video plot goes in a surprising yet reasonable way. By relating to their personal experiences, users rationalize the surprises and are amazed at the creativity of the video maker).”



**Figure 4.** *Danmu* senders on Bilibili describing an annoyed kitten as “happy.” This figure is a screenshot taken by the author (Lianggeren Yitiaogou, 2020).

The last category of evaluative expressions is the ironic expression, which refers to the instances when users send *danmu* expressions that are opposite to their evaluations. For instance, the term “灵魂歌手 (‘soul singer’ or ‘spiritual singer’)” is normally used to describe someone who sings well and whole-heartedly, touching someone’s soul. In the *danmu* register, however, “spiritual singer” ironically refers to someone who is tone-deaf: someone who is singing their soul out but still sings badly. Another example of such an ironic expression is “你看他满脸写着高兴 (‘look, they write “happy” all over their face’).” The phrase literally means “look, they write ‘happy’ all over their face,” describing someone who is so happy that others can easily tell it from their facial expression. However, *danmu* senders generally use the term to evaluate negative emotion. As shown in Figure 4, though the kitten looks quite annoyed when the uploader celebrates her birthday, *danmu* senders still post comments such as “他看起来满脸高兴 (‘it looks happy in the face’)” and “一副开心的样子 (‘a happy look’)” onto the screen.

Whether they are video-concentrated, video-self connected, or ironic expressions, evaluative expressions are essentially the subjective ideas of the senders. By allowing such subjective evaluations, the *danmu* comment feature provides a platform for individuals to express themselves and let their voices be heard.

Although above I examined separately the characteristics of the *danmu* speech repertoire—abbreviations, homophonic translations, interactive designs, and evaluative expressions—it is worth noting that these characteristics are not independent of each other. Often, a *danmu* comment shows two or more such characteristics. For example, the abbreviation “awsl (‘ah, I’m dead’)” is also an evaluative expression that evaluates the video content as cute or attractive. In some cases when there is a large number of the “awsl” comments on the screen, it creates spamming, serving as a kind of interactive design. The characteristics of the *danmu* speech repertoire work in conjunction with each other and provide the *danmu* users



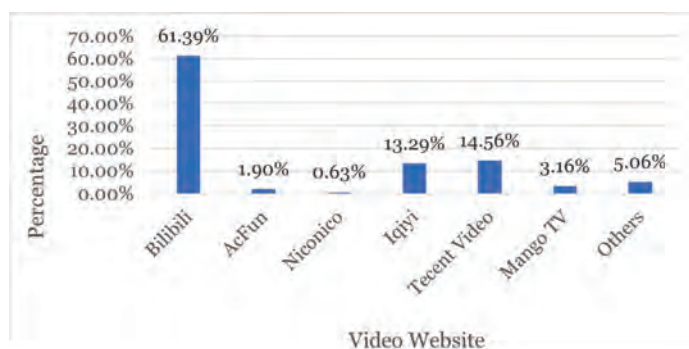
with an integrated experience of efficiency, entertainment, communication, and expression.

#### 4. Metapragmatic Evaluations of *Danmu*

According to Agha (2003), to describe a register, it is important to find a criterion of differentiation that distinguishes the register's repertoire from "other types of speech that occur in the same stretch of discourse." In doing so, linguists study people's metapragmatic evaluations of a certain register, sometimes through the language users' metalinguistic activities. In this section, I will examine the metapragmatic evaluations of the *danmu* register and explore how *danmu* is uttered by a youth-centered speech community in China.

##### 4.1 The *Danmu* Manner Test

The most explicit kind of metapragmatic evaluation is the metapragmatic discourse, the account describing the pragmatics of speech forms (Agha, 2003). In the *danmu* register, a representation of such metapragmatic discourses is Bilibili's *danmu* manner test. On Bilibili, the users are not allowed to send *danmu* until they have passed the so-called *danmu* manner test. Although there are no similar explicit regulations of *danmu* usage on other Chinese *danmu* websites, as an account on the platform that most surveyed individuals chose to turn on *danmu* while watching videos (Figure 5), Bilibili's *danmu* manner test reflects a common ideology and will therefore be the focus of examination in this section.



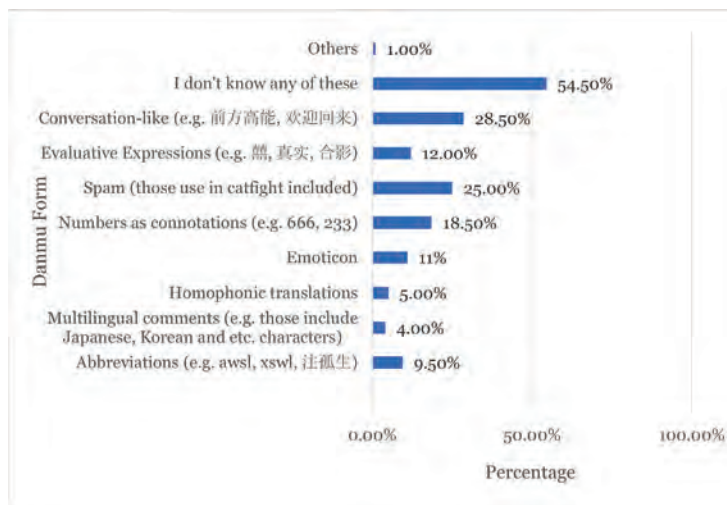
**Figure 5.** *Danmu* viewers' preference on which website to turn on *danmu*. This figure displays data collected from survey question 4.

The Bilibili *danmu* manner test is the first part of a longer membership registration exam for the website, the so-called "B-site Membership Promotion Test." In the test, the questions establish various imaginary scenarios for the test-takers, and ask them what they think they should do under different circumstances (e.g., when you see *danmu* "xxxx," what should you do?). Formerly, users were provided with 40 questions on the "*danmu* manner" section. Currently, however, whilst the website reforms the section name to "the community regulation test," the number of questions focusing particularly on "*danmu* manner" has decreased to about 20.

Examining the 2020 version on Bilibili, I find two prominent characteristics of the *danmu* manner test.<sup>1</sup>

First, the *danmu* manner test focuses on the establishment of a polite speech community. Asking the users what to do under different imaginary circumstances, the test reminds the users to avoid sending “junk” *danmu* (e.g. putting their ID names onto the screen and spamming their access date to the video) and malicious *danmu* (e.g. wishing someone’s family dead and attacking the uploaders for their work), encouraging the users to report any misbehaviors they notice. It also defines the appropriate *danmu* usage for the users, such as sending a “thank you” *danmu* for the uploaders at the beginning of the video, and “this is a good episode” or “excited for the next episode” at the end of the video. Such metapragmatic discourse is illustrated most clearly when test question 32 asks, “what happens when you use *danmu* to have a large amount of one-on-one discussion with your friends in a video?”, to which “this will cover the screen and will affect others’ watching experience” is the correct choice. The *danmu* manner test, in black and white, directly defines the standards of the appropriate use of the *danmu* register for the users, describing a polite community.

Secondly, the *danmu* manner test questions use the *danmu* speech repertoire to test their users. For example, test question 6 asks about *guichu* (“鬼畜,” referring to auto-tune remix-themed content) videos, and test question 28 mentions *konger* in the options. The option B of test question 17, “我上我也行 (‘if I do it I can do it too’),” also counts as an evaluative expression. Since such terminologies are only familiar to frequent *danmu* viewers or young people using the Internet, the adults who are not so familiar with popular culture or the Internet might easily fail the test (as shown in Figure 6, over half of the adult *danmu*-non-viewers had no idea about any form of *danmu* comments). Thus, the *danmu* manner test illustrates the ideology of which *danmu* is a practice for the youths.



**Figure 6.** Identification of *danmu* forms among adult non-*danmu*-viewers. This figure displays data collected from survey questions 1 and 15.

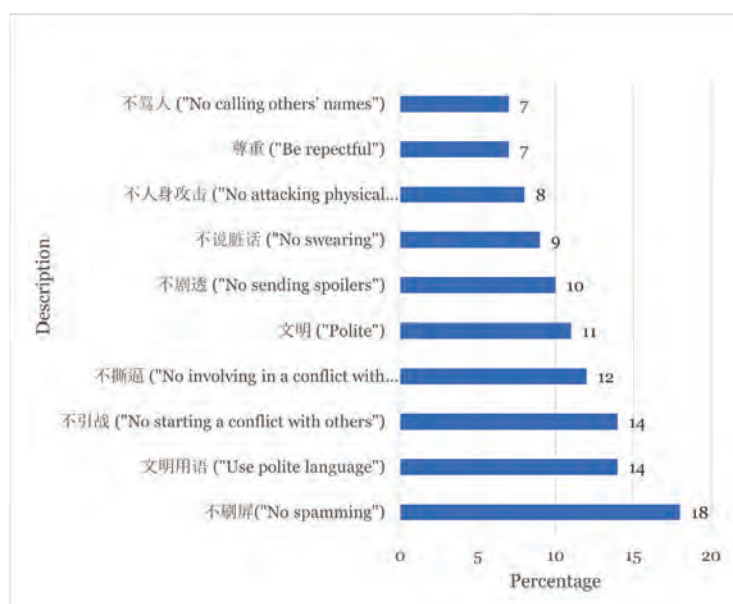
<sup>1</sup> The *danmu* manner test questions mentioned in this paper (test questions 6, 17, 28, and 32) are translated in Appendix B. As Bilibili randomly assigns test questions to its test-takers, the translated questions are just for reference and are numbered based on the particular test taken by the author.



#### 4.2 Politeness: The Metapragmatic Typification Among *Danmu* Language Users

As discussed above, the *danmu* manner test serves as a metapragmatic discourse that describes the *danmu* register as one that is polite and youth-centered. However, a register does not establish its social existence unless similar evaluative behaviors are exhibited by multiple individuals (Agha, 2003). Therefore, in this section, I examine the metapragmatic typification of politeness among *danmu* language users (see Section 4.3 for the discussion of *danmu* as a youths' practice).

The survey results imply that it is a common phenomenon for *danmu* users – both viewers and senders of posts – to value politeness in sending or interpreting *danmu* comments. According to the data collected, when asked to identify some *danmu* manners, frequent *danmu* viewers provided uniform answers that determined the politeness of *danmu* usage (Figure 7), with 98.1% of them maintaining that one should follow such *danmu* manners while sending a “scrolling comment.” However, such metapragmatic evaluation does not establish its existence outside the *danmu* speech community, as about 80% of the surveyed *danmu*-non-viewers did not even recognize the existence of the so-called “*danmu* manners.” Thus, it can be concluded that the metapragmatic typification of *danmu* as a polite register is an ideology that is exclusive to its language speakers.

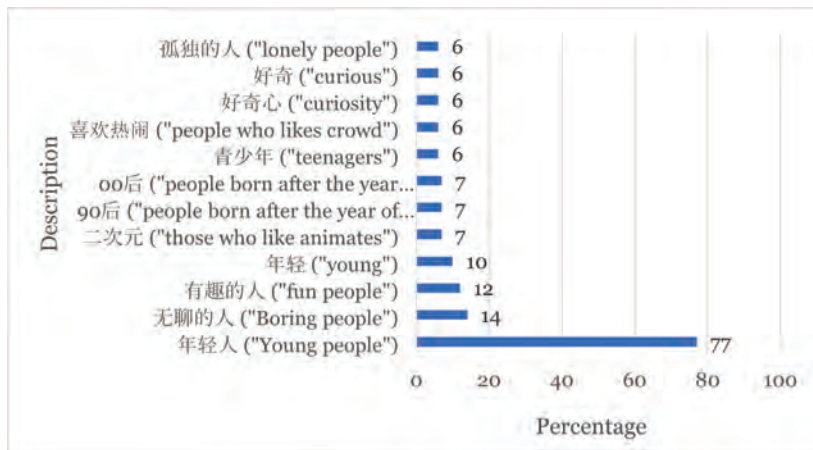


**Figure 7.** *Danmu* viewers' identifications of *danmu* manners. This figure displays data collected from survey question 11. The responses by *danmu* viewers were analyzed through a word frequency test

#### 4.3 The Practice of the Youths: Chinese Metapragmatic Stereotype of Use on *Danmu*

When a certain kind of metapragmatic typification recurs in the evaluation behavior of many speakers, it forms a “metapragmatic stereotype” (Agha, 2003). Hence, in

exploring the social existence of *danmu* as a youth-centered register, in this section, I examine the Chinese public's metalinguistic stereotypes on the users of *danmu* and its effect on keeping adults away from the *danmu* speech community.



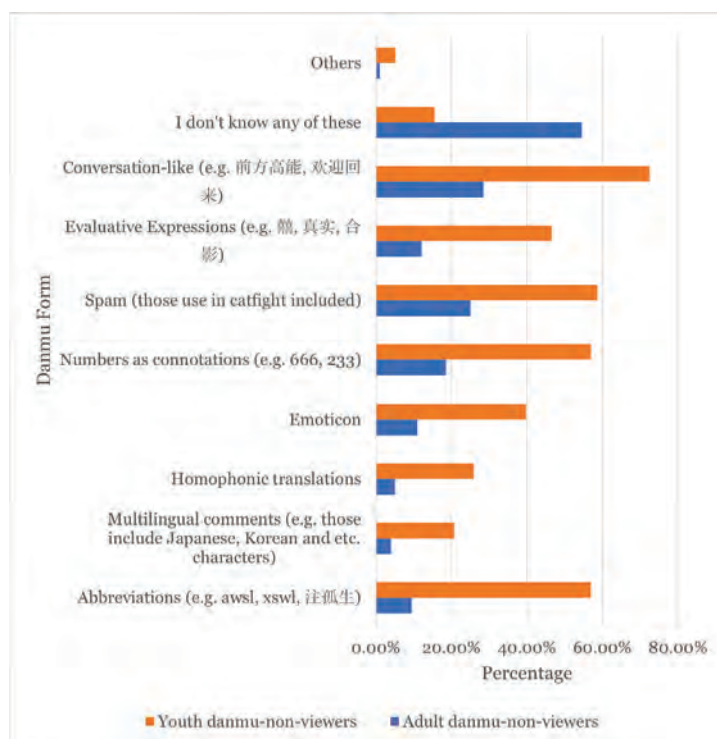
**Figure 8.** People's stereotype on *danmu* users. This figure displays data collected from survey question 20. The responses by all respondents were analyzed through a word frequency test.

The survey responses reveal a pattern of typification of *danmu* in Chinese society. As shown by Figure 8, when asked to describe what kind of people they think use *danmu* while watching videos, the respondents of all age groups provided descriptions that suggested *danmu* as a practice of the youths, confirming the metapragmatic stereotype of *danmu* users among Chinese people.

Although the video comment function of *danmu* can easily be accessed by any member of the Chinese society, such a metapragmatic stereotype contributes to *danmu*'s status as a generational marker in China. With the assumption that *danmu* is a youths' practice in mind, adults may be reluctant to participate in the register. According to the survey results, out of 200 adult *danmu*-non-viewers, 65.5% reported themselves not willing to learn about the *danmu* register even if there is a chance provided and most gave “不感兴趣 (‘not interested’)” as the explanation. Though there are many possible reasons for the adults to be “not interested” in the *danmu* register, a closer examination of the survey data suggested the age-related metapragmatic stereotype of *danmu* as an important factor, upon which I elucidate below.

In comparing the responses from adult and young *danmu*-non-viewers, an interesting phenomenon develops: like their adult counterparts, most young *danmu*-non-viewers are also reluctant to learn about the *danmu* register and provide “not interested” as their most common explanation. In other words, on an expression level, the youths and adults who are *danmu*-non-viewers reach a consensus on why they do not want to know about the *danmu* register. However, when considering content from other comments, a difference in ideology is revealed. While several adult respondents provided additional explanations such as “年纪大不喜欢 (‘I am old, I don't like it),” “觉得是小朋友玩的 (‘I think it is for little kids’),” and “离我生活远 (‘far away from my life’),” none of the young respondents gave age-related rejections, focusing instead on expressing their “laziness” to learn and the disgust they feel towards *danmu*. Moreover, the survey results from question 15 also show that adult *danmu*-non-

viewers are far less familiar with the *danmu* forms than their young counterparts (Figure 9). Such information then helps us shed some light on the different pragmatic meanings of “not interested” used by the youths and the adults. Whereas the young *danmu*-non-viewers are “not interested” in learning more about the *danmu* register because they already know something about it and find it meaningless, the adult *danmu*-non-viewers consider *danmu* to be a youths’ language that conflicts with their social identity, and they are “not interested” in learning anything about it. The metapragmatic stereotype of *danmu* in Chinese society hence acts as an exclusionary force that discourages the adults from entering the *danmu* speech community.

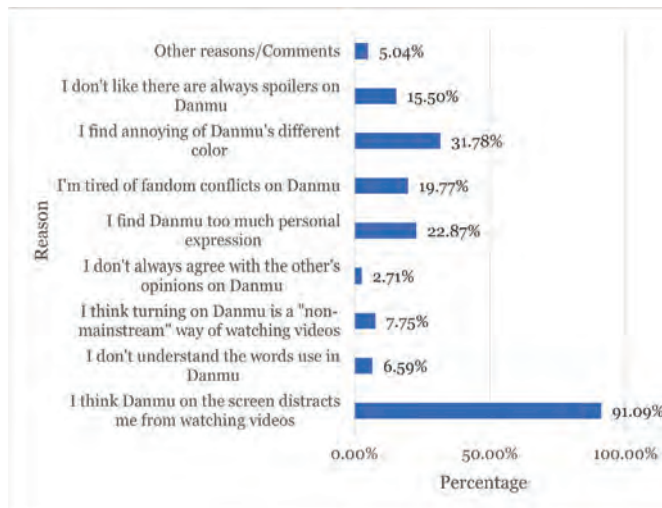


**Figure 9.** A comparison of the identification of *danmu* forms among adult and young *danmu*-non-viewers. This figure displays data collected from survey questions 1 and 15.

#### 4.4 *Danmu* as a Value-Boundary Register

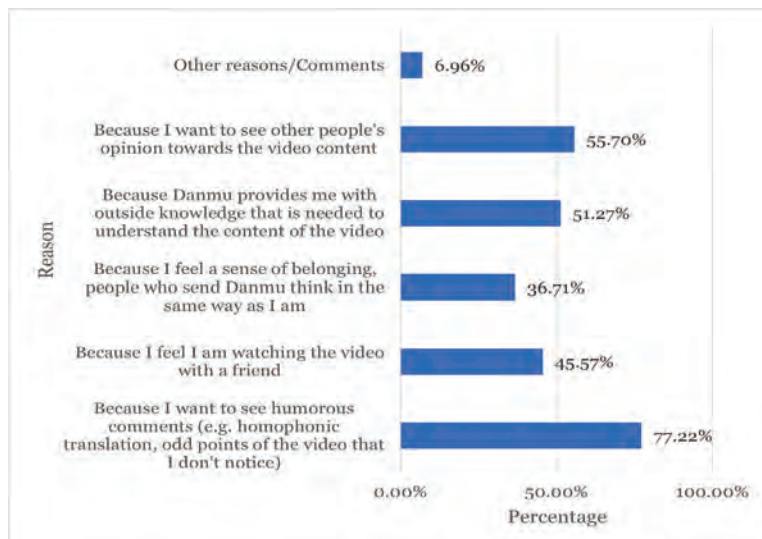
As all registers, in some sense, exist at a “value boundary” (i.e. receive both positive and negative evaluations from different evaluators), in this section, I explore the value boundary phenomenon in the *danmu* register (Agha, 2015).

According to the survey results, a negative evaluation of the *danmu* register is presented by the choice of the *danmu*-non-viewers. In describing the reasons for turning off *danmu*, many reported the *danmu* function to be annoying and distracting, and some considered it to be a non-mainstream way of watching videos (Figure 10). Further analyses show that such negative evaluations are largely the practice of the adults, as 77.54% of all *danmu*-non-viewers are equal to or over 30 years old and 77.82% of all surveyed adults are *danmu*-non-viewers.



**Figure 10.** Reasons for danmu-non-viewers to turn off danmu while watching videos. This figure displays data collected from survey question 13.

The collected data also reflected the positive evaluation of *danmu* from the *danmu* viewers. Enjoying the synchronic process of reading *danmu* comments while watching videos, *danmu* viewers regard such “scrolling comments” as a kind of entertainment, a platform for gaining knowledge, and a place of belonging (Figure 11). Contrary to the practice of negative evaluations, such positive evaluations on *danmu* have been revealed to be a practice of the youths as 63.52% of all young respondents are *danmu* viewers and 63.93% of *danmu* viewers are under the age of 30.



**Figure 11.** Reasons for danmu viewers to turn on danmu while watching videos. This figure displays data collected from survey question 6.

The general trend, according to which the *danmu* register mainly receives positive evaluations from the youths and negative evaluations from the adults, is crucial in describing *danmu* as a register that exists within a youth-centered speech community in China. The majority of adults, as they negatively believe the *danmu* register to be a deviation from the traditional non-distracting video watching experience, further exclude themselves from participating in the *danmu* register. For the youths, however, their usage of *danmu* generates a sense of community (discussed in Section 4.5), further involving themselves in the *danmu* speech community.

#### 4.5 Adults as “Bystanders”: The Youth-centered Speech Community on the *Danmu* Interface

Also shown by the data in Section 4.4, the general trend does not always apply to individuals—as there is a significant amount of adult *danmu* viewers and young *danmu*-non-viewers. In describing *danmu* as a register within a youth-centered speech community in China, I am not concerned with why some Chinese youths choose not to participate in the *danmu* register. Instead, I focus on adult *danmu* users’ role in the *danmu* speech community. As mentioned earlier in Section 3, the *danmu* speech repertoire holds characteristics that provide a source of entertainment, a platform of expression, and a bridge of interaction. In this section, through analyses of the *danmu* users’ positive evaluations towards such characteristics, I first examine how the *danmu* comment feature creates a sense of community for all users. Then, I turn to a discussion of how such “a sense of community” is perceived differently by young and adult *danmu* users.

As demonstrated by Figure 11 (in Section 4.4), one of the most important reasons for *danmu* viewers to turn on *danmu* while watching videos is to see the humorous comments that scroll down the screen—in other words, for the sake of entertainment. An interesting question, therefore, arose from such a positive evaluation: while watching an online video is already a form of entertainment (that is why the *danmu*-non-viewers find *danmu* distracting), what additional fun can such humorous comments provide? An interviewee’s (age 22, watches videos with *danmu* every day) commentary from Du and Cai’s (2019) research might give us a clue:

The netizens are so funny. *Danmu* comments always find some interesting points that I never thought of or never noticed. Sometimes the video plot is obviously a sad one, but the *danmu* comments would say something very funny or ironic. Suddenly, I don’t want to cry at all and the sad scene turns out to be very funny.

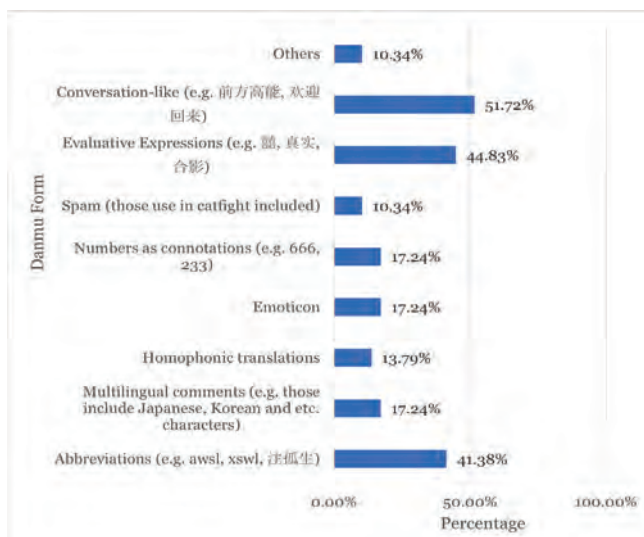
This remark reveals the different types of *entertainment* provided by the video content and the *danmu* comments. While the video content lets one relax and “satisfy general interest,” the *danmu* comments entertain the viewers by offering new perspectives (Luo, 2015). In this sense, when the *danmu* viewers report themselves watching *danmu* for the “humorous comments,” they are not looking for *danmu* content that is linguistically humorous, such as a joke or random displays of Chinese characters. Instead, the “humorous comments” they refer to are those that make sense of the video content in an unconventional or unexpected way. For example, the *konger* discussed in Section 3.2 is considered a kind of humorous comment. While the homophonic translations of foreign languages are quite accurate translations of the sound people hear from the video content, they produce unexpected meanings in

Chinese. Such a surprising yet reasonable way of translating brings a new angle of understanding to the *danmu* viewers, prompting them to burst into laughter, and ultimately be labeled as “humorous.” Another example is illustrated in Figure 12, where the bracketed *danmu* writes “冰箱里有空瓶子 (‘there’s an empty bottle in the fridge’).” Though the comment is stating a fact—as the arrow points out, there *is* an empty bottle in the fridge—the *danmu* viewers consider this a humorous comment for its unique perspective in viewing the video content.



**Figure 12.** An example for humorous comments. This figure is retrieved from a forum comment response by Bilibili’s official account (Bilibili, 2018).

With the *danmu* viewers’ awareness that such humorous comments are sent by other netizens (shown by the above interviewee’s comment), the act of being entertained whilst watching humorous *danmu* therefore becomes similar to the action of attending a workshop; in both cases, the participants encounter ideas from various addressers. As one surveyed *danmu* viewer commented, “everyone’s so interesting and funny, [I] always laugh my head off by [the comments of] the ‘little geniuses’ in *danmu*,” *danmu* as an entertainment indeed creates a sense of community among the *danmu* viewers.



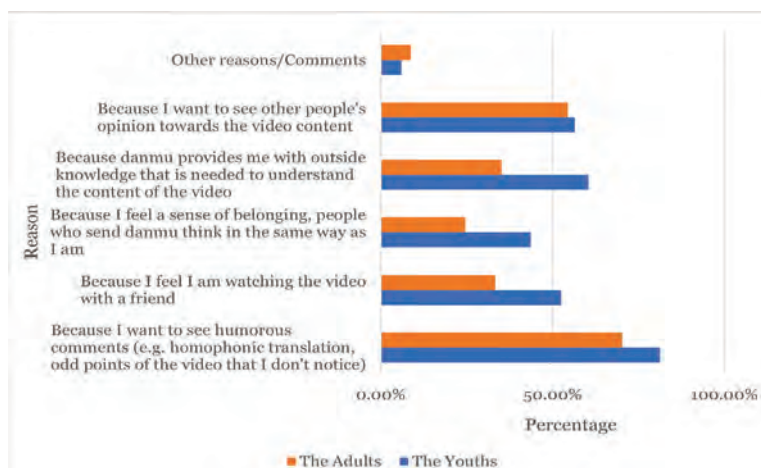
**Figure 13.** Danmu forms that the *danmu* senders like to send. This figure displays data collected from survey question 9.



Since *danmu* senders are at the same time *danmu* viewers, they do gain a sense of community by watching *danmu* comments as a form of entertainment; however, can the actions of sending *danmu* also generate a sense of community? I address this question below. According to Guo's (2015) research into the *danmu* senders' motive of posting comments, over half of the interviewees considered sending *danmu* to be a way of expressing themselves to others. They claim that if they are watching videos alone and there is something they want to share immediately about the video content, *danmu* is "a very good way of sharing." Such commentary is supported by the survey results (Figure 13). The surveyed *danmu* senders are reported to be especially in favor of posting *danmu* that involve conversation-like comments and evaluative expressions, positively valuing the *danmu* interface as a platform for communications and self-expression. Though *danmu* comments do not function like forum comments, where the users can have a direct conversation with one another, the interface of *danmu* creates a sense of a "backchannel" for its users (Ma & Cao, 2017). When the *danmu* senders post their personal opinions on the screen, they imagine themselves sharing ideas with an imitate group that holds the same interest as them, and hence a sense of community is created.

The fundamental reason for *danmu* users—regardless of their age—to feel a sense of community through viewing and sending *danmu* comments is the anonymous nature of the *danmu* register. Seeing such non-ID-attached comments scrolling down the screen, the *danmu* viewers imagine themselves listening to multiple addressers while the *danmu* senders imagine themselves speaking to a wide range of addressees (J. Li, 2017). As such, both the *danmu* viewers and the *danmu* senders consider themselves a member of a highly immersive community. However, if such an anonymous nature allows the participants to imagine their role in the *danmu* speech community, do youth and adult *danmu* users perceive their roles differently? I examine this question below by comparing the youth and the adult *danmu* users' metapragmatic evaluations of the *danmu* register.

Besides entertainment, the survey results reveal that youth and adult *danmu* users have different motivations for turning on *danmu* while watching videos. As shown by Figure 14, compared to their adult counterparts, young *danmu* viewers are more likely to use *danmu* as a tool to seek company and a sense of belonging.



**Figure 14.** A comparison of the reasons for turning on *danmu* among adult and young *danmu* viewers. This figure displays data collected from survey questions 1 and 6.

Such distinctions suggest the different roles that the youth and the adult *danmu* viewers take in watching videos with *danmu*. To borrow terms from Goffman's (1981) theory on the participation framework, while the youths consider themselves the "ratified participants" of the *danmu* speech community, the adults believe themselves to be the "bystanders" of the conversation between other viewers. Consider the difference in the following explanations given by the surveyed respondents (all answering survey question 5):

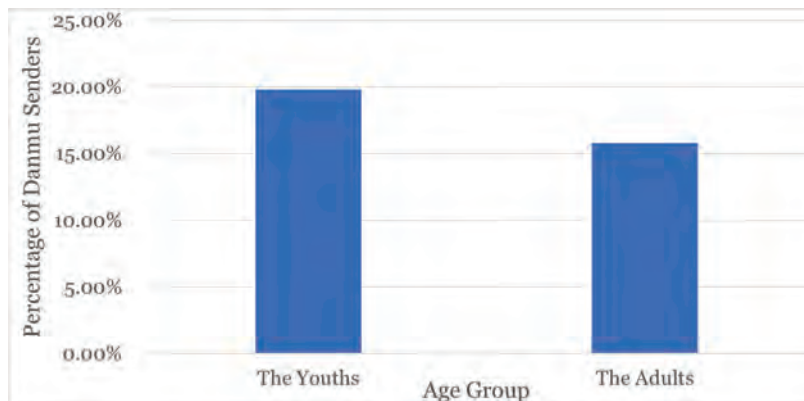
Feeling like [I] have interaction with others. [So that] it doesn't seem I'm watching the video alone. (age range 12-17)

Fun *danmu* comments can complement boring videos and [I] hope to look for friends with the same ideas as me in the *danmu* comments. (age range 12-17)

Looking at others' fantasies, sometimes the netizens have interesting comments. (age range 40-49)

B-site's (Bilibili) *danmu* comments are relatively more interesting, [I like to] see others' commenting on and interacting with each other. (age range 30-39)

From the commentaries, we can tell that, though for both age groups, the *danmu* viewers realize the existence of the *danmu* speech community on the *danmu* interface, but young *danmu* viewers take an active role and seek "friends" and "interaction with others," while the adults enjoy the process of learning about "others' fantasies" and viewing others' interactions with one another. Unlike the way in which their young counterparts imagine themselves in conversation with the *danmu* senders, the adult *danmu* viewers consider themselves "bystanders" of the speech community they notice.



**Figure 15.** A comparison of the percentage of *danmu* senders among adult and young *danmu* viewers. This figure displays data collected from survey questions 1 and 8.

While *danmu* viewers can either be "bystanders" or "ratified participants," in posting *danmu*, the senders imagine themselves as the "addresser" of a conversation, and therefore automatically take the role of a ratified participant. A question thus arose from such a role restriction: if the adult *danmu* viewers largely perceive themselves as the "bystander" of the *danmu* speech community, will they be less



willing to become *danmu* senders than their youth counterparts? Though only 18.35% of all surveyed *danmu* viewers report themselves to be (frequent) *danmu* senders, the survey results indicate a positive answer to the question (Figure 15). Meanwhile, Chen's (2015) research on Bilibili found most registered users on the website (i.e. permitted *danmu* senders. As Section 4.1 discussed, on Bilibili, one cannot send *danmu* until they pass the *danmu* manner test and become a registered user) belong to the age range of 15 to 25. A previous study by Guo (2015) also reveals a negative correlation between the *danmu* viewers' age and the amount of *danmu* they send, implying the peripheral role of the adults in the *danmu* speech community.

As Agha (2005) writes, "registers have a social existence only insofar as—and as long as—the metapragmatic stereotypes associated with their repertoires continue to be recognized by a criterial population, that is, continue to have a social domain." As such, the way that the adult *danmu* viewers read comments from the *danmu* speech repertoire but categorize themselves as "bystanders" – one surveyed adult *danmu* viewer directly referred to the *danmu* comments as "youths' thoughts" – enhances the social domain of "persons that can recognize the register's forms" and further describes *danmu* as a youth-centered register.

## 5. Conclusion and Discussion

The purpose of this paper is to investigate how the *danmu* register comes into existence in a youth-centered speech community in China. The analysis of the four dominant characteristics of the *danmu* speech repertoire reveals how *danmu*, as a similar yet distinctive register from both Chinese youth slang and Chinese internet slang, provides its users with an integrated experience of efficiency, entertainment, communication, and self-expression. The examination of the Chinese public's metapragmatic evaluations on *danmu* further demonstrates how such a youth-centered discourse is supported by "social voices linked to registers" (Agha, 2005): the *danmu* manner test, the metapragmatic typification of politeness among its language users, the age-related metapragmatic stereotype in the Chinese society, and the competing valorization between Chinese youth and adults. Moreover, the identification of the adult *danmu* viewers as "bystanders" in the *danmu* speech community discusses the expansion in *danmu*'s social domain and suggests how its "youthness" is reified in every associated utterance.

There are, indeed, some limitations. For one, as mentioned earlier, I do not discuss why some youths choose not to participate in the *danmu* register in this paper. Moreover, as the survey results suggest that, although Chinese female youths are more likely to become *danmu* viewers than their male counterparts, young female *danmu* viewers are less likely to become *danmu* senders than young male *danmu* viewers, I do not explore such gender differences here. Although I do not attempt to focus on such questions, those factors are also important to the description of the *danmu* register, and further research can be done.

Besides such limitations, in describing *danmu* as a register that exists in a youth-centered speech community in China, I hope to shed some light on how the Internet, the communications with other cultures such as the Japanese culture, and the Chinese society at a larger scale, have shaped the Chinese youth culture youth language. Most importantly, it is the youth themselves that brings the register to its existence, as one *danmu* user once commented (aged 22, uses *danmu* every day):

Speaking of the advantages that *danmu* websites have...maybe is [because] that people nowadays like to let themselves heard, and *danmu* happens to satisfy this point. (Du & Cai, 2019)

In an Internet era where most Chinese youths are the only child in the family, are given the choice to learn about a wide range of things, and are stressed with burdens from work and life, the *danmu* register provides youth with a way to escape from reality, to find “friends” with the same interest, and to perform their best selves (Zhao, 2018). Whether it is to spam for their favorite characters, to post their opinions about the video content, or to imagine themselves in-conversation with other viewers, Chinese youths enjoy themselves being in the center of the stage, with adult bystanders as part of their audience.

## References

- Agha, A. (2003). Register of Language. In A. Duranti (Eds.), *A Companion to Linguistic Anthropology* (pp. 23-45). Blackwell Publishing.
- Agha, A. (2005). Voice, Footing, Enregisterment. *Journal of Linguistic Anthropology*, 15(1), 38–59. <https://doi.org/10.1525/jlin.2005.15.1.38>
- Agha, A. (2015). Tropes of Slang. *Signs and Society*, 3(2), 306–330. <https://doi.org/10.1086/683179>
- Aiya Liule. (2019). [*Yuanshou de fennv*] *Wo dao hebeisheng lai...qi si ou lei!!!* [“The Fury of the Fuehrer” I came to Hebei province...I am so angry!!!][Video]. Bilibili.com. <https://www.bilibili.com/video/BV1Vt411p76U?t=128>
- Bilibili (2018). *Younaxie gaoxiao de danmu?* [What are Some Funny *Danmu* Comments?][Online forum post]. Zhihu.com. <https://www.zhihu.com/question/39463086/answer/522214731>
- Bilibili. (2019). *2019 Bilibili niandu danmu* [2019 Bilibili Annual *Danmu*]. Bilibili.com. <https://www.bilibili.com/blackboard/activity-danmaku2019.html>
- Bilibili *Danmu* Website. (2019). [*Ganbei shizhounian*] *bilibili shizhounian, ting wo jiang gushi quanchenghuiyu* [“Cheers for Tenth-Anniversary” Bilibili Ten Year Anniversary, Listen to My Story-telling Full Review]. Bilibili.com. <https://www.bilibili.com/video/BV1Px411d7pf?from=search&seid=739144898761173849>
- Chen, S. Y. (2014). From OMG to TMD – Internet and Pinyin acronyms in Mandarin Chinese. *Language@Internet*, 11, article 3.
- Chen, W. (2015). *Shipin wangzhan jiqi yonghu yanjiu—yi Bilibili danmuzhan weili* [Research on *Danmu* video sites and users—Take Bilibili video as an example][Unpublished manuscript]. School of Journalism & Communication, Nanchang University.
- Chen, Y., Gao, Q., & Rau, P.-L. P. (2015). Understanding Gratifications of Watching Danmaku Videos – Videos with Overlaid Comments. *Lecture Notes in Computer Science*, 153–163. [https://doi.org/10.1007/978-3-319-20907-4\\_14](https://doi.org/10.1007/978-3-319-20907-4_14)
- Du, X., & Tsay, W. (2019). Dalu *Danmu* wenhua jiqi qun zu texing yanjiu. *Journal of National Kaohsiung University of Science and Technology*, 1, 57–75. [https://doi.org/10.6912/JNKUST.201907\\_1.0004](https://doi.org/10.6912/JNKUST.201907_1.0004)
- Eble, C. (1996). *Slang & Sociability: In-group Language Among College Students*. The University of Northern Carolina Press.

- Fengxue Huanghun. (2015). [*Quanzhi gaoshou] bandao yingxiong* ["Quanzhi Gaoshou" Hero Half-way][Video]. <https://www.bilibili.com/video/BV1ss41127nh?from=search&seid=2397655062913868099>
- Goffman, E. (1981). Footing. In E. Goffman, *Forms of Talk* (pp. 124-159). University of Pennsylvania Press.
- Guanchazhe Web. (2018). *Bzhan fabu 2018 niandu danmu: "zhenshi" paiming diyi* [B-site Released its 2018 Annual Danmu: "True" Ranked the First]. (2018). Guancha.cn. <https://baijiahao.baidu.com/s?id=1620981293195870761&wfr=spider&for=pc>
- Guo, L. (2015). *Woguo danmu wangzhan de shouzhong yanjiu* [Research on the Audience of Barrage Video in China][Unpublished manuscript]. School of Chinese Language and Literature, Yunnan University.
- Hilson, J. (2013). Homophonic Translation: Sense and Sound. In H. J. Minors, *Music, Text, and Translation* (pp. 95-105). Bloomsbury Publishings.
- Huanqiu Web. (2019, December 5). *2019 Nian wangluo liuxingyu baodan zhaduichulu, nizhidaojige?* [2019 Internet Buzzword Billboard came out in cluster, how many do you know?]. Huanqiu.com. <https://baijiahao.baidu.com/s?id=1652033969274332548&wfr=spider&for=pc>
- Iqiyi. (2018). *Iqiyi 2018nian danmu reci fabu baokuan zizhi neirong cuisheng gaopin liuxingci* [Iqiyi's 2018 Danmu Hot Word is Released, Popular Self-made Content Promote High Frequency Buzzwords]. Iqiyi.com. <https://www.iqiyi.com/common/20190213/98b659be706b1312.html>
- Koyinu. (2017). [*Undertale] DRUGTALE heji (TVple danmufu)* ["Undertale" DRUGTALE Collection (TVple, Tvple.com)][Video file]. Bilibili.com. <https://www.bilibili.com/video/BV1Wx411S77j>
- Liaogeren Yitiaogou. (2020). *Jinjianceng mashu 2sui la, bama haidilao yuangong shasheng, mao: guai ganga de.* [*Lianggeren Yitiaogou*] [Felinae Mashu is 2 Years Old! Parents Celebrated the Happily. Cat: It's Kind of Award. "Two People One Dog"] [Video]. Bilibili.com. <https://www.bilibili.com/video/BV1954y1U7Ew>
- Liaobuqi de Gaicibi* [The Great Gatsby]. (n.d.). [Movie]. Bilibili.com. <https://www.bilibili.com/bangumi/play/ss28658/?from=search&seid=7546649483503139847>
- Li, J. (2017). The Interface Affect of a Contact Zone: Danmaku on Video-Streaming Platforms. *Asiascape: Digital Asia*, 4(3), 233–256. <https://doi.org/10.1163/22142312-12340079>
- Li, S. (2017). *Danmu yuyan tedian he biaoda yiyi* [The expression and significance of language features of barrage][Unpublished manuscript]. Department of Journalism & Communication, China West Normal University.
- Liu, H. (2020, April 1). *Danmu yinyuehui laile* [The Danmu Concert is Here]. Bjd.com.cn. <http://www.bjd.com.cn/a/202004/01/WS5e81b632e4b03cc9592db4ef.html>
- Loveridge, L. (2012, October 2). *English Niconico.com to Retire, Re-Open Under Main Site on October 17.* Anime News Network. <https://www.animenewsnetwork.com/news/2012-10-02/english-niconico.com-to-retire-re-open-under-main-site-on-october-17>
- Luo, Y. (2015). Motivation, Emotion, Attitude, & Gratification in the Use of Online video media [Unpublished manuscript]. Hong Kong Batisipt Uni [https://repository.hkbu.edu.hk/etd\\_oa/260](https://repository.hkbu.edu.hk/etd_oa/260)
- Ma, X., & Cao, N. (2017). Video-based Evanescent, Anonymous, Asynchronous Social Interaction. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 770–782. <https://doi.org/10.1145/2998181.2998256>

- Mofa shaonv xiaoyuan: di yi hua sihu zai mengli jianguo, nayang* [Puella Magi Madoka Magica: Episode 1, Seems Once Met in Dreams, Like That...]. (n.d.). [Animation] Bilibili.com. <https://www.bilibili.com/bangumi/play/ep63470>
- Okada, Y. (2008). *[/Nikoniko douga (natsu) / happyou toukute yu—za—no mi de douga kyoyuu su ru kinou na do tsuika* [Niconico Douga (summer) announcement: additional video sharing function for specific users]. ITmedia News. <https://www.itmedia.co.jp/news/articles/0807/04/news130.html>
- Pandapia. (2019). *Yiqi die xiongmao ya! Ruanruan de ke hao die le* [Let's Pile Up Pandas Together! They are so Soft and Easy to Pile Up!][Video]. Bilibili.com. <https://www.bilibili.com/video/BV16t411n7vJ?t=31>
- Qin, A. (2014, August 21). *Theaters in China Screen Movies, and Viewers' Text Messages*. Sinosphere Blog. <https://sinosphere.blogs.nytimes.com/2014/08/20/theaters-in-china-screen-movies-and-viewers-text-messages/>
- Qiu, S. (2019). 2019 Wangluo liuxingyu de xingcheng yu gongyong [The Formation and Usage of 2019 Internet Buzzwords in China]. *Gaodeng Jiaoyu*, 701, 58.
- Spamming. (n.d.). In *Merriam-Webster.com dictionary*. Retrieved from <https://www.merriam-webster.com/dictionary/spamming>
- Wu, X. (2016). *Canyu de shouzhong: danmuzu de shiyong yu manzu yanjiu* [Participative Audience: Research on Danmaku users' Uses and Gratifications][Unpublished manuscript]. School of Journalism and Communication, Shandong University.
- Xijing Mudan (2018). *Jiazhanghui—jijia huanxi jiajia chou* [Parent's Meeting—Some Family End Up Happy and Some End Up Sad][Video]. Bilibili.com. <https://www.bilibili.com/video/BV17t411o7XB?from=search&seid=1702046900366584577>
- Yinyuezhishang* [The Sound of Music]. (n.d.). [Movie]. Bilibili.com. <https://www.bilibili.com/bangumi/play/ss28658/?from=search&seid=7546649483503139847>
- Youku. *2019 Youku niandu baogao chulu: 95hou nvsheng doushi dianzan kuangren* [2019 Youku Annual Report is Released: Ladies Born after 1995 are "Like-holics"]. (2019, December 30). Youth.Cn. <https://baijiahao.baidu.com/s?id=1654321892000298924&wfr=spider&for=pc>
- Zhao, M. (2018). *Jiyu hudong shijiao xia de danmuzu ziwo rentong yanjiu* [Research on the self-identity of barrage family based on interactive perspective][Unpublished manuscript]. Journalism College, Shanxi University.

## Appendix A

**Note.** Questions 4-8 and 10-12 are designed to appear only when a respondent selects "yes" for question 3; question 9 only appears when a respondent selects "yes" for question 8. Questions 12-16 are designed to appear only a respondent selects "no" for question 3; question 17 only appears when a respondent selects "yes" for question 16; question 18 only appears when a respondent selects "no" for question 16. Questions 1-3 and 19-20 are for all respondents and question 21 is optional. "Multiple choice" refers to those questions that allow the respondents to choose multiple answers. "Single choice" refers to those questions that only allow the respondents to choose one answer.

## Survey on *Danmu* Usage

Hello! Thank you very much for participating in this survey on *danmu* usage. This survey will be conducted anonymously, and all information collected will only be used for research purposes.

Below is a diagram of an example of *danmu* (the texts on the screen are *danmu* comments):



1. Your Age [Single Choice] \*

- Below 12
- 12-17
- 18-24
- 24-29
- 30-39
- 40-49
- 50-59
- Above 59

2. Your Sex [Single Choice] \*

- Male
- Female

3. Generally speaking, do you turn on *danmu* when watching online videos? [Single Choice] \*

- Yes
- No

4. On which website do you turn on *danmu* most often when watching online videos [Single Choice] \*

- Bilibili
- AcFun
- Niconico
- Iqiyi
- Tencent video

- Mango TV
- Others \_\_\_\_\_

5. Briefly describe why do you turn on *danmu* for that website. [Short Answer] \*

---

6. Why do you like watching videos with *danmu*? [Multiple Choice] \*

Select these options if applicable, feel free to add your own comments

- Because I want to see humorous comments (e.g. homophonic translation, odd points of the video that I don't notice)
- Because I feel I am watching the video with a friend
- Because I feel a sense of belonging, people who send *danmu* think in the same way as I am
- Because *danmu* provides me with outside knowledge that is needed to understand the content of the video
- Because I want to see other people's opinion towards the video content
- Other reasons/Comments \_\_\_\_\_

7. Which of the following forms of *danmu* do you like best/ use most? [Multiple Choice] \*

You can choose more than one options

- Abbreviations (e.g. awsl, xswl, 注孤生)
- Multilingual comments (e.g. include Japanese, Korean and etc. characters)
- Homophonic translations
- Emoticon
- Numbers as connotations (e.g. 666, 233)
- Spam (those use in catfight included)
- Evaluative Expressions (e.g. 囍, 真实, 合影)
- Conversation-like (e.g. 前方高能, 欢迎回来)
- Others \_\_\_\_\_

8. Do you often send *danmu* while watching videos? [Single Choice] \*

- Yes
- No

9. What kind of *danmu* do you like to send? [Multiple Choice] \*

- Abbreviations (e.g. awsl, xswl, 注孤生)
- Multilingual comments (e.g. include Japanese, Korean and etc. characters)
- Homophonic translations
- Emoticon



# How Images Influence People's Understanding of Science in the Context of COVID-19

---

*Xinyi Tian*

**Author Background:** *Xinyi Tian grew up in China and currently attends Beijing Academy, in Beijing, China. Her Pioneer research concentration was in the field of neuroscience/psychology and titled "Vision Impossible: An Introduction to the Psychology and Neuroscience of Seeing."*

## 1. Introduction

In countless conversations we have all tried to convince other people that what we know is true. When asked "how do you know," our answer will likely be "because I've read it somewhere." But how do we know that we know? Ontologists and epistemologists have suggested several ways: through informal and selective observation, overgeneralization, authority, and research methods (Blackstone, 2012). Through these answers given by philosophers, we can see that not all of our knowledge is reliable. In other words, we might "think we have a pretty good handle on how the world works, but when it comes down to it, it turns out we know very little about a lot of things." (Sloman & Fernbach, 2017) This phenomenon is especially evident during the initial stage of the pandemic outbreak, when everyone seemed to have their own knowledge of what was going on and what we ought to do. The confusion prompted a question: When we read a scientific article and convince ourselves that we "understand it," does our self-perceived understanding match our actual knowledge?

A 2008 study by Swartling and Helgesson shows an endeavor to address this. The researchers demonstrated that in a largescale predictive screening study, there was a distinct difference between self-assessed understanding and actual knowledge (Swartling & Helgesson, 2008). Later, a group of scholars at Yale discovered that participants' self-assessed ability to answer questions increased after searching for explanations online in a previous, unrelated task (Fisher, Goddu, & Keil, 2015). This example more directly illustrates how susceptible our self-assessing ability is when examined under the influence of external manipulations such as having access to the Internet.

One of these external manipulations is the presence of images. Advocates for including more visuals in textbooks as a means to improve students' comprehension have cited that we read visuals 6,000 times faster than we do textual information (Gazzaniga, 1994), and that "a picture is worth a thousand words." However, how much credit for our "understanding" is attributed to the textual information itself and how much more is to do with the appealing images alongside the text?

When it comes to images and scientific communication, it was demonstrated in a 2002 study by Smith et al. that visual displays are associated with a greater degree of scientific credibility (Smith, Best, Stubbs, Archibald, & Roberson-Nay, 2002). Following this line of logic, much attention was devoted to people's preferences on scientific explanations (see Lombrozo, 2016). Related research has focused on factors that influence participants' assessment of explanations (see McCabe & Castel, 2007; Weisberg, Keil, Goodstein, Rawson, & Gray, 2008; Hopkins, Weisberg, & Taylor, 2016) and has dedicated particular time to the effect of neuroscientific information on their judgment (see Fernandez-Duque, Evans, Christian, & Hodges, 2015; Weisberg, Taylor, & Hopkins, 2015).

The search interest of the keyword "Coronavirus" spiked to 100 on Google Trends as relative to other search interests on March 12, 2020, indicating that COVID-related topics were the most frequently-searched by users. Then, the keyword remained significantly more popular than other topics throughout the month (Google, 2020). Late in June, the updated Morbidity and Mortality Weekly Report (MMWR) posted online by the Centers for Disease Control and Prevention showed that "approximately one third of survey respondents engaged in non-recommended high-risk practices with the intent of preventing SARS-CoV-2 transmission, including using bleach on food products, applying household cleaning and disinfectant products to skin, and inhaling or ingesting cleaners and disinfectants." (Gharpure et al., 2020) Both the continuing trend of coronavirus search popularity and the shocking survey result demonstrate an urgent demand for effective scientific communication, rendering this study timely.

The purpose of this project is to answer the following questions: (1) Is there a difference between self-assessed understanding of COVID-19 and actual understanding? (2) How are assessed and actual understanding of COVID-19 influenced by the inclusion of images with the presentation of textual information? (3) What kind(s) of images, in particular, support understanding and what kind(s) of images might reduce understanding (while potentially increasing self-assessment)? The answers to these questions are of immediate significance to everyone in circumstances such as COVID-19, for the results may lend insight on generating more effective means to convey scientific knowledge to the public. As suggested in a 2002 study by Trout, "only the truth or accuracy of an explanation make the sense of understanding a valid cue to genuine understanding" (Trout, 2002). This study further wishes to draw conclusions that may help bring the two types of understanding closer together, in the hope of empowering more people to protect themselves and their loved ones from the pandemic.

In order to achieve these goals, self-assessed and actual understanding of textual information are measured and compared with the inclusion of either informative images, uninformative images, or no image at all.

## 2. Methodology

Due to time constraints and quarantine policy, surveys were distributed online to friends of mine and colleagues of my parents. Participants from China and the US between the ages of 15 and 50 were randomly assigned to three groups: the control group, the informative image group, and the uninformative image group. Each group was given surveys that are identical in textual information, but different in their inclusion of images. Each subject was asked to read two coronavirus-related, 200-word-long excerpts once, each retrieved from the official site of the Centers for



Disease Control and Prevention. A text-only survey was presented to the control group, an infographic was presented alongside the text for each excerpt to the informative image group, and an uninformative image was presented alongside the text for each excerpt to the uninformative image group.

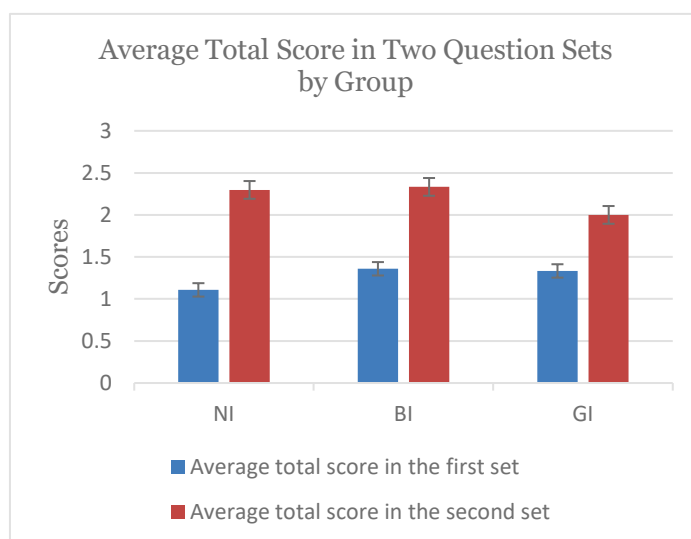
To examine subjects' self-assessed understanding, subjects were asked to self-report their understanding of the excerpt on a gradient scale with integer options, where 1 represents, "I did not understand at all," and 4 represents, "I completely understood." To assess subjects' actual understanding of the excerpt, four multiple choice questions examine subjects' opinions on statements related to the excerpt and their decisions under related scenarios. One point is earned for each question that the subject answers correctly, with no points awarded for incorrect answers. The self-assessed score is compared with the total score for each excerpt and with each other.

See the surveys in appendix.

### 3. Results

Of the 145 surveys distributed, 132 were collected and 115 were valid. The results presented are analyzed from 37 surveys of the control group (labeled "NI," which stands for "no image") and 39 surveys of each experimental group (labeled "BI" and "GI", which stand for "bad image (uninformative)" and "good image (informative)"). Average scores that indicate participants' actual knowledge and confidence level are plotted, and error bars mark standard errors.

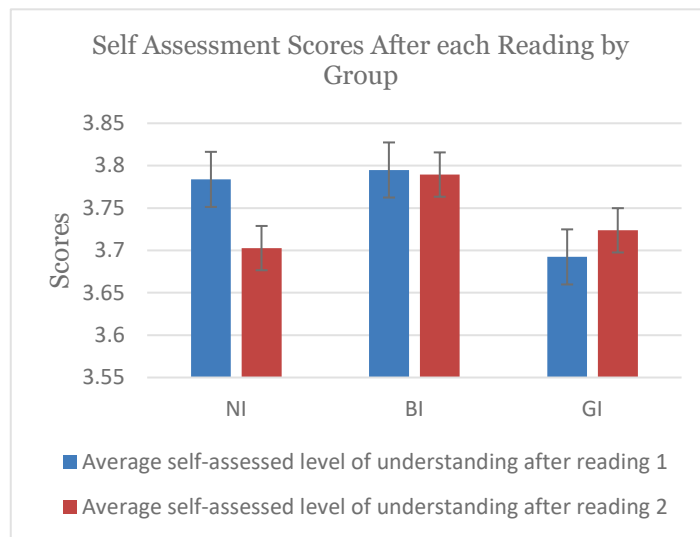
**Figure 1** plots the average scores that participants received on the two fact-based question sets. The figure shows that, overall, participants were unable to answer all the questions correctly. It also appears that the first question set was more difficult than the second, as reflected by fewer correct responses for this set. It is critical that participants made errors because this particular fact warrants that they did not perfectly understand the reading material.



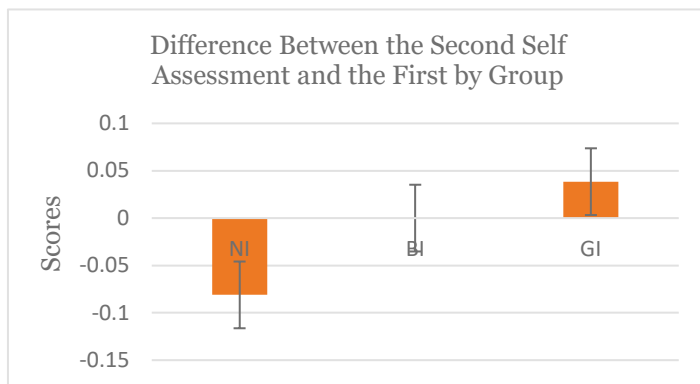
**Figure 1**

**Figure 2** plots average self-assessment scores by group. It is evident that, in general, subjects' self-assessments were relatively high: above 3.5 on average. In the NI group, participants became significantly less confident about their understanding in the second question set, compared to the first. This suggests that experience with questions after reading the first excerpt produced a better appreciation of what the participants actually understood. Interestingly, this was not the case for the BI group: self-assessment remained high and did not fluctuate over time, despite having the same experience with question sets post-reading. This implies that uninformative images inflate self-assessed understanding, without – referring back to the result concluded from Figure 1 – improving actual understanding. The GI group, on the contrary, showed generally lower self-assessments during both question periods.

**Figure 3** makes this pattern more apparent by plotting the average difference between the second and first assessment by group.

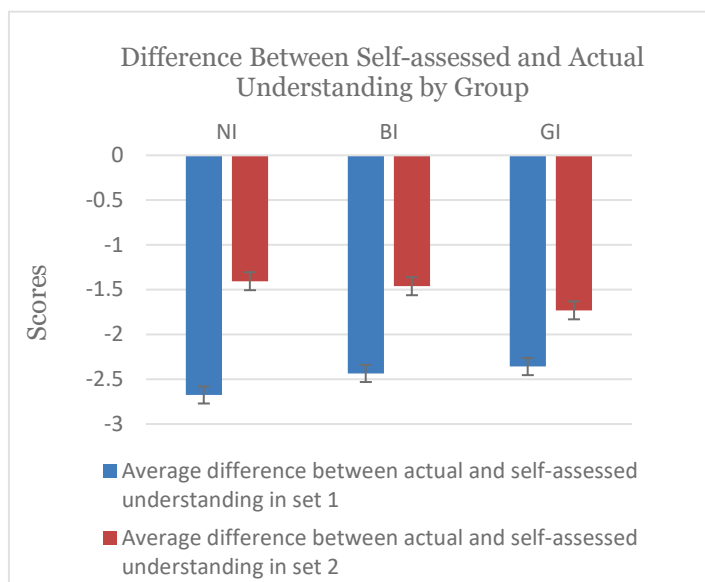


**Figure 2**



**Figure 3**

Participants, regardless of their groups, were quite confident about their understanding. However, is their confidence warranted? **Figure 4** plots the average differences between assessed and actual understanding, calculated by the number of answers correct (one point each) minus self-assessment score. This figure demonstrates that there is a significant difference between subjects' self-assessed understanding of COVID-19 information and their actual understanding. The pattern is more evident in the first reading period than it is in the second: in general, participants have over-rated their understanding by approximately 2.5 points more than their actual performance on the first question set, and 1.5 points more for the second.



**Figure 4**

## 4. Discussion

This study sought to determine how images affect actual understanding and self-assessment of understanding in the specific context of COVID-19. Participants were asked to answer two sets of fact-based multiple-choice questions related to the reading materials, and also to assess their own understanding twice. They were all provided with the same excerpts about the pandemic, except that one group was not shown any images, one was shown an uninformative image, and one was shown an infographic. Overall, self-assessed understanding was inflated compared to actual. Intriguingly, this inflation waned after the experience of reading-comprehension questions for the control group and was generally lower for the infographic group, but persisted for the uninformative image group. This indicates that an uninformative image negatively affected participants' self-assessment ability, for it prevented them from modulating their confidence level after answering a first set of questions. Participants in the other groups, meanwhile, seemed to appreciate that they were overconfident after completing the first set of questions.

The findings of this paper have two primary implications. First, it is essential for the general public to realize that notable differences exist between one's self-

assessed and actual understanding when reading about scientific information such as COVID-19. As such, people should not rely entirely on what they presume to know for the sake of their own health. Instead, it is advised that they frequently refer back to the original information, to ensure more accurate interpretation of the information.

Second, given the findings that self-assessment is inflated by uninformative images, it is suggested that the media does not include uninformative images alongside scientific information since they promote overconfidence. When it comes to scientific knowledge, people will never lose by being more meticulous with information; in this sense, they are always encouraged to double-check the original source, to ask questions, or to consult with professionals. In theory, a less confident person is probably more likely to do so than an overconfident one. This inference also suggests a potential future direction, which is to explore what kind of behavior(s) overconfidence promotes. A possible follow-up study might investigate whether subjects will ask questions after reading scientific material: will the uninformative image group ask fewer questions?

## References

- Blackstone, A. (2012). *Principles of Sociological Inquiry: Qualitative and Quantitative Methods*. Saylor Academy.
- Chaitin, G. (2006). The Limits of Reason. *Scientific American*.
- Fernandez-Duque, D., Evans, J., Christian, C., & Hodges, S. D. (2015). Superfluous Neuroscience Information Makes Explanations of Psychological Phenomena More Appealing. *Journal of Cognitive Neuroscience*, 926-944.
- Fisher, M., Goddu, M. K., & Keil, F. C. (2015). Searching for Explanations: How the Internet Inflates Estimates of Internal Knowledge. *Journal of Experimental Psychology: General*.
- Gazzaniga, M. (1994). *Nature's Mind: The Biological Roots of Thinking, Emotions, Sexuality, Language, and Intelligence*. Basic Books.
- Google. (2020, June 14). *Coronavirus Search Trends*. Retrieved from Google Trends: [https://trends.google.com/trends/story/US\\_cu\\_4Rjdh3ABAABMHM\\_en](https://trends.google.com/trends/story/US_cu_4Rjdh3ABAABMHM_en)  
Access date: June 14, 2020.
- Hopkins, E. J., Weisberg, D. S., & Taylor, J. C. (2016). The Seductive Allure is a Reductive Allure: People Prefer Scientific Explanations that Contain Logically Irrelevant Reductive Information. *Journal of Cognition*, 67-76.
- Lombrozo, T. (2016). Explanatory Preferences Shape Learning and Inference. *Trends in Cognitive Science*.
- McCabe, D. P., & Castel, A. D. (2007). Seeing is Believing: The Effect of Brain Images. *Cognition*, 343-52.
- Gharpure, R., Hunter, C. M., Schnall, A. H., Barrett, C. E., Kirby, A. E., Kunz, J., Berling, K., Mercante, J. W., Murphy, J. L., & Garcia-Williams, A. G. (2020). Knowledge and Practices Regarding Safe Household Cleaning and Disinfection for COVID-19 Prevention — United States, May 2020. *MMWR Morb Mortal Wkly Rep* 2020, 705-9.
- Sloman, S., & Fernbach, P. (2017). *The Knowledge Illusion: Why We Never Think Alone*. Penguin Publishing Group.

- Smith, L. D., Best, L. A., Stubbs, D. A., Archibald, A. B., & Roberson-Nay, R. (2002). Constructing Knowledge: The Role of Graphs and Tables in Hard and Soft Psychology. *American Psychologist*, 749-761.
- Swartling, U., & Helgesson, G. (2008). Self-Assessed Understanding as a Tool for Evaluating Consent: Reflections on a Longitudinal Study. *Journal of Medical Ethics*, 557-562.
- Trout, J. D. (2002). Scientific Explanation and the Sense of Understanding. *Philosophy of Science*, 212-233.
- Weisberg, D. S., Keil, F. C., Goodstein, J., Rawson, E., & Gray, J. R. (2008). The Seductive Allure of Neuroscience Explanations. *Journal of Cognitive Neuroscience*, 470-77.
- Weisberg, D. S., Taylor, J. C., & Hopkins, E. J. (2015). Deconstructing the Seductive Allure of Neuroscience. *Judgment and Decision Making*, 429-441.

## Appendix A: Control Group Survey

This survey is designed to investigate how well you understand information about COVID-19 on the Internet. Please read the excerpts at your normal reading speed and answer the questions that follow. You may highlight your answers in the file and send it back to me when you complete the questions. **You should not refer to the excerpts when answering the questions.**

### Your gender

Male

Female

### Your age

10-19

20-29

30-39

40-49

50+

### I. How the Virus Spreads

(retrieved from <https://www.cdc.gov/coronavirus/2019-ncov/preventgetting-sick/how-covid-spreads.html> and <https://www.cdc.gov/coronavirus/2019-ncov/preventgetting-sick/prevention.html>)

- The virus is thought to spread mainly from person-to-person.
  - Between people who are in close contact with one another (within about 6 feet).
  - Through respiratory droplets produced when an infected person coughs, sneezes or talks.

- These droplets can land in the mouths or noses of people who are nearby or possibly be inhaled into the lungs.
- COVID-19 may be spread by people who are not showing symptoms.

It may be possible that a person can get COVID-19 by touching a surface or object that has the virus on it and then touching their own mouth, nose, or possibly their eyes. This is not thought to be the main way the virus spreads, but we are still learning more about how this virus spreads.

There is currently no vaccine to prevent coronavirus disease 2019 (COVID-19), and the best way to prevent illness is to avoid being exposed to this virus. You can take steps to slow the spread:

- Maintain good social distance (about 6 feet). This is very important in preventing the spread of COVID-19.
- Wash your hands often with soap and water. If soap and water are not available, use a hand sanitizer that contains at least 60% alcohol.
- Routinely clean and disinfect frequently touched surfaces.
- Cover your mouth and nose with a cloth face covering when around others.

On the scale below, indicate to what extent did you understand the information presented in the excerpt.

I did not understand at all

I understood everything

1

2

3

4

1. Do you think COVID-19 could spread through surfaces of packages?
  - a. Definitely yes
  - b. Likely, but we don't know
  - c. Not likely, but we don't know
  - d. Definitely not
2. Do you think that drinking alcohol could prevent infection?
  - a. Definitely yes
  - b. Likely, but we don't know
  - c. Not likely, but we don't know
  - d. Definitely not
3. Of the following, which would be the best way to prevent infection of COVID-19?
  - a. Getting a vaccination
  - b. Drinking fluids and staying hydrated
  - c. Staying isolated
  - d. Staying warm
4. Of the following situations, when do you need precautions (e.g. wearing a mask)? Check all that apply.

- a. Being with friends who are not showing symptoms
- b. Being close to an elderly person who has tested negative for COVID-19
- c. Being around family members who have been your quarantine buddies
- d. Going to the bathroom in your own house

## II. What to do if you have pets?

(retrieved from <https://www.cdc.gov/coronavirus/2019ncov/daily-life-coping/pets.html> and <https://www.cdc.gov/coronavirus/2019-ncov/daily-lifecoping/positive-pet.html>)

- A small number of pets worldwide, including cats and dogs, have been reported to be infected with the virus that causes COVID-19, mostly after close contact with people with COVID-19. It appears that the virus that causes COVID-19 can spread from people to animals in some situations.
- Based on the limited information available to date, the risk of animals spreading COVID-19 to people is considered to be low.
- Treat pets as you would other human family members – do not let pets interact with people outside the household.
- If a person inside the household becomes sick, isolate that person from everyone else, including pets.
- If your pet tests positive for the virus that causes COVID-19, isolate the pet from everyone else, including other pets.
- Do not wipe or bathe your pet with chemical disinfectants, alcohol, hydrogen peroxide, or any other products not approved for animal use.
- Only a few pets have been confirmed to be infected with the virus that causes COVID-19. Some pets did not show any signs of illness, but those pets that did get sick all had mild disease that could be taken care of at home. None have died from the infection.

On the scale below, indicate to what extent did you understand the information presented in the excerpt.

I did not understand at all

I understood everything

1

2

3

4

1. Do you think pets are less susceptible to COVID-19 than human?
  - a. Definitely yes
  - b. Likely, but we don't know
  - c. Not likely, but we don't know

- d. Definitely not
2. During the COVID-19 outbreak, if you have to ask one of the following people to take care of your pet for one day while you are at work, who would it be?
- Mr. A, whose wife had been diagnosed positive for the virus
  - Mrs. H, who kindly promises to bathe your pet with disinfectants
  - Ms. C, who is willing to wear a mask while being with your pet
  - None of the above, because none of the situations is safe.
3. Of the following situations, when do you need to isolate your pet from people? Check all that apply.
- When a family-friend comes over to visit
  - When one of your family members is showing symptoms of COVID-19
  - When you are alone with your pet at home
  - At all times
4. Do you think you should request to be tested for the virus under the following situation?
- Your best friend's dog tested positive for COVID-19 yesterday, while you just visited and played with the dog two days ago.
- Definitely yes
  - Probably
  - Maybe, but the risk is low
  - Definitely not, people can't get the virus from pets

## **Appendix B: Uninformative Group Survey**

This survey is designed to investigate how well you understand information about COVID-19 on the Internet. Please read the excerpts at your normal reading speed and answer the questions that follow. You may highlight your answers in the file and send it back to me when you complete the questions. **You should not refer to the excerpts when answering the questions.**

Your gender

Male

Female

Your age

10-19

20-29



30-39

40-49

50+

**I. How the Virus Spreads**

(retrieved from <https://www.cdc.gov/coronavirus/2019-ncov/preventgetting-sick/how-covid-spreads.html> and <https://www.cdc.gov/coronavirus/2019-ncov/preventgetting-sick/prevention.html>)

- The virus is thought to spread mainly from person-to-person.
  - Between people who are in close contact with one another (within about 6 feet).
  - Through respiratory droplets produced when an infected person coughs, sneezes or talks.
  - These droplets can land in the mouths or noses of people who are nearby or possibly be inhaled into the lungs.
  - COVID-19 may be spread by people who are not showing symptoms.

It may be possible that a person can get COVID-19 by touching a surface or object that has the virus on it and then touching their own mouth, nose, or possibly their eyes. This is not thought to be the main way the virus spreads, but we are still learning more about how this virus spreads.

There is currently no vaccine to prevent coronavirus disease 2019 (COVID-19), and the best way to prevent illness is to avoid being exposed to this virus. You can take steps to slow the spread.

- Maintain good social distance (about 6 feet). This is very important in preventing the spread of COVID-19.
- Wash your hands often with soap and water. If soap and water are not available, use a hand sanitizer that contains at least 60% alcohol.
- Routinely clean and disinfect frequently touched surfaces.
- Cover your mouth and nose with a cloth face covering when around others.

On the scale below, indicate to what extent did you understand the information presented in the excerpt.

I did not understand at all

I understood everything

1

2

3

4

1. Do you think COVID-19 could spread through surfaces of packages?
  - a. Definitely yes

- b. Likely, but we don't know
  - c. Not likely, but we don't know
  - d. Definitely not
2. Do you think that drinking alcohol could prevent infection?
  - a. Definitely yes
  - b. Likely, but we don't know
  - c. Not likely, but we don't know
  - d. Definitely not
3. Of the following, which would be the best way to prevent infection of COVID-19?
  - a. Getting a vaccination
  - b. Drinking fluids and staying hydrated
  - c. Staying isolated
  - d. Staying warm
4. Of the following situations, when do you need precautions (e.g. wearing a mask)? Check all that apply.
  - a. Being with friends who are not showing symptoms
  - b. Being close to an elderly person who has tested negative for COVID-19
  - c. Being around family members who have been your quarantine buddies
  - d. Going to the bathroom in your own house

## II. What to do if you have pets?

(retrieved from <https://www.cdc.gov/coronavirus/2019ncov/daily-life-coping/pets.html> and <https://www.cdc.gov/coronavirus/2019-ncov/daily-lifecoping/positive-pet.html>)



- A small number of pets worldwide, including cats and dogs, have been reported to be infected with the virus that causes COVID-19, mostly after close contact with people with COVID-19. It appears that the virus that causes COVID-19 can spread from people to animals in some situations.
- Based on the limited information available to date, the risk of animals spreading COVID-19 to people is considered to be low.

- Treat pets as you would other human family members – do not let pets interact with people outside the household.
- If a person inside the household becomes sick, isolate that person from everyone else, including pets.
- If your pet tests positive for the virus that causes COVID-19, isolate the pet from everyone else, including other pets.
- Do not wipe or bathe your pet with chemical disinfectants, alcohol, hydrogen peroxide, or any other products not approved for animal use.
- Only a few pets have been confirmed to be infected with the virus that causes COVID-19. Some pets did not show any signs of illness, but those pets that did get sick all had mild disease that could be taken care of at home. None have died from the infection.

On the scale below, indicate to what extent did you understand the information presented in the excerpt.

I did not understand at all

I understood everything

1

2

3

4

1. Do you think pets are less susceptible to COVID-19 than humans?
  - a. Definitely yes
  - b. Likely, but we don't know
  - c. Not likely, but we don't know
  - d. Definitely not
2. During the COVID-19 outbreak, if you have to ask one of the following people to take care of your pet for one day while you are at work, who would it be?
  - a. Mr. A, whose wife had been diagnosed positive for the virus
  - b. Mrs. H, who kindly promises to bathe your pet with disinfectants
  - c. Ms. C, who is willing to wear a mask while being with your pet
  - d. None of the above, because none of the situations is safe.
3. Of the following situations, when do you need to isolate your pet from people? Check all that apply.
  - a. When a family-friend comes over to visit
  - b. When one of your family members is showing symptoms of COVID-19
  - c. When you are alone with your pet at home
  - d. At all times

4. Do you think you should request to be tested for the virus under the following situation?

Your best friend's dog tested positive for COVID-19 yesterday, while you just visited and played with the dog two days ago.

- a. Definitely yes
- b. Probably
- c. Maybe, but the risk is low
- d. Definitely not, people can't get the virus from pets

### **Appendix C: Informative Group Survey**

This survey is designed to investigate how well you understand information about COVID-19 on the Internet. Please read the excerpts at your normal reading speed and answer the questions that follow. You may highlight your answers in the file and send it back to me when you complete the questions. **You should not refer to the excerpts when answering the questions.**

Your gender

Male

Female

Your age

10-19

20-29

30-39

40-49

50+

#### **I. How the Virus Spreads**

(retrieved from <https://www.cdc.gov/coronavirus/2019-ncov/preventgetting-sick/how-covid-spreads.html> and <https://www.cdc.gov/coronavirus/2019-ncov/preventgetting-sick/prevention.html>)

- The virus is thought to spread mainly from person-to-person.
  - Between people who are in close contact with one another (within about 6 feet).
    - Through respiratory droplets produced when an infected person coughs, sneezes or talks.
    - These droplets can land in the mouths or noses of people who are nearby or possibly be inhaled into the lungs.

- COVID-19 may be spread by people who are not showing symptoms.

It may be possible that a person can get COVID-19 by touching a surface or object that has the virus on it and then touching their own mouth, nose, or possibly their eyes. This is not thought to be the main way the virus spreads, but we are still learning more about how this virus spreads.

There is currently no vaccine to prevent coronavirus disease 2019 (COVID-19), and the best way to prevent illness is to avoid being exposed to this virus. You can take steps to slow the spread.

- Maintain good social distance (about 6 feet). This is very important in preventing the spread of COVID-19.
- Wash your hands often with soap and water. If soap and water are not available, use a hand sanitizer that contains at least 60% alcohol.
- Routinely clean and disinfect frequently touched surfaces.
- Cover your mouth and nose with a cloth face covering when around others.

On the scale below, indicate to what extent did you understand the information presented in the excerpt.

I did not understand at all

I understood everything

1

2

3

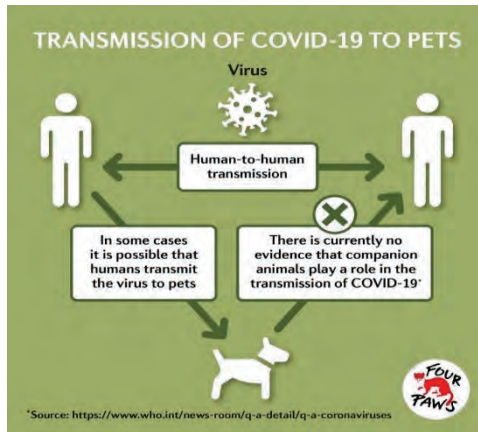
4

1. Do you think COVID-19 could spread through surfaces of packages?
  - a. Definitely yes
  - b. Likely, but we don't know
  - c. Not likely, but we don't know
  - d. Definitely not
2. Do you think that drinking alcohol could prevent infection?
  - a. Definitely yes
  - b. Likely, but we don't know
  - c. Not likely, but we don't know
  - d. Definitely not
3. Of the following, which would be the best way to prevent infection of COVID-19?
  - a. Getting a vaccination
  - b. Drinking fluids and staying hydrated
  - c. Staying isolated
  - d. Staying warm
4. Of the following situations, when do you need precautions (e.g. wearing a mask)? Check all that apply.
  - a. Being with friends who are not showing symptoms

- b. Being close to an elderly person who has tested negative for COVID-19
- c. Being around family members who have been your quarantine buddies
- d. Going to the bathroom in your own house

## II. What to do if you have pets?

(retrieved from <https://www.cdc.gov/coronavirus/2019ncov/daily-life-coping/pets.html> and <https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/positive-pet.html>)



- A small number of pets worldwide, including cats and dogs, have been reported to be infected with the virus that causes COVID-19, mostly after close contact with people with COVID-19. It appears that the virus that causes COVID-19 can spread from people to animals in some situations.
- Based on the limited information available to date, the risk of animals spreading COVID-19 to people is considered to be low.
- Treat pets as you would other human family members – do not let pets interact with people outside the household.
- If a person inside the household becomes sick, isolate that person from everyone else, including pets.
- If your pet tests positive for the virus that causes COVID-19, isolate the pet from everyone else, including other pets.
- Do not wipe or bathe your pet with chemical disinfectants, alcohol, hydrogen peroxide, or any other products not approved for animal use.
- Only a few pets have been confirmed to be infected with the virus that causes COVID-19. Some pets did not show any signs of illness, but those pets that did get sick all had mild disease that could be taken care of at home. None have died from the infection.

On the scale below, indicate to what extent did you understand the information presented in the excerpt.

I did not understand at all

I understood everything

1

2

3

4

1. Do you think pets are less susceptible to COVID-19 than human?
  - a. Definitely yes
  - b. Likely, but we don't know
  - c. Not likely, but we don't know
  - d. Definitely not
2. During the COVID-19 outbreak, if you have to ask one of the following people to take care of your pet for one day while you are at work, who would it be?
  - a. Mr. A, whose wife had been diagnosed positive for the virus
  - b. Mrs. H, who kindly promises to bathe your pet with disinfectants
  - c. Ms. C, who is willing to wear a mask while being with your pet
  - d. None of the above, because none of the situations is safe.
3. Of the following situations, when do you need to isolate your pet from people? Check all that apply.
  - a. When a family-friend comes over to visit
  - b. When one of your family members is showing symptoms of COVID-19
  - c. When you are alone with your pet at home
  - d. At all times
4. Do you think you should request to be tested for the virus under the following situation?

Your best friend's dog tested positive for COVID-19 yesterday, while you just visited and played with the dog two days ago.

- a. Definitely yes
- b. Probably
- c. Maybe, but the risk is low
- d. Definitely not, people can't get the virus from pets



# The Influence of Government Intervention on Sports Participation for School-Aged Children in China Between 2011 and 2020

---

*Xinyue (Abby) Luo*

**Author Background:** LuoXinyue grew up in China and currently attends the International Department of The Affiliated High School of South China Normal University in Guangzhou, China. Her Pioneer research concentration was in the fields of history and sociology and titled “Only a Game? Players and Spectators: History, Culture, and Sports.”

## Introduction

In the past decade, the physical and mental well-being of school-aged children has become an important concern for parents worldwide. Research has shown that daily physical activity helps boost mental and physical well-being.<sup>1</sup> In response to this growing concern for children’s health, the Chinese government has made an effort to increase children’s participation in sport and physical activity through their new policies. Aiming to understand the influence of government intervention on sports participation rate for school-aged children in China, this paper investigates the effectiveness of a series of policies passed by the Chinese government between 2011 and 2020.

This paper will first look into the history of sports developments from the introduction of western sport to China to the National Fitness Plan of 2011-2015 (a plan issued by the State Council that aims to boost mass sports development in China). The second part will then analyze the National Fitness Plan of 2011-2015, the National Fitness Plan of 2016-2020, and the Healthy China 2030 Plan (2019), which are three fitness policies that aim to improve school-aged children’s sport participation by providing a higher quality and quantity of sports facilities, training more qualified PE instructors, and fostering a culture that embraces sport participation.<sup>2</sup> Through an in-depth analysis, this paper will show how the policies have been unrealistic and ambitious, establishing standards that are too vague and hard to reach while providing little direction on how to achieve those standards.

---

<sup>1</sup> Miriam Reiner, Christina Niermann, Darko Jekauc, and Alexander Woll, 2013, "Long-Term Health Benefits of Physical Activity – a Systematic Review of Longitudinal Studies," *BMC Public Health* 13 (1): 813-814, <http://www.biomedcentral.com/1471-2458/13/813>.

Karen Pickett, Tony Kendrick, and Lucy Yardley, 2017, "“A Forward Movement into Life”: A Qualitative Study of how, Why and when Physical Activity may Benefit Depression," *Mental Health and Physical Activity* 12: 100-103, doi: 10.1016/j.mhpa.2017.03.004.

<sup>2</sup> Central Committee of the Communist Party of China, State Council of China, “National Fitness Plan 2011-2015,” [in Chinese], 2011, The Xinhua News Agency. [http://www.gov.cn/zhengce/content/2011-02/24/content\\_6411.htm](http://www.gov.cn/zhengce/content/2011-02/24/content_6411.htm).



Next, to assess the effectiveness of government interventions, this paper will look into the participation rate in sports and physical activity during and after the three policies. Through comparison, I find that the number of school-aged children who meet the standard of the minimum level of physical activity set by the policies is often lower than projected in the policies. Based on my understanding of Chinese economics, the increase in school-aged children's participation is also often a result of China's economic growth rather than the programs. Besides, the amount of academic pressure imposed by parents and schools on children renders sports time for school-aged children a luxury.

Finally, a self-reported survey on the current attitude towards sport participation for school-aged children will investigate factors that hinder sport participation among Chinese school-aged children. Analysis indicates that by shifting government investment in elite sports training to broader community sport, incorporating sports into the education system, encouraging the growth of private and local sports enterprises, and fostering a national sport culture, the future of physical education and sports development is bright.

However, changes are easier proposed than implemented as different provinces in China face different needs. What remains crucial is the government's long-term financial and policy support. Without incorporating physical education in the school curriculum, physical and sport activity will always remain a luxury for school-aged children.

## **The Origin of Western Sports and Government-Sponsored Sports Programs in China**

Christian Missionaries and organizations that built modern sports facilities introduced Western sport and physical education to China in the late nineteenth century. Between 1922 and 1927, The Nationalist Government, under the leadership of the Chinese Nationalist Party (KMT), promoted participation in sport to boost national unity, patriotism, and physical strength. Many large sports complexes were built by the KMT using government funding, including the Nanjing Central Sports Stadium. These constructions mattered for they led to the popularization of sports among the masses.<sup>3</sup>

In the 1920s, the rate of sport participation increased in China as people associated sports with patriotism. After World War I, Japan issued the Twenty-One Demands to the KMT and gained temporary control over southern Manchuria.<sup>4</sup> These humiliating demands led to a surge of patriotism, team spirit, self-control, vigor, and sacrifice that resulted in an increased participation rate in sports.

In the 1930s, sports were used by the KMT as a tool to gain international recognition. Under the "militarized Tiyu" movement, which emphasized the benefits

---

<sup>3</sup> Nationalist Government refers to the government of the Republic of China between 1 July 1925 and 20 May 1948. Lu Zhouxiang, "From Hongkew Recreation Ground to Bird's Nest: The Past, Present and Future of Large Sports Venues in China," *The International Journal of the History of Sport*, vol. 30, no. 4 (2013): 422-428, doi:10.1080/09523367.2013.765724.

<sup>4</sup> The Twenty One Demands are claims made by the Japanese government to special privileges in China during World War I. Through these demands Japan claimed railway and mining properties in Shandong province; controlled the Han-Ye-Ping mining base in central China; gained access to harbors, bays, and islands along China's coast; and controlled through advisers, of Chinese financial, political, and police affairs.

The Twenty-One Demands, <https://u-s-history.com/pages/h1111.html>.

of sports for training a stronger army, the KMT saw physical education as a crucial tool to improve the Chinese military capacity.<sup>5</sup> After the Mukden incident in 1931, the KMT sent China's first athlete Liu Changchun to take part in the 1932 Los Angeles Olympics.<sup>6</sup> Through participating in the Olympics, the KMT attempted to prevent Japan's legitimization of its puppet state in Manchuria by establishing Chinese international influence. However, as a full-scale Second Sino-Japanese War broke out between China and Japan in late 1937, the idea of "militarized Tiyu" soon proved impractical. The KMT realized the inefficiency of training soldiers through sports. Sending athletes to the Olympics also became less important for the KMT compared to the war. Nevertheless, a mass sports movement called the New Life Movement, which aimed to strengthen all bodies in the nation, continued from 1934 to 1949.<sup>7</sup>

### **Sports Development Under the Chinese Communist Party**

During the Second Sino-Japanese War, the Chinese Communist Party (CCP) took a different approach to sports development than the KMT. The Red Army propagated the Red Tiyu movement through "clubs" (Julebu) and "Lenin Rooms" (Liening shi), a model adopted from the Soviet Union.<sup>8</sup> After the CCP gained victory over the KMT in the Chinese Civil War between 1927 and 1950, the CCP established a new government, the People's Republic of China (PRC). The newly established PRC's involvement in sports dates back to the 1950s when the PRC attempted to use sports as a venue to show the world its power. The government associated nationalism with sports and physical culture in an attempt to eradicate the perceived stereotype of the "sick man of East Asia" (dong ya bing fu), which claimed that all Chinese men have feeble physical strength.<sup>9</sup>

The pattern of using sports as a venue to gain world recognition reemerged in the 1960s. In 1963, the PRC attempted to promote physical culture at the Games of the New Emerging Forces (GANEFU), an Asian version of the Olympic Games, attended by forty-eight nations. While the GANEFU provided China a way to represent itself to its people, it also challenged the Olympic movement and negotiated

---

<sup>5</sup> Andrew David Morris, "Cultivating the National Body: A History of Physical Culture in Republican China," (PHD diss., University of California San Diego, 1998), ProQuest Dissertations Publishing. pp. 528-530

<sup>6</sup> The Mukden incident was staged by the Japanese government as an excuse to invade and occupy Manchuria.

Milestones: 1921-1936 - Office of the Historian. <https://history.state.gov/milestones/1921-1936/mukden-incident>. Accessed 1 Sept. 2020.

<sup>7</sup> Lu Zhouxiang, "Sport, Nationalism and the Building of the Modern Chinese Nation State (1912-49)," 2011, *The International Journal of the History of Sport*, vol. 28, no. 7, pp. 1046-1047. doi:10.1080/09523367.2011.563635.

<sup>8</sup> The "clubs" and "Lenin Rooms" were small organizations that brought people together for community sports; Andrew David Morris, "Cultivating the National Body: A History of Physical Culture in Republican China," (PHD diss., University of California San Diego, 1998), ProQuest Dissertations Publishing. pp. 546-548

<sup>9</sup> Y. K. Kogh, "A Descriptive Trend Analysis of Government Policy on Physical Education in the People's Republic of China: 1949-1965," (PHD diss., University of Denver, 1987), <http://search.ebscohost.com.ezproxy.bowdoin.edu/login.aspx?direct=true&db=hia&AN=46184881&site=ehost-live>. pp. 283

a new position for China on the world stage.<sup>10</sup> The Cultural Revolution between 1967 and 1971 temporarily stopped sports developments. However, after Mao Zedong's death (1976), sports development resumed. The PRC issued a series of policies emphasizing elite programs and winning gold medals in the Olympics, which led to massive sums of money being spent on sports developments during the Reformation and Opening-up Era.<sup>11</sup>

In 1995, however, the PRC shifted the focus to community sports and launched the Nation Fitness Program to address people's complaints that sports development between 1970 and 1990 only benefited elite athletes and neglected the population as a whole. The National Fitness Program, aiming to increase participation in sports among the general public, continues to this day. The State Council and the Sports Bureau program revise the program every four years.

### Past Studies on Sport Policies

Studies in Western countries have evaluated policies for sports development in China. Jepkorir Rose, Chepyator-Thomson, and Shan-Hui Hsu, after analyzing the physical education developments in various countries, recommended a public-policy reform focusing on reforms within the physical education curriculum, physical education in teacher education, and after-school programs for China. By reforming the curriculum and training better physical education teachers, they argued that more school-aged children would adopt a habit of exercising from a young age.<sup>12</sup> Similarly, in 1987, Y. K. Kogh analyzed government policy on physical education in Chinese schools by analyzing leaders' statements, aims and purposes, laws and regulations, administrative and organizational structures, and programs in China between 1949 and 1965. Kogh's work showed that government policy and physical education development did help to boost physical activities.<sup>13</sup>

In the past, there have been studies analyzing whether policy evaluation is effective or not. Barrie Houlihan and Mick Green acknowledged the difficulties of evaluating policies given the "complexity and dynamism of the socio-economic environment within which" the programs function.<sup>14</sup> In general, the relationship

---

<sup>10</sup> Amanda Shuman, "Elite Competitive Sport in the People's Republic of China 1958-1966: The Games of the New Emerging Forces (GANEF0)", *Journal of Sport History* 40, no.2 (2013): 261-263, <http://www.jstor.org/stable/10.5406/jsporthistory.40.2.258>.

<sup>11</sup> The Cultural Revolution was a period of political upheaval led by the CCP to bring about the return of Revolutionary Maoist beliefs. The movement included attacks on intellectuals and a personality cult around Mao Zedong. During the Reformation and Opening up period, Deng Xiaoping began a series of economic reforms termed "socialism with Chinese characteristics".

Lu Zhouxiang, "From Hongkew Recreation Ground to Bird's Nest: The Past, Present and Future of Large Sports Venues in China," *The International Journal of the History of Sport*, vol. 30, no. 4 (2013): 422-424. doi:10.1080/09523367.2013.765724.

<sup>12</sup> Chepyator-Thomson, Jepkorir Rose, and Shan-Hui Hsu, *Global Perspectives on Physical Education and After-School Sport Programs*, 2013, Lanham, MD: UPA, ProQuest Ebook Central. pg.83-97

<sup>13</sup> Y. K. Kogh, "A Descriptive Trend Analysis of Government Policy on Physical Education in the People's Republic of China: 1949-1965," (PHD diss., University of Denver, 1987), <http://search.ebscohost.com.ezproxy.bowdoin.edu/login.aspx?direct=true&db=hia&AN=46184881&site=ehost-live>. pg. 284-286

<sup>14</sup> Barrie Houlihan and Mick Green, *Routledge Handbook of Sports Development*, 2011, London: Taylor & Francis Group, ProQuest Ebook Central. pg.561-570

between sport policies and sports development is often weak, meaning sports development shares a stronger relationship with broader social developments, such as economic growth and urbanization. Growth in sport participation often is not the result of reforms in policies but rather the economic growth and improved living standards of the people. However, because the Chinese government is more centralized than most governments in western countries, sport development in China is different. The Chinese Planned Economy exerts great influence over economic growth, thereby giving the government the power to decide on which economic sectors they will focus.

### **A Shift to Mass Sports Development Between 1995 and 2011**

In 1995, the National Fitness Program began with the intention of encouraging mass sports development in China, but it achieved little initially. A study in 2003 on physical activity levels and patterns in Chinese school-aged students (6 to 18 years old) showed that 72% of students engaged in in-school moderate to vigorous physical activity; however, fewer than 8% eight percent engaged in such activity outside school.<sup>15</sup> Surprisingly, TV watching did not account for the lack of physical activity. Unlike most Western countries, Chinese students rarely helped with household chores or participated in after-school physical activities owing to the pressure to achieve academic excellence. Outside of school, the only type of physical activity was commuting to school by walking or biking.

Though the overall participation rate in physical activity did not increase substantially, physical education improved. A documentary produced by the UK Department of Education in 2006 showed that compared to Britain, Nanjing elementary school students devoted more time to physical exercise, enjoyed better-planned programs, better nutrition, well-organized everyday rituals, and specialized teachers.<sup>16</sup> However, Nanjing, the capital of Jiangsu province, was one of the most developed and modern cities in China, with a booming economy and high living standards. In most rural elementary schools, children enjoyed neither proper nutrition nor specialized physical education instructions.

Still, criticisms of the government's partiality for elite sports programs remained. China was known for its rigorous but efficient elite sport system. Athletes were recruited at a young age and underwent intensive training once they were in the elite sports system. Children recruited into athletic schools participated in an academic curriculum tailored to them, which included special textbooks with sport-related examples and less demanding schoolwork. In these schools, students and teachers valued athletic achievements more than academic grades.<sup>17</sup> After joining the government training program, Yao Ming, a former professional basketball player, noted that he trained ten hours a day when he was 14 years old. The screening of

---

<sup>15</sup> C.Tudor-Locke, BE Ainsworth, LS Adair, S. Du, and BM Popkin, "Physical Activity and Inactivity in Chinese School-Aged Youth: The China Health and Nutrition Survey," *International Journal of Obesity* 27, no.9 (date): 1096-1098. doi: 10.1038/sj.ijo.0802377.

<sup>16</sup> "How do they do it in China?" directed by Anonymous, produced by Real Life Productions, Teachers TV/UK Department of Education, 2006, Alexander Street, <https://video.alexanderstreet.com/watch/how-do-they-do-it-in-china>.

<sup>17</sup> "China: Going for Gold," directed by Anonymous, produced by Real Life Productions, Teachers TV/UK Department of Education, 2007, Alexander Street, <https://video.alexanderstreet.com/watch/china-going-for-gold>.

players was conducted at a young age with strict standards, and only the best players are recruited. Heights of the players and those of the parents were considered by coaches in recruitment. Children who did not meet the standard had no chance to join the program.<sup>18</sup>

The structure of China's elite athletic training system has a lot to do with why children and parents are unwilling to spend time on sports. Talented children who are potential future champions do not need to take the GaoKao and ZhongKao examinations (two major examinations in China, the former an entrance exam to high schools and the latter to colleges). The rest of the population who are not recruited into the elite sports training system gain no academic benefits from participating in sports. Unlike in America, where being good at sports and joining varsity teams can help students get into better universities, Chinese college entrance is based on academic grades alone. Therefore, separating talented athletes from the rest led to a lack of focus on sports in normal schools and a lack of focus on academics in athletic schools.

### **Analysis of the Policies**

In this section, I will analyze the three policies issued by the Chinese government. Between 2011 and 2020, the State Council issued the National Fitness Plan 2011-2015, hoping to boost the rate of participation in sports, especially of school-age children.

The goal of the plan was clear. The government hoped to "ensure that students take part in at least one hour of physical exercise every day during school."<sup>19</sup> However, the detailed plan of how to achieve this goal was ambiguous. The phrases "improve the working mechanism and supervision system for physical education" and "strengthen the construction of youth sports clubs" could mean investing more money, building more facilities, or training more physical education instructors. Having an ambiguous plan causes two problems.<sup>20</sup> First, different provinces might interpret the plans differently, leading to disparities in the plan's execution. While some people might think "strengthen" means building a youth sports club in every school, others might interpret it as building one or two. Second, without a clear standard, it is hard to evaluate whether the goals are actually achieved. Finally, China is a large and populous country with 23 provinces. One ambiguous plan will not fit the varied conditions around China.

Compared to the 2011 to 2015 plan, the 2016 to 2020 National Fitness Plan shifted the focus from building sports clubs to developing healthy behaviors and incorporating performance evaluation in school. This plan was also ambiguous, setting unclear goals and vague methods for promoting physical education and student participation in sport. Despite its ambiguity, there was a clear shift to focusing on school reforms instead of outside-school programs. The plan addressed the importance of "school education to help teenagers develop healthy behaviors" and suggested schools "incorporate students' physical health level into the assessment

---

<sup>18</sup> Yao Ming, Interview by Graham Bensinger, In depth with Graham Bensinger, Youtube, October 5, 2016, [https://www.youtube.com/watch?v=w\\_BT7\\_cCa90#action=share](https://www.youtube.com/watch?v=w_BT7_cCa90#action=share).

<sup>19</sup> Central Committee of the Communist Party of China, State Council of China, "National Fitness Plan 2011-2015," [in Chinese], 2011, The Xinhua News Agency. [http://www.gov.cn/zhengce/content/2011-02/24/content\\_6411.htm](http://www.gov.cn/zhengce/content/2011-02/24/content_6411.htm).

<sup>20</sup> Ibid.

system.”<sup>21</sup> The focus on school programs might be the result of the government realizing the improbability of having clubs. In-school programs were mandatory and more likely to be implemented by physical education instructors. Most students were not willing to participate in clubs that offered neither academic credits nor rewards unless they were genuinely interested in the sport. In 2013, research on sedentary and physical activity through the China Health and Nutrition Survey showed a positive correlation between students' attitudes towards physical education and physical activity, meaning that the willingness to participate in sports increased with an increase in positive attitudes towards sports.<sup>22</sup> One's attitude towards sports is formed at a young age and shaped by schools and family environment. By fostering the habit of athletic involvement from a young age, children are more likely to participate in sports when they are adults without having to be rewarded for doing so.

In contrast to the ambiguity of goals and plans of the 2011-2015 and 2016-2020 Programs, the Healthy China 2030 Plan included detailed instructions to promote sports and physical activity for individuals, families, schools, and government. On a personal level, the Healthy China 2030 Plan advised people to:

Ensure adequate physical activity and reduce sedentary and video time (watching TV, using computers, mobile phones, etc.). During the break, get out of your seat and do some exercise. Accumulate at least 1 hour of moderate-intensity exercise every day and develop a lifelong exercise habit.<sup>23</sup>

On the family level, the plan advised parents to “actively guide children to outdoor activities or exercise” and “encourage and support children to participate in various forms of sports activity outside school” so that children could develop a lifelong exercise habit.<sup>24</sup> The emphasis on individual and family involvement revealed the government's attempt to foster a more positive attitude towards sport by targeting people's daily habits.

At the school level, the plan suggested schools improve upon their physical education and extracurricular activity programs, “ensuring 4 class hours per week for primary and secondary grades, 3 class hours per week for grades three to six and junior high schools, and 2 class hours per week for senior high schools. Primary and secondary schools arrange 30 minutes of physical activity between classes every day.”<sup>25</sup> On the government level, the government promised to hold “targeted winter and summer camps, training camps, and sports events to attract children and adolescents to participate in sports.”<sup>26</sup>

---

<sup>21</sup> Central Committee of the Communist Party of China, State Council of China. “National Fitness Plan 2016-2020,” [in Chinese], 2011, The Xinhua News Agency, [http://www.gov.cn/zhengce/content/2016-06/23/content\\_5084564.htm](http://www.gov.cn/zhengce/content/2016-06/23/content_5084564.htm).

<sup>22</sup> Jiali Duan, Huanhuan Hu, Guan Wang, and Takashi Arao, “Study on Current Levels of Physical Activity and Sedentary Behavior among Middle School Students in Beijing, China,” *PLOS ONE* 10, no. 7 (2015): 8-9, e0133544, <https://doi.org/10.1371/journal.pone.0133544>.

<sup>23</sup> Central Committee of the Communist Party of China, State Council of China, “Healthy China 2030 Plan,” [in Chinese], 2016, The Xinhua News Agency, [http://www.gov.cn/xinwen/2016-10/25/content\\_5124174.htm](http://www.gov.cn/xinwen/2016-10/25/content_5124174.htm).

<sup>24</sup> *Ibid.*

<sup>25</sup> *Ibid.*

<sup>26</sup> *Ibid.*

Unlike the previous Nation Fitness Plans, the Healthy China 2030 Plan had clearer goals and provided more detailed instruction, making implementation easier and more practical. This plan also showed that the government understood that promoting physical education and sport participation could not rely on government intervention alone; instead, every person, family, and school needed to be a part of it. Fundamentally, if people were not aware of the benefits of sports, they were not likely to be willing to increase their participation in sports.

### **Participation in Sport and Physical Activity During and After the Plans**

The rate of participation in sport and physical activity underwent an overall improvement during the plans. Research in 2013 on middle school children in Beijing found that 76.6% of students reported having three 45-minute physical education classes every week. A total of 35.6% of students spent one hour per day performing moderate to vigorous physical activity (MVPA) during school, and 34.9% completed one hour per day of MVPA outside school time.<sup>27</sup> This was a significant improvement compared to the 2003 study, which found only 8% of school-aged children engaging in MVPA outside school.<sup>28</sup> Both studies found that the more urban the area the school was located, the more likely children were to participate in physical education programs and physical activities. This might be the result of the city having better sports facilities, more professionally trained physical education instructors, and more middle-class families that had surplus money and time to do sports. However, though children in rural areas participated less in sports programs, they did farm chores that were equivalent to sports programs in urban areas.

Still, the small increase in the rate of participation in sports and physical activities among school-aged children from 2003 did not meet the current physical activity recommendations set by the government, and about half of the students spent excessive time engaging in sedentary behaviors. In 2010, the World Health Organization recommended at least 60 minutes of moderate- to vigorous-intensity physical activity daily for youth aged 5–17.<sup>29</sup> In the 2013 study, less than 40% of the students fulfilled this recommendation.

There were also debates on whether government interventions were effective. From August 2015 to June 2016, an eight-month randomized control trial in China found that modification in school physical education policy and curriculum incrementally improved cardio-respiratory fitness and other physical fitness outcomes of students.<sup>30</sup> Yet, by comparing the physical activity participation rate before and after government involvement, another study found that government policies had little impact on promoting physical education and sport participation within or outside school. The study revealed that although numerous government policies promote physical education for schools, there was little or no evidence

---

<sup>27</sup> Jiali Duan, Huanhuan Hu, Guan Wang, and Takashi Arao, "Study on Current Levels of Physical Activity and Sedentary Behavior among Middle School Students in Beijing, China," 2015, Edited by Qinghua Sun, PLOS ONE 10, no. 7: e0133544, 4-6, <https://doi.org/10.1371/journal.pone.0133544>.

<sup>28</sup> C.Tudor-Locke, BE Ainsworth, LS Adair, S. Du, and BM Popkin. pg. 1096-1098

<sup>29</sup> "Global Recommendations on Physical Activity for Health". Genève: WHO, 2010. pg. 18-21

<sup>30</sup> Zhixiong Zhou, Shiyu Li, Jun Yin, Quan Fu, Hong Ren, Tao Jin, Jiahua Zhu, Jeffrey Howard, Tianwen Lan, and Zenong Yin, 2019, "Impact on Physical Fitness of the Chinese CHAMPS: A Clustered Randomized Controlled Trial," International Journal of Environmental Research and Public Health 16 22: 4412-4413. doi: 10.3390/ijerph16224412.

showing that these policies had increased the level of physical activity among school-aged children.<sup>31</sup>

## **Survey on Attitudes Towards School-Aged Children's Engagement in Sports and Physical Activity**

In August 2020, hoping to understand the current attitudes of adults towards school-aged children's participation in sports and discover factors that are preventing children from doing more sports, I conducted a self-reported survey on 245 people in China. The survey was distributed to people via the mini-application Survey Star (Wen Juan Xing) in WeChat, a Chinese social media application equivalent to WhatsApp. I sent the survey to twenty chat groups consisting of 30 to 400 people. I then collected the results of the survey via Survey Star and analyzed them. The results of the survey indicated that people thought sport participation among school-aged children was important and was hindered by the heavy academic pressure from the test-based education system in China.<sup>32</sup>

### **Participants**

The survey reached ten provinces in total, with the majority of participants from Guangdong (73.88%) and Hubei province (17.55%) and the rest from Beijing, Yunnan, Zhejiang, NingXia, Inner Mongolia, Shanghai, Jiangxi, and Henan (8.57%). The participants of the survey were 23.67% male and 75.51% female. This gender imbalance was likely the result of the uneven distribution of females and males in the group chats the surveys were sent to. Most participants resided in cities (92.24%), and a small population of participants lived in suburban areas or in the countryside (7.76%). I divided the participants into four age groups. Participants between 40 and 60 were in the majority, constituting 62.45% of the total population who took the survey. Due to senior people's limited access to technological devices, people who were over 60 only constituted 6.12% of the total population (See Figure 1).

---

<sup>31</sup> Peijie Chen, "Physical Activity, Physical Fitness, and Body Mass Index in the Chinese Child and Adolescent Populations An Update from the 2016 Physical Activity and Fitness in China the Youth Study," *运动与健康科学: 英文版* 6 4 (2017): 382, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6189246/?tool=pmcentrez&report=abstract>.

<sup>32</sup> Xinyue Luo, Survey On Attitudes Towards School-Aged Children's Engagement in Sports and Physical Activity, Survey Star, August 13rd to 19th 2020.



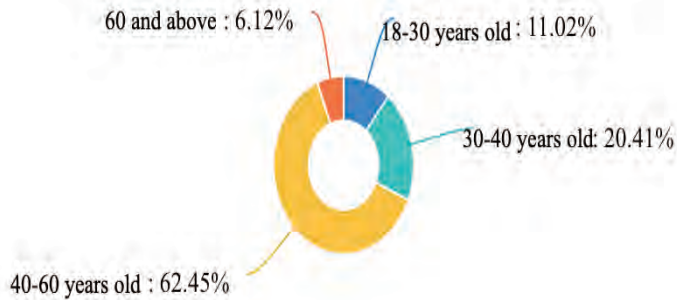


Figure 1. The age distribution of participants

Among the 245 people the survey reached, 66.53% reported having children under the age of 18. The profession of the participants varied a lot, with the majority being business managers, technicians, and company employees (See figure 2).

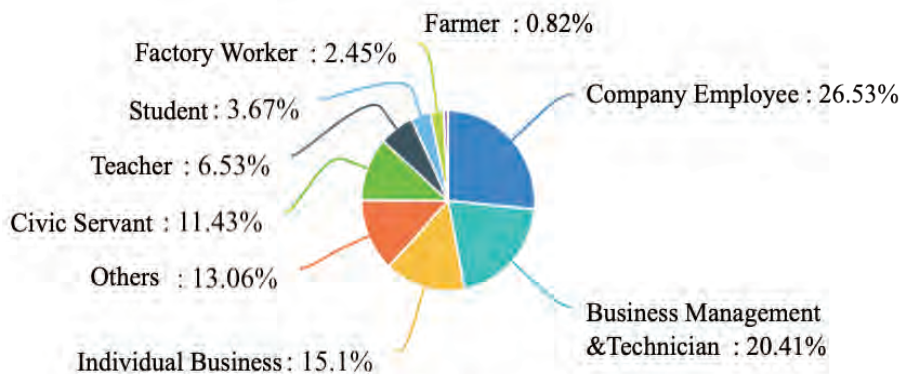


Figure 2. Profession distribution of participants

## Results

The survey showed more people were willing to support school-aged children's sport participation. 81.63% of the participants responded that sports were "very important, required every day." (See figure 3). The answers to this question showed that the participants had a positive attitude towards a child's engagement in sport. Similar attitudes emerged in other studies as well. In 2016, a study exploring perceptions of physical education found that high school principals in Shanghai showed strong support for the development of physical education in their schools.<sup>33</sup> The supportive attitude people showed towards sport participation was likely the result of scientific

<sup>33</sup> "Exploring Principals' Physical Education Perceptions and Views in Shanghai, China." *Research Quarterly for Exercise and Sport* 87.S2 (2016): A20-21. <https://search.proquest.com/openview/ca95572b2bde9a4d5e6656edf6db5d3f/1?pq-origsite=gscholar&cbl=40785>.

researches finding sports beneficial for our physical and mental health.<sup>34</sup> From 2015 to 2016, a randomized trial was conducted on twelve middle school students, attempting to find out whether the modification of school physical education policy and curriculum incrementally influenced the changes in cardio-respiratory fitness and other physical fitness outcomes. The study found that middle schools need to increase the opportunity for physical education in schools and to introduce high-intensity exercises in school-based PA programs for the fitness of the middle school students.<sup>35</sup> All these studies encouraged people to put more emphasis on sports and physical activity.

Surprisingly, my survey found that participants over forty-years-old deemed sport more important. Among the 50 participants aged 30 to 40, 74% chose the option “very important, required every day,” while among the 153 participants aged 40 to 60, 84.99% chose this option. Among the 15 participants over 60, 86.67% chose this option (See Figure 3). One possible reason behind this difference in opinion between different age groups was that people’s attitudes changed over time. As people’s age increased, they began to place higher value upon their health. When they looked back in time, they wished they could have exercised more so they might be in better health condition now. One participant reported, “when I was young, my parents cared little about physical education and academics. We just ran around in the fields. I wish I could have done more workout after I started working.” However, due to the small sample size, the data from this survey had limitations and could not represent the whole population accurately. To better understand the disparities between the different age groups, a larger scale survey covering a more representative population and geographic location is needed.

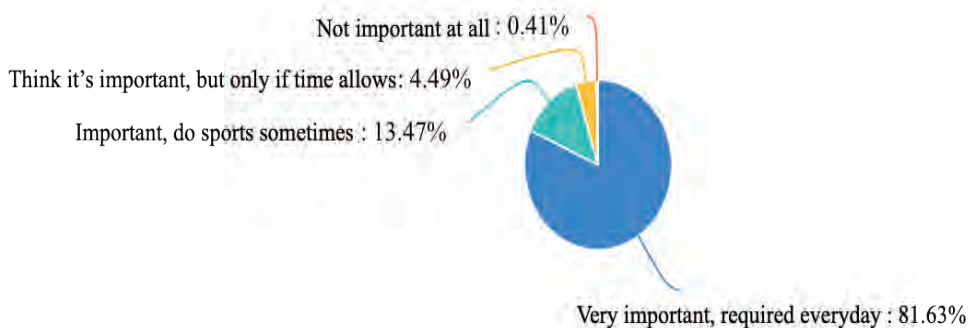


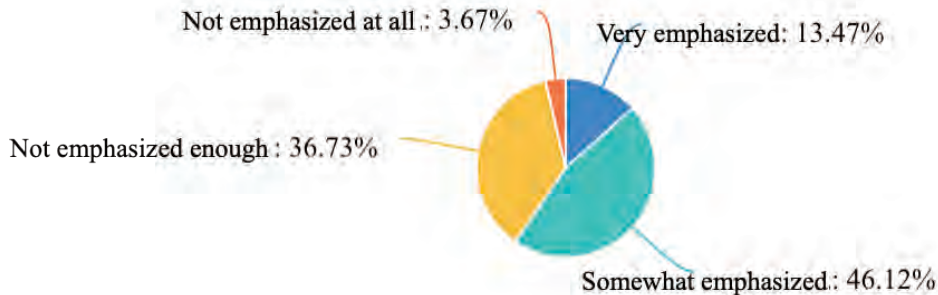
Figure 3. Answers to the question “Do you think sport is important for school-aged children?”

Many participants revealed that they thought sports education and activity were not emphasized enough in society. When asked the question, “Do you think

<sup>34</sup> Miriam Reiner, Christina Niermann, Darko Jekauc, and Alexander Woll, 2013, “Long-Term Health Benefits of Physical Activity – a Systematic Review of Longitudinal Studies,” *BMC Public Health* 13 (1): 813-814, <http://www.biomedcentral.com/1471-2458/13/813>.

<sup>35</sup> Zhixiong Zhou, Shiyu Li, Jun Yin, Quan Fu, Hong Ren, Tao Jin, Jiahua Zhu, Jeffrey Howard, Tianwen Lan, and Zenong Yin, “Impact on Physical Fitness of the Chinese CHAMPS: A Clustered Randomized Controlled Trial,” *International Journal of Environmental Research and Public Health* 16 22 (2019): 4412-4415, doi: 10.3390/ijerph16224412.

school-aged children's sport participation is emphasized enough," people's response showed that 36.73% of the participants found that "sport is not emphasized enough" for school-aged children (See Figure 4).



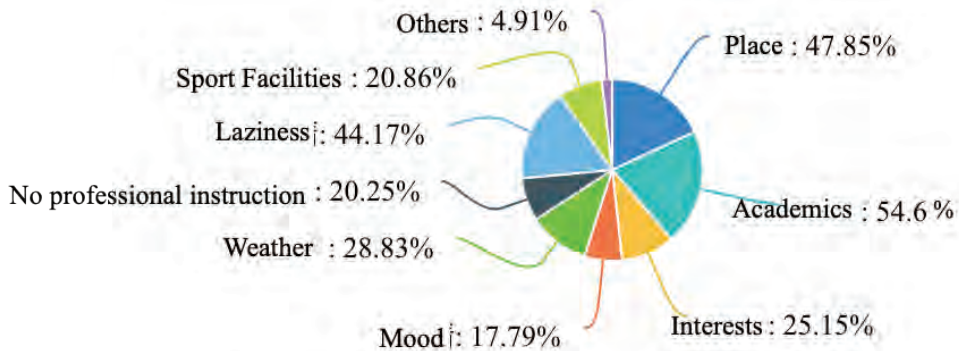
**Figure 4.** Answers to the question "Do you think school-aged children's sport participation is emphasized enough?"

The difference between people's self-rated importance of sports (81.63% reported "very important, required every day.") and the rated importance of sports they observed around them may be the result of multiple factors. First, there might be a group of people who think sports is important for school-aged children but feel hindered by some outer force. These outer forces include but are not limited to lack of nearby sports facilities, heavy academic pressure, and lack of incentives to do sports. Second, there might be a group of people who find sport important for school-aged children but happened to be living among people who have negative attitudes towards sport participation. This contributes to their perception that people around them were not putting enough emphasis on sports development for school-aged children. The first scenario is more likely to be the case since the probability that a great majority of participants lived among sports haters is low.

To better understand what was holding the parents back from letting their children participate in physical activities and sports, I looked specifically at the 163 (66.53%) participants who were parents of school-aged children. When asked "what factors restrict your child's participation in sports," 54.6% of the 163 parents indicated that "academics" were limiting their children's engagement in sports. Other important restrictions were "place" (47.85%) and "laziness" (44.17%) (See Figure 5). In addition, other factors included the lack of nearby facilities and sports fields as 20.86% of the parents reported being limited by the facilities, and 47.85% of the parents reported being limited by the number of places for sports around them. One participant suggested that the government should open more "free facilities and encourage organized group physical activities."

The "laziness" people described was not necessarily being too lazy to participate in sports. In many cases, "laziness" could mean being too lazy to drive half an hour to a gym and work out for half an hour. One participant suggested, "sports fields are in shortage. School playgrounds should be open to the public during weekends and holidays." Similarly, another participant complained, "there aren't enough football and basketball fields around." In a previous study on ten provinces in China, researchers found that the closer the sports facilities (measured in commute

time), the more willing the people were to exercise.<sup>36</sup> Therefore, to encourage more school-aged children to participate in sports and physical activity, government policies need to center around building more sports facilities in communities to allow sports to become more accessible to the public.



**Figure 5.** Answers to the question "What factors are restricting your child's participation in sports? (more than 1 answer allowed)."

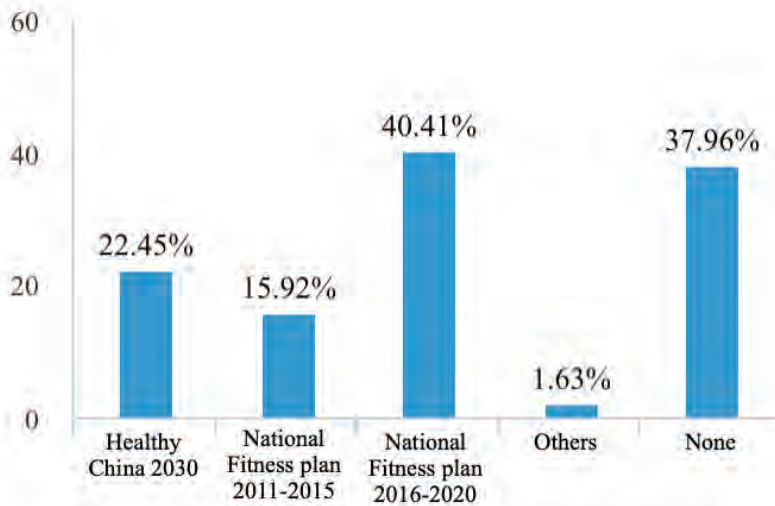
However, school-aged children's participation in sports and physical activities will not increase rapidly just by building more sports facilities. The academic pressure exerted on children by schools and parents will continue to hinder their participation in sports and physical activity. One participant said in the survey that "academic pressure is making time for sports impossible." Another participant called for the "prohibition of after-school tutorial classes such as XueErSi (A tutorial lesson company) and prohibition of video games so that children can have time to spend outdoors."

Only if the government policies seek to incorporate physical education and sports thoroughly into the academic system and the exam-based education system in China can sports become a part of school-aged children's lives. One participant in the survey suggested that "high schools and universities should recruit more students talented in sports." In China, a small proportion of students in each high school and university is recruited at a lower score to reward their excellence in sports. This is a relatively new policy aiming to increase the diversity of students in schools. Incorporating sports and physical activity into the education system does not trade off academic achievements. An experiment on a high school in Shanghai in 2015 showed a small increase in grades among male students after the school adopted a more rigorous and specialized physical education program.<sup>37</sup>

<sup>36</sup> Xiujin Guo, Jian Dai, Pengcheng Xun, Lynn M. Jamieson, and Ka He, "Sport Facility Proximity and Physical Activity: Results from the Study of Community Sports in China," *European Journal of Sport Science* 15, no.7 (2014): 667, doi:10.1080/17461391.2014.982203.

<sup>37</sup> Yunting Zhang, Xiaochen Ma, Jin Zhao, Hong Shen, and Fan Jiang, "The Effect of Strengthened Physical Education on Academic Achievements in High School Students: A Quasi-Experiment in China," *International Journal of Environmental Research and Public Health* 16 23 (2019): 2-4, doi: 10.3390/ijerph16234688.

In addition to asking people about their opinions on school-aged children's participation in sports, the survey also investigated whether they know about the Chinese government policies. Many people (40.41%) reported having heard of the National Fitness Program 2016-2020, while 37.96% of the people reported having neither heard of the National Fitness Program nor the Healthy China 2030 Plan (See Figure 6). This indicates that the government needs to put more effort into propagating these policies so that more people can understand and implement them.



**Figure 6.** Answers to the question "Have you heard of the following government policies."

## Conclusion

Before 1995, different political parties in China focused on the development of elite sports and the cultivation of talented athletes to compete on the global stage. The main incentives of the political parties were to improve China's status globally and to build patriotism within the country. After the PRC released the National Fitness Program in 1995, hoping to boost participation in sports for the whole population, school-aged children's participation in sports and physical activity gradually increased. However, despite the construction of more training fields and facilities, the rate of participation in sports and physical activities for school-aged children had remained low. The more recent policies (National Fitness Plan 2011-2015, National Fitness Plan 2016-2020, Healthy China 2030) further increased participation in sports and physical activities for school-aged children. The number of school-aged children who did sports both within and out of school has increased. However, the programs remained vague, making them hard to implement and assess. Nevertheless, a trend of improvement could be observed in the PRC's policies. The improvement was especially evident when comparing detailed plans' instructions in Healthy China 2030 Plan with the previous National Fitness Program's instructions. The changes,

including a more holistic plan dedicated to school-aged children in the program, indicated that the government had made an effort to promote sports among school-aged children.

My survey suggested a readiness by adults to accept more emphasis on physical education in school and a positive attitude towards sport participation. It also revealed that the education system was a huge roadblock for further increasing sports participation among school-aged children. The government also needed to put more effort into building more port facilities and sports fields.

### Suggestions

To boost participation in sports and physical activity among school-aged children, the government should make the following additional changes. The government needs to move further away from Elite-only training programs and to emphasize community sport and school-aged children's sport. The government needs to incorporate sports into the exam-based education system since academics exert such a huge influence on the incentives of school-aged children's participation in sports. Including sports in the college admission process would be a place to start. This can be done by including PE assessment scores as part of the grades for college admission or by establishing more varsity teams in colleges. In addition, more funds need to be directed towards the construction of sports facilities and sports fields. Lastly, the government needs to spend more effort in promoting these programs and plans released by the State Council, making sure that all provinces in China engage in this process of building a national pro-sport culture.

### Bibliography

- Cao, Xiang Fan and Zhen-Bo. "Physical Activity among Chinese School-Aged Children: National Prevalence Estimates from the 2016 Physical Activity and Fitness in China the Youth Study." *Journal of Sport and Health Science* no.6 (2017): 388–394. doi: 10.1016/j.jshs.2017.09.006. Accessed July 8, 2020.
- Central Committee of the Communist Party of China, State Council of China. "Healthy China 2030 Plan," [in Chinese]. 2016. The Xinhua News Agency. [http://www.gov.cn/xinwen/2016-10/25/content\\_5124174.htm](http://www.gov.cn/xinwen/2016-10/25/content_5124174.htm). Accessed July 21, 2020.
- Central Committee of the Communist Party of China, State Council of China. "National Fitness Plan 2011-2015," [in Chinese]. 2011. The Xinhua News Agency. [http://www.gov.cn/zhengce/content/2011-02/24/content\\_6411.htm](http://www.gov.cn/zhengce/content/2011-02/24/content_6411.htm). Accessed July 21, 2020.
- Central Committee of the Communist Party of China, State Council of China. "National Fitness Plan 2016-2020," [in Chinese]. 2011. The Xinhua News Agency. [http://www.gov.cn/zhengce/content/2016-06/23/content\\_5084564.htm](http://www.gov.cn/zhengce/content/2016-06/23/content_5084564.htm). Accessed July 21, 2020.
- Chen, Peijie. 2017. "Physical Activity, Physical Fitness, and Body Mass Index in the Chinese Child and Adolescent Populations: an Update from the 2016 Physical Activity and Fitness in China the Youth Study." *运动与健康科学: 英文版* 6 (4): 381-383. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6189246/?tool=pmcentrez&report=abstract>. Accessed July 8, 2020.



- Chepyator-Thomson, Jepkorir Rose, and Hsu, Shan-Hui, eds. 2013. *Global Perspectives on Physical Education and After-School Sport Programs*. Lanham, MD: UPA. ProQuest Ebook Central. Accessed July 7, 2020.
- "China: Going for Gold." directed by Anonymous, produced by Real Life Productions. Teachers TV/UK Department of Education, 2007. Alexander Street, <https://video.alexanderstreet.com/watch/china-going-for-gold>. Accessed July 18, 2020.
- Duan, Jiali, Huanhuan Hu, Guan Wang, and Takashi Arao. "Study on Current Levels of Physical Activity and Sedentary Behavior among Middle School Students in Beijing, China." *PLOS ONE* 10, no. 7 (2015): e0133544. <https://doi.org/10.1371/journal.pone.0133544>. Accessed July 8, 2020.
- "Global Recommendations on Physical Activity for Health". Genève: WHO, 2010. Accessed July 7, 2020.
- Guo, Xiujin, Jian Dai, Pengcheng Xun, Lynn M. Jamieson, and Ka He. "Sport Facility Proximity and Physical Activity: Results from the Study of Community Sports in China." *European Journal of Sport Science* 15, no.7 (2014): 663-669. doi:10.1080/17461391.2014.982203. Accessed July 8, 2020.
- Houlihan, Barrie, and Green, Mick, eds. 2011. *Routledge Handbook of Sports Development*. London: Taylor & Francis Group. ProQuest Ebook Central. Accessed July 7, 2020.
- "How do they do it in China?" directed by Anonymous, produced by Real Life Productions. Teachers TV/UK Department of Education, 2006. Alexander Street. <https://video.alexanderstreet.com/watch/how-do-they-do-it-in-china>. Accessed July 18, 2020.
- Kogh, Y. K. "A Descriptive Trend Analysis of Government Policy on Physical Education in the People's Republic of China: 1949-1965." (PHD diss., University of Denver, 1987). <http://search.ebscohost.com.ezproxy.bowdoin.edu/login.aspx?direct=true&db=hia&AN=46184881&site=ehost-live>. Accessed July 28, 2020.
- Lu, Zhouxiang. "From Hongkew Recreation Ground to Bird's Nest: The Past, Present and Future of Large Sports Venues in China." *The International Journal of the History of Sport*. vol. 30, no. 4 (2013): 422-4, Taylor and Francis+NEJM. doi:10.1080/09523367.2013.765724. Accessed July 19, 2020.
- Ming, Yao. Interview by Graham Bensinger. In depth with Graham Bensinger. YouTube, October 5, 2016. [https://www.youtube.com/watch?v=w\\_BT7\\_cCag0#action=share](https://www.youtube.com/watch?v=w_BT7_cCag0#action=share). Accessed July 8, 2020.
- Morris, Andrew David. "Cultivating the National Body: A History of Physical Culture in Republican China." (PHD diss., University of California San Diego, 1998) ProQuest Dissertations Publishing. Accessed July 29, 2020.
- Pickett, Karen, Tony Kendrick, and Lucy Yardley. 2017. "'A Forward Movement into Life': A Qualitative Study of how, Why and when Physical Activity may Benefit Depression." *Mental Health and Physical Activity* 12: 100-109. doi: 10.1016/j.mhpa.2017.03.004. Accessed on August 31, 2020.
- Reiner, Miriam, Christina Niermann, Darko Jekauc, and Alexander Woll. 2013. "Long-Term Health Benefits of Physical Activity – a Systematic Review of Longitudinal Studies." *BMC Public Health* 13 (1): 813-813. <http://www.biomedcentral.com/1471-2458/13/813>. Accessed on August 31, 2020.
- Shuman, Amanda. "Elite Competitive Sport in the People's Republic of China 1958-

- 1966: The Games of the New Emerging Forces (GANEF0)." *Journal of Sport History* 40, no. 2 (2013): 258-83. <http://www.jstor.org/stable/10.5406/jsporhistory.40.2.258>. Accessed July 2, 2020.
- Tudor-Locke, C., BE Ainsworth, LS Adair, S. Du, and BM Popkin. "Physical Activity and Inactivity in Chinese School-Aged Youth: The China Health and Nutrition Survey." *International Journal of Obesity* 27, no.9 (2003): 1093-1099. doi: 10.1038/sj.ijo.0802377. Accessed July 8, 2020.
- Xinyue Luo, Survey On Attitudes Towards School-Aged Children's Engagement in Sports and Physical Activity, Survey Star, August 13rd to 19th 2020.
- Zeng, Zhenhao, and Xiaozan Wang. "Exploring Principals' Physical Education Perceptions and Views in Shanghai, China." *Research Quarterly for Exercise and Sport* 87.S2 (2016): A20. <https://search.proquest.com/openview/ca95572b2bde9a4d5e6656edf6db5d3f/1?pq-origsite=gscholar&cbl=40785>. Accessed July 21, 2020.
- Zhang, Yunting, Xiaochen Ma, Jin Zhao, Hong Shen, and Fan Jiang. "The Effect of Strengthened Physical Education on Academic Achievements in High School Students: A Quasi-Experiment in China." *International Journal of Environmental Research and Public Health* vol.16, no.23 (2019): 4688. doi: 10.3390/ijerph16234688. Accessed July 21, 2020.
- Zhouxiang, Lu. "From Hongkew Recreation Ground to Bird's Nest: The Past, Present and Future of Large Sports Venues in China." *The International Journal of the History of Sport*, vol. 30, no. 4 (2013): 422-42. Taylor and Francis+NEJM. doi:10.1080/09523367.2013.765724. Accessed August 4, 2020.
- Zhouxiang, Lu. "Sport, Nationalism and the Building of the Modern Chinese Nation State (1912-49)." *The International Journal of the History of Sport*, v28, no. 7 (2011):1030-54. doi:10.1080/09523367.2011.563635. Accessed August 4, 2020.
- Zhou, Zhixiong, Shiyu Li, Jun Yin, Quan Fu, Hong Ren, Tao Jin, Jiahua Zhu, Jeffrey Howard, Tianwen Lan, and Zenong Yin. 2019. "Impact on Physical Fitness of the Chinese CHAMPS: A Clustered Randomized Controlled Trial." *International Journal of Environmental Research and Public Health* 16 (22): 4412-4433. doi:10.3390/ijerph16224412. Accessed July 21, 2020.





# **The Slippery Slope: Naturalistic Observations into Nonverbal Manifestations of Formed Dominance Hierarchies Among Children at Play**

---

*Xuan Jiang*

**Author Background:** *Xuan Jiang grew up in China and currently attends The Experimental High School Attached to Beijing Normal University in Beijing, China. Her Pioneer research concentration was in the field of anthropology and titled "Nonverbal Communication."*

## **Abstract**

Dominance-subordination hierarchies are known to form naturally in groups of children once they are gathered together. However, these hierarchies have a self-reinforcing nature which, if left unchecked, may gradually shift them into systems of bullying and segregation. This study was conducted with the goal of building a data-supported framework with which parents or educators could promptly recognize over-dominant and over-subordinate children to prevent that happening. Over a period of four weeks, a naturalistic observation was conducted examining the nonverbal manifestations of dominance hierarchies among small groups of children at play in a Beijing neighborhood. The participants were thirty 5 to 7-year-olds, observed in groups of 3-6. The children's individual relative dominance levels were calculated and documented along with their specific nonverbal behavior. A five-part analysis was then conducted to identify specific behavioral sets of dominant and subordinate children in the dominance hierarchy. It is proposed that dominant children stand in more open postures, use more and bigger gestures, and initiate more touches when play fighting. They also tend to speak louder and more fluently. Subordinate children were found to directly orient to their peers more when interacting and to self-touch more. The causes, interpersonal significance, and further implications of particular nonverbal behaviors, often with regard to their functions of reinforcing dominance-subordination hierarchies, is also discussed.

## 1. Introduction

From as young as 3 to 5 years-old, children gathered in groups begin to form dominance-subordination hierarchies (Halpern, Jodi et al., 2015). As with adults, these dominance relations between children serve to add stability and predictability to social relationships, aid in the distribution and use of limited resources, and define roles for group members (Savin-Williams, 1979). However, as dominance relations are reinforced and positions on the social ladder are polarized, some children begin to suffer lasting harms outweighing the benefits of a balanced hierarchy. “Health, safety, and behavior in early development,” as summarized by W.T. Boyce (2004), “can be deeply affected by subjective childhood experiences of occupying greater or lesser positions on a scale of social influence” (p. 47).

For example, take the overly-subordinate child in the dominance hierarchy. Research shows he is likely to develop resentments and compensatory drives that influence his behavior for years to come. He has higher rates of chronic medical conditions or violent injuries (Boyce 2004), higher levels of anxiety (Ammon A.V., Filippova E.V., 2015), and a higher likelihood of becoming bullied, because “subordinates become implicitly acceptable targets for exclusion and rejection” (Halpern, Jodi et al., 2015, p.24). Once bullying occurs – 20% to 30% of school children are frequently involved in it (Klomek et al., 2007) – the child suffers increased risks of depression and low self-esteem (Fast, 2016). In fact, “the enduring stress response patterns of childhood subordination suggest that early childhood may represent a developmental ‘critical period’ during which social experiences have disproportionately large effects” (Halpern, Jodi et al., 2015, p.26). On the other end of the spectrum, due to stability of the aggressiveness trait (Huesmann et al., 1984), the dominant king of the playground usually grows into an overaggressive adult. The ones who become bullies more often grow up into young adults suffering from antisocial personality disorder, substance abuse, depressive and anxiety disorders (Sourander, 2007), as well as suicide ideation (Klomek et al., 2009). An overly segregated dominance hierarchy is a lose-lose situation, yet they have been found to self-reinforce and aggregate by nature (Magee & Galinsky, 2008). In fact, the very behavioral sets of subordinate children to be examined in this paper contain “subtle signals that they will internalize shame rather than fight back” (Fast, 2016, p. 39) which are part of the very reason children become bullying victims.

Several intervention methods are known for preventing such unfortunate outcomes, such as increasing children’s understanding of aggressive tactics and prosocial skills (Hawley, 2007), and promoting positive relationships with adults (Pianta, 1999). But the most fundamental issue in implementing such methods still stands: how do we recognize in time the specific children who need them? More than half of bullying victims never disclose their experiences to parents or teachers (Smith & Shu, 2000; Sandra & Petrie, 2002). And one cannot always wait until the damage is done and realized to act. Thus, there is a critical necessity in understanding and recognizing the behavioral manifestations of dominance hierarchies. According to ethologist Donald Omark’s 1979 review of dominance relations, once one’s position in the pecking order is acquired, a set of predictable behavioral responses by both the individual and other members of the group ensue. In Europe and America, abundant research on nonverbal dominance/subordination communication of both adults and children has been conducted, and much of it is cited in this study. However, there is comparatively

less literature analyzing even only general trends of nonverbal behavior in East Asia, much less China in particular. Presently, it is known that East Asians touch with lower frequency than other groups (McDaniel & Anderson, 1998, Regan et al., 1998), have different concepts of space from Americans (Lv, 2017), and exhibit higher frequencies of deference behavior (bowing, gaze avoidance and politeness routines) (Knapp & Hall, 2013). But few endeavors have been made to analyze specifically the nonverbal communication of dominance relations among East-Asian children. Describing such signature nonverbal behavioral sets of dominant and subordinate Chinese children so that parents and educators may recognize them in those needing intervention is thus this study's main goal.

## 2. Methodology

The participants of the study were thirty 5-to-7-year-old children of a middle-class apartment area in Beijing, China. Unlike the majority of past research in this field, this study was conducted as a naturalistic observation, so as to acquire realistic pictures of participant behavior. The thirty children were observed in five child-organized small groups of 3-6, where dominance hierarchies are most liable to form and most easily observable. The context of observation was a large courtyard in the children's same neighborhood.

The operationalization of dominance was a critical and complicated issue in this study. Because the objective of this study is to examine nonverbal behaviors as results of established dominance hierarchies among children, the behaviors themselves were not to be used as immediate means of measuring dominance. Instead, children's relative dominance levels were determined by quantifying their relative ability to derive favorable results from social situations within the group. This was done by counting the number of times they either terminated episodes of social conflict favorably, or made self-benefitting decisions for other members of the group, during the observational period. Each child's Relative Dominance Level (RDL) was then calculated by dividing the number of times they derived Favorable Results (FR) by the number of Total Social Situations (TSS) participated in ( $RDL = FR/TSS$ ). Children whose relative dominance level was higher than 0.5 were deemed dominant (or "higher-dominance"), and children whose relative dominance level was lower than 0.5 subordinate (or "lower-dominance"). To ensure validity, data on children who were involved in fewer than five social situations during the entire observational period were excluded from analysis. This operationalization of dominance was derived in part from that of Strayer and Strayer (1976), as the relative ability to terminate episodes of social conflict favorably. It was modified so as to include instances where children used prosocial behaviors instead of solely aggressiveness in conflicts to indicate dominance. Their definition of an "episode of social conflict" was still applied: any instance in which one child physically attacks or interferes with another's person, activities, or materials, or threatens by word or gesture to do so, or endeavors by force or verbal demand to possess another's belongings or to direct another's activities in opposition to the apparent desires of the other child (Strayer and Strayer, 1976). "Self-benefitting" is defined as consistent with the apparent desires of the individual child.

In addition to relative dominance levels, the children's nonverbal behaviors were also documented and summarized. Specific, quantifiable behaviors of individuals such as frequency of gesturing, number of touches initiated, or number of pauses in speech were counted during one-hour natural

interactions between groups. More complex behaviors, such as length of eye gaze while speaking, were quantified during ten-minute interactions. Uncountable behaviors, for instance openness of posture, were described qualitatively. Some representative sets of such data will be presented throughout the paper, the children represented by alphabetical aliases. Deductions were then proposed from the analysis into relationships between relative dominance and nonverbal behaviors, with the objective of describing specific behavioral sets of dominant and subordinate children. Child behavior was analyzed in relation to dominance levels in five sections: posture, gestures, touch, eye gaze, and vocal cues. Some preliminary analyses of causes, significance, and hierarchy-reinforcing effects of individual nonverbal behaviors were also conducted.

### 3. Posture

#### 3.1 General characteristics of posture

Two characteristics were examined in the research into posture: whether or not the children stood straight and whether or not they were reserved in their posture. Children of two groups were observed for their relative dominance levels and general posture characteristics during videotaped one-hour interactions. The data was then organized into Table 1.

**Table 1.** *Dominance and posture in two groups of children over videotaped one-hour interactions*

Characteristic /Participants	Group 1				Group 2				
	Alice	Bob	Cat	Dave	Ed	Fred	Greg	Hans	Ivy
FR/TSS	2/9	10/11	0/5	6/9	3/8	5/8	1/5	2/8	8/11
Relative dominance level	0.22	0.91	0.00	0.67	0.38	0.63	0.20	0.25	0.73
Straight	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes
Reserved	Yes	No	Yes	No	Yes	No	Yes	Yes	No

As apparent from Table 1, no correlation was found between standing straighter and having higher relative dominance. Only two of the 9 children observed (Fred & Greg) did not stand straight, and they were one dominant (RDL=0.63) and one subordinate (RDL=0.2). Openness of posture, on the other hand, differed significantly between dominant and subordinate children. All higher-dominance children observed (Bob, Dave, Fred & Ivy) indeed stood in less reserved and more open postures. These open postures were marked by relaxed arms, standing with legs apart, taking larger strides when walking, and more general expansiveness. Lower-dominance children (Alice, Cat, Ed, Greg & Hans) kept reserved and contractive poses. Characteristics of such "reserved" posture were identified as keeping the knees together, keeping the arms close to sides, and the muscles being tense around the shoulders. Some subordinate children were observed to also keep their hands clasped together instead of naturally hung. Such differences are consistent with past research in adults suggesting that greater

degrees of postural relaxation denote increases in potency or status (Mehrabian, 1972). Although some subordinate children were seen to limit the movement of their arms with visible reserve, it is still unclear whether their reserved postures are a subconscious reflection of internal shyness, or a conscious effort to present a certain image.

Observations into posture also revealed a more surprising accidental finding; some (not all) subordinate children tended to stand straighter when joining a new group. They intentionally kept their heads up and shoulders back, in a manner observably different from when in familiar groups. One such child seemed to have caught his own shoulders slumping during one observation and straightened his posture suddenly, a behavior never exhibited among familiar peers. Interestingly, despite standing straighter in a move suggesting confidence, these subordinate children still maintained reserved postures in the new groups. They continued to keep their arms close to their sides and knees close together, suggesting subordination. It's likely these self-contradictory messages are indicators of "an imperfect job of lying" (Knapp & Hall, 2013, p.16). The head and the shoulders are more likely to be attentively managed by a child aiming to feign confidence around unfamiliar peers, because they are so commonly associated with confidence ("Straighten up!" Chinese mothers are prone to teaching). Limb behavior, by contrast, is seldom directly connected to appearances of dominance in children's heads. Not to mention that standing straight is easier to accomplish than forcing naturally reserved limbs to act more open. Continuously dominant children, on the other hand, also maintained their naturally open and expansive postures when coming into new groups, taking larger strides and standing with muscles relaxed. These observations could suggest a strong persistence of the postural trait of openness across situations.

### 3.2 Body orientation

Body orientation is one of the more direct nonverbal indicators of relative dominance. It is, to start with, a most natural and direct behavioral display of attention, second only to eye gaze, which will be discussed separately. When paying attention to something, people instinctively look and orient towards it. Table 2 records the relative dominance of four children in a small group, the number of peer interactions in which they directly orient to the other party, and the number of interactions in which they do not directly orient to the other party. Interactions where direct body orientation was subjectively impeded, such as when walking side by side, were excluded from documentation.

**Table 2.** Dominance and body orientation in a group of children over a videotaped one-hour interaction

Characteristics/Participants	Jason	Karen	Larry	Mary
FR/TSS	5/8	2/7	2/8	5/7
Relative dominance level	0.63	0.29	0.25	0.71
Directly oriented	37	30	38	41
Not directly oriented	25	5	17	21

To summarize, the more dominant children of the group, Jason and Mary, oriented directly to their peers in 59.7% (37/62) and 66.1% (41/62) of interactions, respectively. The subordinate children, Karen and Larry, oriented directly to their peers in 85.7% (30/35) and 69% (38/55) of interactions. On average, the subordinate children oriented directly to their peers in 23% more interactions than dominant children. Results generally suggest direct body orientation negatively correlating with dominance. This is to say, dominant children are less likely to directly orient to their average peer than subordinate children when interacting. This is consistent with similar conclusions drawn from cohorts of younger children (preschoolers) (Gage and Lieberman, 1978), as well as findings that low-status individuals pay more attention to high-status individuals than vice versa (Anderson et al., 2001). Higher-dominance children are liable to orient away when communication, and more often signal that their attention, also, veers from who they are speaking to. However, it is worth noting that this study did not differentiate between high-dominance/low-dominance interactions and similar-dominance interactions. Thus, it is still not conclusive whether low-dominance children more often directly orient to higher-dominance children than vice versa, but only that lower-dominance children have a greater likelihood of directly orienting to a given interactant than higher-dominance children.

Throughout the observational timeframe, the children also showed a surprising attentiveness to each other's body orientation, contrasting with the minimal registering they did of other nonverbal communication forms. Those who were not directly oriented to in conversation sought it out, often moving themselves in front of their inattentive peers so that they faced each other. Rarely, however, did they appear to experience a conscious decision-making process before doing so.

The surprising attentiveness children pay to bodily orientation may be a result of its psychological functions. More than attention, researchers have suggested direct body orientation as an indicator of greater closeness or liking, and a lower frequency of it to be associated with less intimacy or even disliking (Mehrabian, 1972). Though they do not realize it, direct body orientation from peers likely has such importance to children because it represents a sense of intimacy children need from peer relationships. For the over-dominant child who upsets his peers, for instance, more directly orientation to other children when interacting may have the effect of strengthening his relationships and improving a dominance hierarchy that has excessively divided its members.

#### **4. Gestures**

Gestures accentuate speech – sometimes even substituting it. When speaking, gestures help the user provide clarity for the audience, and demonstrate an investment in the conversation which connects the two (Knapp & Hall, 2013). In adults, studies have found the use of them to correlate with dominance (Gifford, 1994). In this study, observations were conducted on dominance and the use of gestures in two groups of children to examine whether the same is true of children. The data collected were then organized into Table 3.

**Table 3.** *Gestures and dominance in a group of children over a videotaped one-hour interaction*

Characteristics/Participants	Nathan	Oliver	Patrick	Quincy
FR/TSS	2/8	9/12	8/11	1/5
Relative dominance level	0.25	0.75	0.73	0.20
Gestures used	61	93	105	45
Gesture size	Average	Large	Large	Small
Imitated gestures	29	33	29	17
Percentage of imitated gestures	0.48	0.35	0.28	0.38

The two dominant children of the group observed, Oliver and Patrick (RDL=0.75; RDL=0.73), used on average of 101 gestures during the one-hour observation, more than twice as many as the average 48 of subordinate children Nathan and Quincy (RDL=0.25; RDL=0.2). The gestures they used were also large, contrasting with those of Quincy and Nathan which were small and average, respectively. From this we may infer that dominant children gesture with higher frequency and bigger movements, while less dominant children use fewer and smaller gestures.

To further interpret these large and frequent gestures used by dominant children, they could be seen as an extension of these children's tendency towards more open and expansive postures discussed previously. The same could be said of the small, infrequent gestures and reserved postures of subordinate children. Whether the subordinate children act this way because they are inexpressive by nature or intentionally self-inhibiting, these nonverbal messages have the combined effect of making the subordinate child physically take up less space than their higher-dominance peers. This manifestation of subordination traits is likely one example of a behavior which could reinforce their subordination by subconsciously encouraging their peers to treat them as the unimportant figures they present themselves to be.

The children were also observed to frequently imitate each other's gestures. For instance, when one child clasped his hands with thumb and forefinger out to make a shooting motion, his peers followed suit. A shooting fight soon began and recurred throughout the day as soon as any one of them began making the gesture again. Among the children of Table 3, dominant peers Patrick and Oliver used 29 and 33 imitated gestures, making up 28% and 35% (average 31.5%) of total gestures used. This is lower than the 38% and 48% (average 43%) imitated gestures of subordinate children Nathan and Quincy, who used 17/45 and 29/61. We may deduce that subordinate children are more likely to imitate their peers' gestures than dominant children.

The majority of imitated gestures belonged to a specific category, called emblems – “speech-independent gestures...nonverbal acts that have a direct verbal translation or dictionary definition” (Knapp & Hall, 2013, p. 201), such as the “shooting” gesture described above. It's likely they are so quickly imitated because they are illustrative and highly consciously perceived (Knapp & Hall, 2013). Other speech-dependent gestures are mostly used and received subconsciously, but subordinate children can be imitative of them also. In one observation, for instance, a dominant child with a tendency to put a fist to his chest

when referring to himself did so several times in the group, and a subordinate child mimicked the behavior minutes later without being aware of it. This could be a reflection of a larger tendency of lower-dominance children to pay comparatively more conscious and unconscious attention into registering and imitating dominant behavior.

## **5. Touch**

Despite touch being known as a critical form of nonverbal communication in children's development, touching between peers occurred surprisingly infrequently in the study conducted. Past research showed adults to touch each other very rarely in public spaces (Hall & Veccia, 1990; Remland, Jones & Brinkman, 1991). From the observations of this study, it seems that despite children generally touching more than adults (Willis & Hoffman, 1975; Willis & Reeves, 1976), they already have some of the same tendency. Nonetheless, two categories existed where touching was frequent and differences between dominance levels were prominent: play fighting and self-touching.

### **5.1 Play fighting**

Play fighting is, as the name suggests, the nonverbal behavior of engaging in the playful fights frequent among children. One characteristic of children's play fighting behaviors often demands immediate addressing: the gender differences. 9 of the total 30 participants (half female, half male) engaged in play fighting; 8 of those 9 children were male, and only one was female. In other words, 7% of females participated in play fighting compared to 53% of males in this study. The remaining females touched their peers only to refer to an article of clothing, to play hand clapping games, or to get their attention with what is known as a management touch (Knapp & Hall, 2013). These touches were infrequent and showed little differences across dominance levels. For boys, however, play fighting is nearly the only cause for them to touch their peers, including only a handful of management touches. This resulted in considerably greater dominance differences in general touching behavior among boys than girls, due to the significant dominance differences in play fighting behavior this study will now examine. Table 4 documents the touches used in a ten-minute play fight between three boys significantly divided in their dominance levels.



**Table 4.** *Touching and dominance in a group of children over a ten-minute interaction/play fight*

Characteristics/Participants	Stuart	Ted	Udo
FR/TSS	2/7	9/10	0/5
Relative dominance level	0.29	0.90	0
Normal touches	4	3	3
Shoves	0	4	1
Punches	2	20	0
Other aggressive touches	11	13	4
Total aggressive touches	13	37	5

The dominant child of the triad in Table 4, Ted (RDL=0.9), initiated 37 aggressive touches, including 24 punches and shoves, within the ten minutes of play fighting. This is 4.9 times as many aggressive touches as the average 7.5 of his two peers, Stuart and Udo (RDL=0.29; RDL=0). Only 3 normal, nonaggressive touches were used by Ted during this time, showing small numerical differences with the respective 4 and 3 normal touches used by his lower-dominance peers. Fractionally, however, this dominant child touched aggressively 92.5% of the time in a play fight, higher than the subordinate children's 76.4% and 37.5%. In short, dominant children were found to touch aggressively much more in a play fight, with both higher numbers of aggressive touches and a higher percentage of aggressive touches in the total. This is consistent with past findings that dominant children initiate more touches (Gage and Lieberman, 1978). Furthermore, in the instance documented, the play fighting occurred in 3 segments, and all three times were initiated by the dominant child, Ted. The subordinate children, on the other hand, took on a more defensive stance, initiating fewer offensive touches than they receive. Although more observations will be needed before drawing conclusions, it may be hypothesized that dominant children more often initiate play fights than subordinate ones. Contrastingly, in ideal play fights between children of generally equal dominance, the fight should constitute of about equal attacks and provocations on either side. The singular case presented in Table 4 clearly is not one such example. And there is an unfortunate consequence of such obvious dominance gaps in play fights – it reinforces the established hierarchy with unparalleled efficiency, at times to the point of turning a dominant/subordinate relationship into a bully/victim one.

As mentioned in the introduction, children become bullying victims in part because they give off subtle signals that they will internalize shame rather than fight back (Fast, 2016). In a play fight like that of Table 4, this means that the mild retaliations marking a lower-dominance child will act as an unintentional declaration of submissiveness, and prompt more aggressive attacks from the dominant party. In this sense, play-fighting behaviors are the most obvious simultaneous reflections of and constructors of dominance out of all nonverbal communication forms discussed in this paper. While the other forms are largely encoded and decoded subconsciously, play fighting is direct physical conflict,

accelerating the widening of the dominance differential dangerously quickly. It is then inferable that the small group documented in Table 4 has already undergone several rounds of this process to reach the current level of disparity, having shown high levels of familiarity with each other and the play fights among them. If nonverbal manifestations of dominance in general are a slippery slope into bullying, play fights are its steepest incline. This would certainly explain why 65% of bullying victims report their most frequent place of bullying to be the playground (Smith & Shu, 2000).

## 5.2 Self-touching

Another subcategory of touching, self-touching, is an equally reflective and intriguing mode of the observed children's nonverbal behavior to play fighting, but it is associated with nearly entirely opposite sentiments. In general, self-touching belongs to a wider range of self-focused manipulation known as adaptors or self-manipulators, behavioral adaptations in response to certain situations. There is consensus that adaptors are generally associated with negative feelings (Knapp & Hall, 2013). The self-touching behaviors of some children of different dominance levels are documented in Table 5.

**Table 5.** *Self-touching and dominance in a group of children over a videotaped one-hour interaction*

Characteristics / Participants	Quincy	Patrick	Ron	Nathan	Oliver
FR/TSS	1/5	5/8	2/7	2/8	8/11
Relative dominance level	0.20	0.63	0.29	0.25	0.73
Self-touching	42	46	92	77	35

On average, the children observed self-touched 58.4 times in during the one-hour observation, nearly once every minute. The abundance of self-touching behaviors observed include pulling at the front of shirts/dresses, pulling at the bottom of trousers, scratching the nose, fiddling with fingers, clapping the arms repeatedly onto the sides, and even grabbing at one's own face. Recurrent behaviors specific to girls were tugging at ponytails and brushing hair back including when it was not obscuring the face. Subordinate children, Quincy, Ron & Nathan (RDL=0.2; RDL=0.29; RDL=0.25), exhibited such self-touching behaviors 42, 92, and 77 times, respectively, averaging 70.3 self-touches over the hour. The dominant children, Perry & Oliver (RDL=63; RDL=0.73), self-touched 46 and 35 times, averaging 40.5. We may infer that subordinate children self-touch more frequently than dominant ones, and this would presumably be due to a combination of dispositional and situational factors.

Dispositions-wise, people with trait anxiety, such as those who are shy but also have a desire to be sociable, self-touch more (Cheek & Buss, 1981). This is fitting since subordinate children have generally higher levels of anxiety than higher-dominance ones (Ammon A.V., Filippova E.V., 2015). Another explanation is that people self-touch more frequently in situations invoking situational anxiety

or stress (Knapp & Hall, 2013), and subordinate children are more often in such situations. In the observations conducted, dominant children were more consistently followed around by peers, while subordinate ones were more often left on the perimeter of the group looking on, or alone and uncertain where to go. Under such stressful circumstances, children feel less in control than they would like to be, and they self-touch to outlet nervous energy and provide physical reassurance (Knapp & Hall, 2013). Such scenarios occurred frequently throughout the observations, and did typically accompany increases in the excluded children's self-touching.

## 6. Eye Gaze

Gaze, as defined by Knapp and Hall in *Nonverbal Communication* (2013), refers to the eye movement we make in the general direction of another's face. Mutual gaze occurs when both interactants look into each other's eye area. As mentioned previously, eye gaze and body orientation are two direct reflectors of a person's focus of attention, one often accompanying the other. However, like with body orientation, it is not uncommon for interacting children not to engage in direct mutual gaze, even when standing at a distance of less than two feet. Past research on adult gaze behavior introduced the concept of a visual dominance ratio, a statistical indicator of the relative dominance exhibited by gaze in a dyadic interaction. The visual dominance ratio is calculated as the percentage of time spent looking at another while speaking to him or her divided by the percentage of time spent looking at him or her while listening to him or her (Knapp & Hall, 2013). Higher-dominance adults have been found to have higher visual dominance ratios in laboratory and naturalistic settings (Koch, Baehne, Kruse, Zimmermann, & Zumbach, 2010). Table 6 presents data on eye gaze collected from a ten-minute videotaped interaction between three boys to examine if the same is true among children.

**Table 6.** *Eye gaze and dominance in a group of children over a ten-minute interaction (FR and TSS were counted over one hour of interaction)*

Characteristics/Participants	Alex	Barry	Chase
FR/TSS	7/8	2/6	1/6
Relative dominance level	0.88	0.33	0.17
Gaze time speaking/s	111	32	15
Speaking time/s	129	39	15
Gaze time listening/s	28	105	141
Listening time/s	54	144	168
Visual dominance ratio	1.66	1.13	1.19

In the interaction observed, the visual dominance ratio was 1.66 for the dominant child Alex (RDL=0.88) and 1.13 and 1.19 for the two subordinate children Barry and Chase (RDL=0.33; RDL=0.17). We may infer that dominant children have higher visual dominance ratios than subordinate ones, meaning that

the ratio of gaze percentage when speaking to gaze percentage when listening was higher in dominant than subordinate children. While all children gazed steadily at their interactants when speaking (86%, 82%, and 100% of the time, Alex, Barry & Chase respectively), the more dominant child, Alex, gazed less while listening (52% of the time compared to 73% and 84%), which was the main cause for the differences in visual dominance ratio. Focusing his gaze on objects of interest elsewhere, the dominant child signals that his attention is also away from his speaking peer. This inattention deprives the lower-dominance child of a basic sense of interaction and respect that should accompany conversation. Looking at both sides of the dyad, the differences in gaze time may further reflect their different levels of desired intimacy. According to Argyle and Dean's (1965) intimacy equilibrium model, the dominant child gazing less when the peer gazes too much functions to lower the total intimacy of the relationship. By contrast, the lower-dominance children's tendency to maintain steady eye gaze could be seen as a continuous attempt to establish not just mutual gaze, but a social bond. Considering these significant social functions of eye gaze, it's less surprising that the subordinate children observed would at times tug, move, or even stomp, just so that mutual gaze could be established.

## 7. Vocal Cues

Vocal cues reflect personality and can strongly affect peoples' perceptions of each other. Their significance, at times, even exceeds the verbal messages themselves (Knapp & Hall, 2013). Table 7 records the relative dominance and vocal characteristics of a group of children over a videotaped one-hour interaction.

**Table 7.** *Vocal cues and dominance in a group of children over a videotaped one-hour interaction*

Characteristics / Participants	Xavier	Yolanda	Zora	Cat	Don
FR/TSS	1/7	5/7	4/5	2/10	6/7
Relative dominance level	0.14	0.71	0.80	0.20	0.86
Speaking time	Average	Long	Short	Short	Average
Volume	Low	Average	High	Average	Very high
Rate	Average	Average	High	Low	High
Pitch	High	Average	Low	Very high	Very high
Pauses	39	22	6	38	14
Clarity	Low	Average	High	Low	High
Filler words	18	19	5	25	9
Repetitions	32	21	16	30	7
Imitated phrases	19	10	0	14	6
General fluency	Low	Average	High	Low	High

Regarding volume, most dominant children observed had either high (Zora, RDL=0.8) or very high (Don, RDL=0.86) loudness when speaking, some even prone to yelling. Even the quietest dominant child spoke at an average volume (Yolanda, RDL=0.75). By contrast, subordinate children spoke either at normal loudness (Cat, RDL=0.20) or relatively quietly (Xavier, RDL=0.14). Volume, in short, correlated largely with dominance. This idea is supported by past research among adults finding dominant individuals to have louder voices (Hall, Coats, & Smith LeBeau, 2005; Weaver & Anderson, 1973). Among children, a dominant child's loudness of voice seems to have an attention-grabbing effect that asserts dominance as well as reflects it.

The general fluency of different children's speech is also revealing of dominance. In the study conducted, the more dominant children (Zora, Yolanda & Don) had 6, 22, and 9 pauses in their speech, respectively, averaging 14, fewer than subordinate children's 39 and 38 (average 38.5). They also used fewer fillers, such as "uh"s and "umm"s, averaging 11 compared to subordinate children's 21.5. Repeating oneself, as in "I was going to...I was going to go over..." also happened more frequently among subordinate children (average 16.5) than dominant ones (average 14.7). Meanwhile, the rate of speech was high in two dominant children (Zora & Don), average in another (Yolanda), and average or low in the subordinate children (Xavier & Cat). All of the above being indicators of fluency, it may be concluded that dominant children have higher speech fluency than subordinate ones. This fluency of speech seems to convey a sense of maturity and leadership, enabling the dominant children to further establish dominance. Subordinate children, on the other hand, are more prone to faltering, and some enunciate poorly. In a few cases observed, this became a source of irritation for the other children, whose patience visibly thinned (indicated by glances away and interruptions) while subordinate children composed their sentences.

There are, however, telling exceptions. Certain children of lower dominance were prone to speaking loudly as well. Contrary to the arresting effect of the same high volume utilized by higher-dominance children, they were mostly ignored, even somewhat disliked, for their loudness. This could be explained by borrowing from Dan Olweus's characterizations of a "provocative victim of bullying." Unlike the more numerous submissive victims, provocative victims are hyperactive, distracting, and have problems with concentration. Unlike dominant children, their hyperactivity is excessive and physicality often weak. The loud yet subordinate child of the playground may be an offshoot of the provocative victim, whose subordination has not amounted to victimization, but has still been somewhat excluded and made subordinate. In fact, his subordination is in part precisely because of his unique nonverbal behaviors which incompletely resemble those of dominant children, and this situation may be aggravated until he becomes a provocative victim.

Another exception exists in speech fluency. One particular dominant child observed had relatively low fluency in speech. He paused often, frequently used fillers and slightly slurred his speech. However, he behaved "dominantly" by measures of most other nonverbal messages, had a larger-than-average build suggesting dominance, nearly always succeeded in social conflicts, and his peers listened to him with focused attention as he composed his sentences. This exception has significance because it shows that although nonverbal behaviors aid in building dominance, not all the dominant modes of nonverbal communication discussed need be displayed consistently for a child to gain it among peers. A faltering of dominant behavior in one line of nonverbal communication can be

made up for by others, or by biological traits constructive of dominance. Similarly, offsetting dominant behaviors by otherwise subordinate children may still not overthrow the overall high levels of subordination. They may even contribute to further subordination, as in the case of the provocative subordinate child discussed above. Although the nonverbal behaviors characterizing dominance/subordination are individually supported and explained, holistic evaluations of children's behavioral sets are necessary before arriving at conclusions about the hierarchal status of individual children.

## 8. Discussion

Several nonverbal indicators of dominance and subordination were identified throughout the course of this research. Open and relaxed postures, more and bigger gestures, a tendency to initiate touches in play fights, louder voices, higher speech fluency and less gazing while listening are all indicative of higher dominance. Markers of subordination, by contrast, were reserved and contractive postures, more direct orientation during interactions, fewer and smaller gestures, more self-touching, more gazing than being gazed at, lower volume of speech and lower speech fluency.

In relation to the pre-existing scholarship in related disciplines, this study is unique in its proposition of a complete set of behavioral characteristics for both dominant and subordinate children involving multiple nonverbal perspectives. As a study into children of China, this work is also particularly significant, as the vast majority of past research into children's dominance and nonverbal behavior is conducted by European/American researchers on samples of European/American children, whose behavior cannot always be generalizable to China or the East-Asian world (Chiu, 1972).

However, the present study is also limited by the fact that the sample consisted of only a small number of children drawn from one particular neighborhood and culture. There were also time limitations which prevented the gathering of more comprehensive data, and more diverse examinations will need to be conducted before making broader generalizations about children's behavior.

In addition to statistics regarding nonverbal behavior and some discussions of the causes behind differences between dominance levels, an important theme throughout this paper is that nonverbal actions are often simultaneously reflections of and establishment of dominance rankings, and that dominance hierarchies self-reinforce through such nonverbal communication; sometimes, to dangerous extents.

As phrased by W.T. Boyce (2004), children's experiences with social ordering are "a first, formative encounter with the hierarchical social relations that affect health and susceptibility to violence over the human lifespan" (p.47). The behavioral sets of dominant and subordinate children described in this paper were written with the hope of serving as a preliminary framework, the use of which will hopefully lead to children who need intervention being recognized and helped, and the first encounters children have with hierarchies not forming downwards trajectories into the ever-expanding list of negative aftereffects of bullying, over-aggression, or rejection.

## References

- Ammon A.V., & Filippova E.V. (2015). The Social Status of a Child in a Group and its Correlation with Psychological, Social-Cognitive and Behavioral Characteristics. *Konsul'tativnaya Psikhologiya I Psikhoterapiya = Counseling Psychology and Psychotherapy*, 23(4), 162–175. (In Russ., abstr. in Engl.) doi: 10.17759/cpp.2015230410
- Anderson, C., et al. (2001). Who Attains Social Status? Effects of Personality and Physical Attractiveness in Social Groups. *Journal of Personality and Social Psychology*, 81(1), 116-132. doi:10.1037//0022-3514.81.1.116
- Argyle, M., & Dean, J. (1965). Eye-Contact, Distance and Affiliation. *Sociometry*, 28(3), 289-304. doi: 10.2307/2786027
- Boyce, W. T. (2004) Social Stratification, Health, and Violence in the very Young. *Annals of the New York Academy of Sciences*, 1036(1), 47-68. doi:10.1196/annals.1330.003
- Cheek, J. M., & Buss, A. H. (1981). Shyness and Sociability. *Journal of Personality and Social Psychology*, 41(2), 330-339. doi: : 10.1037/0022-3514.41.2.330
- Chiu, L. (1972). Cross-Cultural Comparison of Cognitive Styles in Chinese and American Children. *International Journal of Psychology*, 7(4), 235-242. doi: 10.1080/00207597208246604
- Fast, J. (2016). *Beyond Bullying: Breaking the Cycle of Shame, Bullying, and Violence*. Oxford University Press, New York, Oxford.
- Gage, F. H., & Lieberman, A. F. (1978). A Multivariate Analysis of Social Dominance in Children. *Aggressive Behavior*, 4(3), 219-229. doi:10.1002/1098-2337(1978)4:3<219::AID-AB2480040304>3.0.CO;2-B
- Gifford, R. (1994). A Lens-Mapping Framework for Understanding the Encoding and Decoding of Interpersonal Dispositions in Nonverbal Behavior. *Journal of Personality and Social Psychology*, 66(2), 398-412. doi: 10.1037/0022-3514.66.2.398
- Hall, J. A., & Veccia, E. M. (1990). More "Touching" Observations: New Insights on Men, Women, and Interpersonal Touch. *Journal of Personality and Social Psychology*, 59(6), 1155-1162. doi: 10.1037/0022-3514.59.6.1155
- Hall, J. A., & Knapp, M. L. (2013). *Nonverbal Communication*. De Gruyter Mouton, Berlin, Boston. doi: <https://doi.org/10.1515/9783110238150>
- Hall, J. A., Coats, E. J., & LeBeau, L. S. (2005). Nonverbal Behavior and the Vertical Dimension of Social Relations: A Meta-Analysis. *Psychological Bulletin*, 131(6), 898-924. doi: 10.1037/0033-2909.131.6.898
- Halpern, J., et al. (2015). Social Dominance, School Bullying, and Child Health: What are Our Ethical Obligations to the Very Young? *Pediatrics*, 135(Suppl 2), S24-30. doi: 10.1542/peds.2014-3549C
- Harris, S., & Petrie, G. (2002). A Study of Bullying in the Middle School. *NASSP Bulletin*, 86(633), 42-53. doi: 10.1177/019263650208663304
- Hawley, P. H., Little, T.D., & Rodkin, P.C. (2007). *Aggression and Adaptation: The Bright Side to Bad Behavior*. L. Erlbaum Associates, Mahwah, N.J. ISBN: 9780805862348

- Huesmann, L. R., et al. (1984). Stability of Aggression Over Time and Generations. *Developmental Psychology*, 20(6), 1120-1134. doi: 10.1037/0012-1649.20.6.1120
- Klomek, A. B., et al. (2008). Childhood Bullying as a Risk for Later Depression and Suicidal Ideation among Finnish Males. *Journal of Affective Disorders*, 109(1-2), 47-55. doi: 10.1016/j.jad.2007.12.226
- Klomek, A. B., et al. (2009). Childhood Bullying Behaviors as a Risk for Suicide Attempts and Completed Suicides: A Population-Based Birth Cohort Study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 48(3), 254-261. doi: 10.1097/CHI.0b013e318196b91f
- Koch, S. C., et al. (2010). Visual Dominance and Visual Egalitarianism: Individual and Group-Level Influences of Sex and Status in Group Interactions. *Journal of Nonverbal Behavior*, 34(3), 137-153. doi: 10.1007/s10919-010-0088-8
- Lv, J. (2017). Differences Between American and Chinese Concepts of Space. *Advances in Social Science Education and Humanities Research*, 156, 167-169. doi: <https://doi.org/10.2991/seiem-17.2018.41>
- Magee, J. C., & Galinsky, A.D. (2008). 8 Social Hierarchy: The Self-Reinforcing Nature of Power and Status. *The Academy of Management Annals*, 2(1), 351-398. doi: 10.1080/19416520802211628
- McDaniel, E., & Andersen, P. A. (1998). International Patterns of Interpersonal Tactile Communication: A Field Study. *Journal of Nonverbal Behavior*, 22, 59-73. doi: 10.1023/A:1022952509743
- Mehrabian, A. (1972). *Nonverbal Communication*. Aldine-Atherton, Chicago.
- Weigel, R. M. (1980). *Dominance Relations: An Ethological View of Human Conflict and Social Interaction*. (D. R. Omark, F. F. Strayer, & D.G. Freedman, Eds.). Garland Pub, New York. doi: [https://doi.org/10.1016/0162-3095\(81\)90040-6](https://doi.org/10.1016/0162-3095(81)90040-6)
- Pettit, G. S., et al. (1996). Stability and Change in Peer-Rejected Status: The Role of Child Behavior, Parenting, and Family Ecology. *Merrill-Palmer Quarterly*, 42(2), 267-294. doi: 10.1007/BF01537347
- Pianta, R. C. (1999). *Enhancing Relationships between Children and Teachers*. American Psychological Association, Washington, DC. ISBN: 1557985421
- Regan, P.C., Jerry, D., Narvaez, M., & Johnson, D. (1999). Public Displays of Affection Among Asian and Latino Heterosexual Couples. *Psychological Reports*, 84, 1201-1202. doi: 10.2466/pro.1999.84.3c.1201
- Remland, M. S., Jones, T. S., & Brinkman, H. (1991). Proxemic and Haptic Behavior in Three European Countries. *Journal of Nonverbal Behavior*, 15(4), 215-232. doi: 10.1007/BF00986923
- Savin-Williams, R. C. (1979). Dominance Hierarchies in Groups of Early Adolescents. *Child Development*, 50(4), 923-935. doi: 10.2307/1129316
- Smith, P. K., & Shu S. (2000). What Good Schools can do about Bullying: Findings from a Survey in English Schools After a Decade of Research and Action. *Childhood*, 7(2), 193-212. doi: 10.1177/0907568200007002005



- Sourander, A., et al. (2007). What is the Early Adulthood Outcome of Boys Who Bully Or are Bullied in Childhood? the Finnish 'From a Boy to a Man' Study. *Pediatrics (Evanston)*, 120(2), 397-404. doi: 10.1542/peds.2006-2704
- Strayer, F. F., & Strayer, J. (1976). An Ethological Analysis of Social Agonism and Dominance Relations among Preschool Children. *Child Development*, 47(4), 980-988, doi: <https://doi.org/10.2307/1128434>
- Teisl, M., et al. (2012). Differential Expression of Social Dominance as a Function of Age and Maltreatment Experience. *Developmental Psychology*, 48(2), 575-588. doi: 10.1037/a0024888
- Weaver, J. C., & Anderson, R. J. (1973). Voice and Personality Inter-relationships. *The Southern Speech Communication Journal*, 38(3), 262-278. doi: 10.1080/10417947309372196
- Willis, F. N., & Reeves, D. L. (1976). Touch Interactions in Junior High Students in Relation to Sex and Race. *Developmental Psychology*, 12(1), 91-92. doi: 10.1037/0012-1649.12.1.91
- Willis, F. N., & Hofmann, G. E. (1975). Development of Tactile Patterns in Relation to Age, Sex, and Race. *Developmental Psychology*, 11(6), 866-866. doi: 10.1037/0012-1649.11.6.866



# Nietzsche's *Agon* with Plato: Truth, Morality, and Modernity

---

*Yilong Wu*

**Author Background:** *Yilong Wu grew up in China and currently attends the University School of Nashville in Nashville, Tennessee in the United States. His Pioneer research concentration was in the field of philosophy and titled "The Examined Life."*

## Abstract

In this paper, I introduce an alternative interpretation of Nietzsche's philosophy and his antagonistic relationship with Plato: understanding Nietzsche's philosophy as a form of *agon*, or competition. Synthesizing the notion of *agon* with Nietzsche's critique of Plato, I demonstrate that Nietzsche is not only engaging in a philosophical *agon* with Plato but also using this *agonal* relationship to resolve the cultural crisis of modernity. Through a close examination of Nietzsche's relationship with *agon*, I conclude that Nietzsche considers *agonal* spirit as an essential element to what he calls "the philosophers of the future."

## Introduction

People tend to conceptualize Nietzsche and Plato as two extremes within the spectrum of Western philosophy. While Plato is often portrayed as a metaphysician, Nietzsche is portrayed as an anti-metaphysician; while Plato supports reason and Truth, Nietzsche ridicules the belief in reason and is skeptical toward the existence of Truth; while Plato dreams of a world of *being*, Nietzsche dreams of a world of *becoming*; while Plato insists on the objectivity of moral values, Nietzsche opposes it. Because of this tension between Plato's philosophy and Nietzsche's philosophy, people often make the assumption that Nietzsche disapproves of Plato's philosophy, thus tending to neglect or deliberately ignore the relationship between Nietzsche and Plato. In this article, however, I will introduce an alternative and less common interpretation of Nietzsche's philosophy -- understanding Nietzsche's philosophy through the lens that Nietzsche is constantly having a philosophical *agon*, or a competition, with Plato. As I will demonstrate in this essay, Nietzsche's entire philosophy is a magnificent spectacle of *agon*. Through his *agon* with Plato, Nietzsche discovers the nature and origin of philosophy, reevaluates the components of the Western philosophical tradition since Plato, and eventually sets out to resolve the cultural and intellectual crisis of modernity.

Although the literature on Nietzsche and his relationship to other philosophers proliferate, the relationship between Nietzsche and Plato is inadequate and remains unexplored. As the renowned Nietzsche scholar and philosopher Walter

Kaufmann insightfully commented on this phenomenon, "Nietzsche's high esteem for the Greeks is a commonplace; but it has been assumed that he wanted to return to the pre-Socratics, while his great debt to Socrates, Plato, Aristotle, and the Stoics has been overlooked"(Kaufmann 402). Even today, there are only a few studies on the relationship between Nietzsche and Plato. According to my own research, the only available monograph on this topic is Mark Anderson's *Plato and Nietzsche: Their Philosophical Art*, yet Anderson's account only offers a comparison between Nietzsche's philosophy and Plato's philosophy, and fails to illustrate how Nietzsche employs the rejection of Plato's philosophy to offer a new solution to the threats of modernity.

Not only is the literature about Nietzsche's philosophical debt to Plato inadequate but the literature about the role of *agon* in Nietzsche's philosophy is inadequate as well. According to my own research, the only existing monograph on Nietzsche's *agon* seems to be Yunus Tuncel's *Agon In Nietzsche*, but similar to Anderson's interpretation, Tuncel's interpretation only focuses on the influence of presocratic philosophy on Nietzsche but does not fully address the connection between *agon* and what Nietzsche calls the 'philosophy of the future.' The majority of Nietzsche commentators and scholars, in short, seem to overlook Nietzsche's deep connection to the notion of *agon*.

The contemporary scholarship, I believe, is in need of an interpretation of Nietzsche that not only examines the philosophical connection between Nietzsche and Plato but also demonstrates the role of *agon* in Nietzsche's philosophy. This is the interpretation of Nietzsche that I intend to offer in this essay: to synthesize Nietzsche's philosophical connection to Plato with the notion of *agon*, and explore how Nietzsche uses the notion of *agon* to resolve the intellectual and cultural crisis of modernity. Thus, the task of this essay is twofold: with the notion of *agon* as the central focus, this essay first examines in depth Nietzsche's philosophical relationship to Plato and then examines the connection between the notion of *agon* and Nietzsche's overcoming of the cultural crisis of modernity.

The structure of this essay consists of four parts. In the first section of this essay, I will first examine the role of *agon* in Nietzsche's reading of Ancient Greek philosophy as well as in his understanding of the origin and nature of philosophy, suggesting the importance of *agon* in Nietzsche's thoughts and to his philosophy as a whole. In the second section of this essay, I will examine the philosophical relationship between Nietzsche and Plato, arguing that Nietzsche views himself as in constant *agon* with Plato. After this, I will proceed to explore Nietzsche's solution to the cultural and intellectual crisis of modernity through two aspects: truth and ethics. In the third section of this essay, thus, I will examine Nietzsche's critique of the rationalist view of truth and explore the place of Plato in his critique, concluding that Nietzsche's criticism of Platonic truth could also be read as his criticism of Plato's lack of *agonal* spirit or courage. In the fourth section of this essay, I will examine Nietzsche's critique of Platonic and Christian moral interpretation of the world, concluding this section with a reconstruction of Nietzsche's ethics: an ethic of *agon*.

## 1. Nietzsche, *Agon*, and the Origin of Philosophy

Unlike many of his contemporaries, Nietzsche viewed and situated his philosophy in the Greek tradition. Nietzsche was already obsessed with the Greek culture during the early stage of his life.<sup>1</sup> During his early years as a classical scholar and a philosopher, for example, Nietzsche wrote numerous essays about the Greek culture, the Greek

spirit, and Homer, such as *The Birth of Tragedy*, “Homer and Classical Philology,” and “Homer’s Contest,” *Tragic Age of the Greeks*, and *The Pre-Socratic Philosophers*. Indeed, throughout his entire philosophical career, Nietzsche’s philosophy is imbued with, and deeply influenced by, the Greek way of thinking, the Greek spirit, and the Greek culture. To comprehend Nietzsche’s philosophy, we have to start with Nietzsche’s relationship with the Greeks and first explore Nietzsche’s own understanding of the nature of philosophy: philosophy as a form of *agon* [ἀγών] <sup>2</sup>

In his account of the origin of philosophy, Nietzsche locates Socrates as the starting point of philosophy and argues that Socrates transformed the previous sophist practices to the practices of philosophy. As Nietzsche writes in *The Birth of Tragedy*, Socrates is “the one turning point and vortex of so-called world history” (*BT*: 15). But why does Nietzsche consider Socrates as the “turning point” of Greek culture?<sup>3</sup> Why doesn’t Nietzsche designate Calicles or Heraclitus as the turning point? Nietzsche expounds upon his decision later in *Twilight of the Idols*; as Nietzsche explains it,

That [Socrates] discovered a new kind of *agon*, that he was the first fencing-master in it for the aristocratic circles of Athens, is one reason. He fascinated because he touched on the agonal instinct of the Hellenes – he introduced a variation into the wrestling matches among the youths and young men (*TI* I: 8).

Nietzsche suggests that this transitional stage is marked by Socrates’s discovery of, and propagation of, a new kind of *agon* in Athens.<sup>4</sup> Socrates transforms the Greek spirit of *agon* from body to mind, channeling the agonistic instincts into a philosophical plane and thus internalizing *agonal* instincts into philosophical practice (Tuncel, 240). Philosophy then, according to Nietzsche’s explanation, is also a form of *agon*, and philosophical discourse is driven by the *agonal* instincts. The nature of philosophy thus, according to Nietzsche, is a spiritual game, a contest between two different minds, and an antithetical practice. It is an ongoing activity of contest, competition, and the antithesis between two different forces.

Slightly different from the Greek *agon*, however, Nietzsche’s notion of *agon* seems to put more stress on the artistic product of contents rather than cruelty or violence. In the essay “Homer’s Contest,” Nietzsche argues that *agon* has its artistic significance, and Nietzsche illustrates that *agon* could make one “a poet, a sophist, an orator” through the case of Plato (*HC*: 32-33). Furthermore, Nietzsche’s notion of *agon* is closely resonant with the creation of art through the *Apollinian* and *Dionysian* duality. As Nietzsche explains this creation process in *The Birth of Tragedy*, “the continuous development of art is bound up with the *Apollinian* and *Dionysian* duality -- just as procreation depends on the duality of the sexes, involving perpetual strife with only periodically intervening reconciliations” (*BT*: 1). Therefore, the constant strife between Dionysus and Apollo could also be understood as a form of a contest or an *agon* between these two deities, and the final product or synthesis of this contest is art -- indeed, *agon* is essentially aesthetic. Nietzsche’s aesthetics helps illuminate that *agon* also has its artistic, creative dimension. Another underlying feature of Nietzsche’s *agon*, thus, is its artistic creativity. The essential aim of *agon* is always toward creating something new, better, and aesthetic -- it is a process of regeneration.

If the essence of *agon* is toward aesthetics and creativity, then philosophy, as a spiritualized form of *agon*, also needs to have an aesthetic, creative dimension. After establishing the notion that the synthesis of the *Dionysian* and the *Apollonian* is an

artistic creation, Nietzsche also introduces Socrates into this antithesis, juxtaposing Socrates with the tragic art. Eventually, Nietzsche suggests that Socrates's rationalistic tendency necessarily gives way to art and thus leads to the regeneration of art (Kaufmann, 394). Indeed, the creative, aesthetic dimension is the underlying rationality of Nietzsche's philosophy: Nietzsche sees philosophy as a creative, artistic enterprise; that is, Nietzsche believes that he could achieve self-creation through philosophy. The task of a philosopher, Nietzsche envisions, is not about the discovery of truth but rather about the artistic creation of concepts. The vision that philosophy as a creative, artistic enterprise permeates throughout Nietzsche's writings. Nietzsche creates concepts, aestheticizes truth, and spiritualizes philosophy. Indeed, Nietzsche is a great destroyer in the narrative of Western philosophy. He is a philosopher who philosophizes with a hammer, a great destroyer who destroys every concept, a philosophical anarchist who tries every position to test and challenge philosophy.

The *agonal* spirit and its end toward creativity characterize Nietzsche's philosophy as a whole; indeed, Nietzsche is a philosopher of creativity. Nietzsche does not only philosophize with his hammer but also philosophize with his *agonal* spirit. Nietzsche contests with other philosophers in order to create; hence, Nietzsche is constantly in *agon* with other philosophers. Throughout the entire narrative of the Western philosophical tradition, one of the greatest enemies to Nietzsche is Plato; thus, Nietzsche intends to contest with Plato and attempt to defeat Plato.

## 2. Nietzsche's *agonal* with Plato

The prevailing view of the relationship between Plato and Nietzsche is that Nietzsche vehemently rejects Plato's philosophy. Throughout his writings, Nietzsche frequently expresses his anti-Platonic view; for instance, Nietzsche describes Plato as "the symptom of decay pseudo-Greek, as the agents of the dissolution of Greece, as pseudo-Greek, as anti-Greek" (*TI I*: 2). Indeed, Nietzsche does have an anti-Platonic view in his philosophy. However, Nietzsche's attitude toward Plato is never single-faceted. That is, Nietzsche expresses not only his anti-Platonic view but also his admiration for Plato; for instance, while Nietzsche is often hostile to Plato and Plato's philosophy, he also highly praises the "Platonic dialogue" and criticizes that modern philosophy has deviated from the Greek way of philosophizing (*D*: 544). Occasionally, Nietzsche's remark of Plato even seems to contradict his other remarks on Plato. Furthermore, Nietzsche's relationship with Plato is complicated by his *agonal* spirit, as I have illustrated in the previous section. Nietzsche's complex and ambivalent view toward Plato, thus, obstructs us from understanding Nietzsche's relationship to Plato only through his remarks on Plato; that is, we cannot truly understand Nietzsche's ambivalent view toward Plato only through his remarks on Plato. Instead, we have to understand and interpret Nietzsche's remarks on Plato within Nietzsche's philosophical framework as a whole.

To understand Nietzsche's relationship with Plato and untangle the ambiguity of Nietzsche's view toward Plato, we have to situate his polemics against Plato within a larger framework -- namely, we have to locate and understand Nietzsche's ambiguous view toward Plato within his philosophical system as a whole. One of the most salient features of Nietzsche's philosophy is his *agonal* spirit: the spirit of a warrior. In *Ecce Homo*, Nietzsche expresses his spirit as a warrior, claiming that "I am warlike [*kriegerisch*] by nature. Attacking is one of my instincts" (*EH* 1: 7).<sup>5</sup> Indeed, Nietzsche is always a warrior within the narrative of the Western philosophical tradition, and this warlike nature constantly characterizes Nietzsche's style and his philosophical system; he is not only a philosopher but also a warrior.

Nietzsche's negative view toward Plato, therefore, is less his disdain of Plato's philosophy than a result of his warlike nature and instincts. Nietzsche views Plato as his opponent, and Nietzsche is attempting to constantly engage in a contest or *agon* with Plato (Zuckert, 22).

Furthermore, Nietzsche's attack on Plato could even be understood as his homage and admiration for Plato, for Nietzsche believes that the level of his criticism should be proportionate to the level of the worth of the philosopher. As Nietzsche explicates this view,

The strength of those who attack can be measured in a way by the opposition they require: every growth is indicated by the search for a mighty opponent—or problem; for a warlike philosopher challenges problems, too, to single combat. The task is *not* simply to master what happens to resist, but what requires us to stake all our strength, suppleness, and fighting skill -- opponents that are our equals (*EH* 1: 7).

Thus, Nietzsche's hostile attack on Plato and his unique style of *ad hominem* should also be understood as a form of admiration for Plato: indeed, Nietzsche views Plato as a philosopher worthy of his criticism and challenge. As Nietzsche states in one of his letters, "Perhaps this old Plato is my true great opponent? But how proud I am to have such an *opponent*" (Brobjer 236). In this way, Nietzsche's hostility toward Plato could even be interpreted as his 'love' toward Plato. Instead of simply viewing the relationship between Plato and Nietzsche as an antagonistic one, therefore, we should rather understand this antagonistic relationship as Nietzsche's way to admire and love Plato.

However, Nietzsche's disapproval of Plato's philosophy is not only because of his 'love' and admiration for Plato but also because of his creative nature. As I have suggested in Section 1, Nietzsche is indeed a philosopher of *agon* and creativity, so one of the most important themes in Nietzsche's philosophy is creativity -- the creation of new concepts and the production of excellence. As Nietzsche writes it, "the men of great creativity, the really great men according to my understanding" (*WP*: 957). Creativity, thus, is the paradigm and the essence of greatness to Nietzsche. Emphasizing the importance of this creative force, Nietzsche practices his teaching of creativity in his philosophical work. To Nietzsche, philosophers have always followed the same path and gloried the old idols. Indeed, the Western philosophical tradition has long been characterized by the pursuit of Truth, the praise of Platonism and rationalism, and dogmatism rather than the enterprise of contesting with other philosophers. According to Nietzsche's understanding of philosophy, the Western tradition of philosophy thus lacked creativity and *agonal* spirit-- *the essence of greatness* -- because it is creativity and the *agonal* spirit that define the greatness of philosophy and lead philosophy toward progress. The modern philosophers, as Nietzsche would argue, are afraid of contesting with others and thus submit themselves to the reign of rationalism, dogmatism, and Truth.

Nietzsche's motive to contest with Plato, however, is more complicated than simply redeeming the greatness of philosophy, as Nietzsche is often concerned deeply with the cultural crisis of modernity. The crisis of modernity, as Nietzsche suggests, is nihilism and the death of God. During the nineteenth century, the advent of the Enlightenment spirit and scientific rationality had threatened the existence of God. For example, David Hume challenged the existence of God and miracles, claiming that "there was never a miraculous event established on so full an evidence" (*EHU*

10.2/14). The existence of God once established a universal standard for the world, yet "the death of God" brought "meaninglessness of events" and nihilism into the world (*Nietzsche: Life as Literature*, 71). The death of God was also a deep concern for Nietzsche. As Kaufmann suggests, "to have lost God means madness; and when mankind will discover that it has lost god, universal madness will break out. This apocalyptic sense of dreadful things to come hangs over Nietzsche's thinking like a thundercloud" (Kaufmann, 97). To Nietzsche and his contemporaries, without the existence of God, the world is devoid of value, meaning, and truth; everything falls apart. What should we do then? This seems to be Nietzsche's real question when he writes "God is dead" (*GS*: 125). Nietzsche's philosophy, thus, is also situated at this chaotic vortex of modernity and nihilism.

In addition, tension also arises when we see Nietzsche simultaneously commit to the ancient Greek culture and to solving the threat of modernity. What is Nietzsche's philosophical position? Does Nietzsche view himself and his philosophy as being closer to modern or ancient? Why does Nietzsche engage with Plato if his main concern is to solve the problem of nihilism? Douglas Kellner explains the tension between Nietzsche's project of modernity and his *agon* with Plato:

Despite [Nietzsche's] keen appreciation for past cultures like classical antiquity and defense of some pre-modern values, Nietzsche is very future and present-oriented, attacking tradition while calling for a new society and culture. An impetus toward innovation, involving negation of the old and creation of the new, is therefore at the very heart of Nietzsche's complex and often enigmatic theoretical work, which, in the spirit of modernity, affirms development and transcendence of the old as crucial values for contemporary individuals and society (Kellner, unpaginated)

Thus, the purpose of Nietzsche's *agon* with Plato is twofold: on the one hand, Nietzsche intends to contest with Plato to redeem the creative and *agonal* aspect of philosophy, and on the other hand, Nietzsche intends to contest with the old Plato to create new values for the future. That is, to overcome the problem of nihilism, Nietzsche starts with the cause of modern nihilism: Plato. By defeating Plato and establishing philosophy upon this cause, Nietzsche intends to find another path to our life, a path that is not doomed to nihilism. With his calling for the philosophy of the future, Nietzsche set out to contest with Plato.

### 3. Nietzsche, Plato, and Truth

As the diagnostician of philosophy, Nietzsche located the starting point of the illness of Western philosophy at Plato. As Nietzsche claims in the preface to *Twilight of Idols*, he not only intended to revalue all values but also to sound out the "eternal idols" of the ages (*TI* Pref.). For most of his philosophy, Nietzsche intended to sound out the *eternal* idol Plato, as Plato is the founder and starting point of Western rationalism, dogmatism, nihilism, and the *decadent* pursuit of Truth.<sup>6</sup> Thus, Nietzsche's first challenge to Plato is Plato's theory of Truth; by challenging Plato's theory of Truth through his method of "revaluation of all values," Nietzsche also indirectly challenges one of the fundamental practices in the entire Western philosophical tradition: the pursuit of Truth. Furthermore, through this challenge toward Plato's theory of Truth, as well as the entire Western tradition, Nietzsche also

constructed his own theory of truth and sought to find a solution to the problem of nihilism.

An understanding of Nietzsche's attitude toward truth is essential to understand the reasons for Nietzsche to challenge Plato and the traditional practice of the pursuit of Truth in Western philosophy. It is important to note that Nietzsche's project is not to reject the existence of truth, as this would lead to self-referential inconsistencies. Nor is Nietzsche's project to simply find truth in the world of *becoming*, proposing a truth of becoming or that truth is a perspective; although this is part of Nietzsche's project, this understanding of Nietzsche overlooks the entirety of Nietzsche's philosophical project. The main reason that Nietzsche rejects the pursuit of truth and the idealist conception of truth lies in his belief that truth should never be independent of the life we live and the world we inhabit (Finnigan, 2). Indeed, Nietzsche's project of truth is centered on human experience and human life rather than on the idealistic version of truth as transcendental and outside the world we inhabit; it is a truth that is approachable and intelligible to every human, not only limited to philosophers. Thus, Nietzsche's conception of truth as *becoming* and as perspective should be understood as a return to life, experience, and the world we inhabit; for Nietzsche, truth as becoming is a solution to nihilism and approach to the world that does not condemn us into nihilism. To propose his theory as a better alternative for the world, thus, Nietzsche starts his intense criticism of, and challenge to, Plato and other philosophers' pursuit of Truth.

One of Nietzsche's challenges to Plato's theory of Truth is Plato's method of *elenchus*.<sup>7</sup> In many of his dialogues, Plato portrays the image that Socrates employs *elenchus* with his interlocutors to attempt to examine and thus find Truth. As Socrates famously claims in *Theaetetus*, "And the most important thing about my art [of midwifery] is the ability to apply all possible tests to the offspring, to determine whether the young mind is being delivered of a phantom, that is an error, or a fertile truth" (*Tht.* 150 c). This passage seems to capture the essential purpose of Socratic dialogue or the method of *elenchus*: to find out Truth through testing various definitions. Nietzsche, however, is skeptical of the method of *elenchus*. Nietzsche believes that truth is essentially contingent and thus the linguistic invention is unreliable. As Nietzsche explains his view,

For now that is fixed which henceforth shall be "truth"; that is, a regularly valid and obligatory designation of things is invented, and this linguistic legislation also furnishes the *first laws of truth*: for it is here that the contrast between truth and lie first originates. The liar uses the valid designations, the words, to make the unreal appear as real; he says, for example, "I am rich," when the word "poor" would be the correct designation of his situation. He abuses the fixed conventions by arbitrary changes or even by reversals of the names (TL 45).<sup>8</sup>

Indeed, Nietzsche's concern with the relationship between language and truth is also resonant with what Ferdinand de Saussure called the "arbitrariness of signs," the fact that the sound of a word or the designation of meaning to a word is entirely arbitrary (Clark, 66). Truth, according to Nietzsche's theory, is totally contingent and arbitrary. If, as Socrates himself claimed, his method to determine truth is essentially to "make other people's things [e.g., concepts or definitions]," then Nietzsche's theory of the arbitrariness of language seems to directly deny the possibility of Socrates's method (*Eut.* 11d). That is, Nietzsche's theory seems to render Socrates's quest of truth



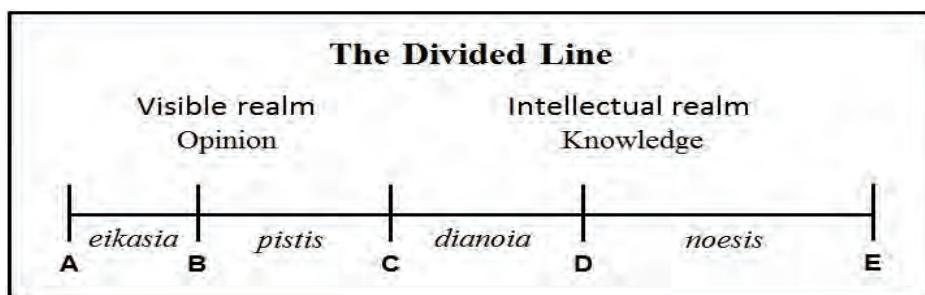
through the examination of definitions, linguistic structure, and dialectics meaningless and futile. In this way, Nietzsche objects to the method of *elenchus* by suggesting that it is futile and absurd, for this method intends to find truth through these arbitrary signs.

With his rejection of the validity of the Socratic/Platonic inquiry for Truth, Nietzsche steps further to suggest that truth is only *becoming* not *being*. While Plato holds that truth is eternal and exists in a state of *being*, Nietzsche argues the opposite of Plato's theory of truth: Nietzsche believes that truth should be in a state of *becoming*. It is constantly changing rather than staying the same permanently. As Nietzsche explicates this view, "But everything has evolved; there are no eternal facts, as there are likewise no absolute truths" (*HH* 1:1). Nietzsche, however, does not only criticize the Platonic misconceptions of truth but also pinpoints that this misconception led them to a history of error, to a state of nihilism. As Nietzsche explains,

Change, mutation, becoming in general were formerly taken as proof of appearance, as a sign of the presence of something which led us astray. Today, on the contrary, we see ourselves as it were entangled in error, *necessitated* to error, to precisely the extent that our prejudice in favour of reason compels us to posit unity, identity, duration, substance, cause, materiality, being; however sure we may be, on the basis of a strict reckoning, that error is to be found here (*TI* II: 5).

The problem of modern philosophy, Nietzsche points out, is its false assumption that truth is in a state of unity, being, and stasis. This false assumption is what leads us into a dilemma -- the state of nihilism. Nietzsche seems to implicitly suggest that we could overcome this problem of modern society by recognizing truth as becoming.

Nietzsche, however, does not only deny the method of *elenchus* in pursuing truth but also doubts the possibility of Plato's theory of Forms. According to Plato's theory of Forms, there exists two different realms for human understanding: the intelligible realm and the visible realm (*R.* 509 d). The visible realm consists of images (*eikasia*) and belief (*pistis*), while the intelligible realm consists of thought (*dianoia*) and understanding (*noesis*); according to Plato's theory, truth exists in the intelligible realm, and knowledge comes from the intelligent realm. To put it simply, the apparent/physical world could never be the source of knowledge. Plato's theory could be visually represented according to the illustration in Figure 1:



**Figure 1:** The Divided Line in Plato's *Republic* Adopted from: <http://www.john-uebersax.com/plato/plato1.htm>

Nietzsche, however, denies Plato's division of the world and advocates us abolishing the real world. As Nietzsche criticizes Plato's division of the world,

The world which matters to us is false, i.e., not a fact but a fictional elaboration and filling out of a meager store of observations; it is 'in flux', as something becoming, as a constantly shifting falsity that never gets anywhere nearer to truth for – there is no 'truth' (WLN 2 [108]).

Nietzsche suggests that since truth is only in a state of becoming, it is impossible for us to find it in the world of *being* or Plato's fictional *noesis*. Through attacking Plato's *noesis* as fictional, Nietzsche discredits Plato's division of the world: "[t]he 'real world' -- an idea no longer of any use, not even a duty any longer -- an idea grown useless, superfluous, consequently, a refuted idea: let us abolish it!" (TI II). By abolishing the real world, Nietzsche seems to return the world back to its original form, as we could not have the "apparent world" without its opposite, the "real world." What is left to us here is only this apparent world or apparent reality.

Nietzsche not only rejects Plato's divided line but also aims to criticize Plato's realism, Plato's theory of Truth, and the possibility of deriving universals from particulars. As Nietzsche argues it,

No leaf every wholly equals another, and the concept "leaf" is formed through an arbitrary abstraction from these individual differences, through forgetting the distinctions; and now it gives rise to the idea that in nature there might be something besides the leaves which would be "leaf" -- some kind of original form after which all leaves have been woven, marked, copied, colored, curled, and painted, but by unskilled, hands, so that no copy turned out to be a correct, reliable, and faithful image of the original form (TL 46).

In this passage, Nietzsche offers a nominalist criticism of Plato's theory of Forms and the Aristotelian syllogism, suggesting that it is impossible to abstract particulars into universals due to the existing individual differences among these objects. While Plato's realist view maintains the existence of universal and abstract objects, Nietzsche's nominalist position directly rejects this realist view, as we could not derive universals from particulars; thus, Nietzsche's nominalism suggests that it is impossible to discover the Forms from the intelligible realm. Another more serious challenge posed by Nietzsche's nominalism is that it exposes the fatal flaws in the Aristotelian mode of syllogism, which also requires the deduction of universals from particulars. As Aristotle suggests this distinction between universals and particulars, "What is universal is furthest from perception, and particulars are closest to it" (PA 72 a). Nietzsche's nominalism suggests that it is impossible for us to acquire things that are "prior by nature," as every particular is essentially different and we *mistakenly* equate something *unequal* as *equal* when trying to find truth or the universal. If we could not derive universals from particulars, then we could not make any claim about truth; Nietzsche, thus, gives a fatal blow to Plato and Platonic rationalism.

Yet, Nietzsche's contention with Plato is deeper than the mere difference in their conceptions of truth; Nietzsche is entirely skeptical of the philosophers' commitment to the pursuit of truth and often psychologizes this pursuit as the "will to truth." As Nietzsche writes in the beginning of *Beyond Good and Evil*,

The will to truth which will tempt us to many a venture, that famous truthfulness of which all philosophers so far have spoken with respect – which question has this will to truth not laid before us! [...] *Who* is it really that puts questions to us here? What in us really wants “truth”? Suppose we want truth: *Why not rather* untruth? And uncertainty? Even ignorance? (*BGE*: 1)

In this passage, Nietzsche first suggests the philosophers' commitment to truth as the will to truth and proposes a question for us: Why do philosophers pursue truth rather than untruth? Indeed, philosophers have pursued truth for a long time, but they never ruminate on or question the *value* of this pursuit, and Nietzsche is starting this questioning here. Psychologizing this mysterious commitment to truth as a desire, Nietzsche directly challenges the traditional *value* of the pursuit of truth started by Plato. Exposing the will to truth, Nietzsche argues that philosophers are more interested in the sense of certainty and security offered by truth (Harris, 249-250). As long as truth is eternal, philosophers could have a firm purchase on it, thereby gaining the sense of certainty and security.<sup>9</sup> However, as Nietzsche suggests, there might be no eternal truth at all, and the *real* truth is that truth is always in a state of *becoming*. We could never truly grasp or hold truth, for it is always becoming and thus elusive to us. With this hypothesis in mind, Nietzsche suggests that philosophers have deceived themselves and eluded the *real* truth through the pursuit of truth.

With the establishment of the concept of the “will to truth,” Nietzsche advances his skepticism toward the philosophical pursuit of truth. If the pursuit of truth, as Nietzsche suggests, is a will, then truth is not purely the Form from the intelligible realm but rather a product of the philosophers' will. Then what is the so-called truth? In the book *Will to Power*, Nietzsche offers us an answer to this question:

Will to truth is a making firm, a making true and durable, an abolition of the false character of things, a reinterpretation of it into beings. "Truth" is therefore not something there, that might be found or discovered -- but something that must be created and that gives a name to a process, or rather to a will to overcome that has in itself no end -- introducing truth, as a *processus in infinitum*, an active determining -- not a becoming -- conscious of something that is in itself firm and determined. It is a word for the “will to power (*WP*: 552).

According to Nietzsche, truth is not a thing that we found or discovered; Nietzsche seems to suggest that throughout the entire narrative of philosophy, philosophers since Plato never discovered truth -- the thing from the intelligible realm or *noesis* -- but *created* truth instead. Truth, in short, is only our will to power. Furthermore, Nietzsche's designation of truth as the will to power also renders the pursuit of Truth after Plato as a genetic fallacy, as the pursuit of Truth and *noesis* seems to be Plato's own psychological invention: it is only Plato's will to truth and thus is devoid of justifications for the intrinsic merit of this pursuit. Plato's *noesis* and philosophical Truth never reveal the truth of the world; they only determine what is truth in the world, thereby transforming the world of *becoming* into the world of *being*: an ultimate manifestation of the will to power.

Nietzsche's questioning of the value of truth and his criticism of philosophers' self-deception can be traced back to Plato's philosophy of virtues; the most important virtue of all Plato's virtues to Nietzsche, I believe, is courage, as Nietzsche's courage is imbued throughout his writings and also manifested in his *agonal* spirit. In *Ecce Homo*, Nietzsche writes, "How much truth can a spirit bear? How much truth can a spirit dare" (*EH F 3*). Every philosopher and spirit need the capacity to accept the possibility that there is no eternal truth at all, and they have to directly confront uncertainty and insecurity; they need the courage to do so. Similarly, Plato also considers courage as one of the most salient characteristics of a philosopher. As Plato suggests in his early dialogue, Socrates's cardinal virtue is his *parrhesia* and the courage to confront death. Furthermore, as Plato's allegory of the cave also teaches us, people need the courage to come out of the cave and confront truth and then bring truth back to the cave. However, Plato's conception of Truth seems to deviate from his constant praise for the philosopher's courage; Plato contradicts himself, as he still needs the Forms and eternal truth to live and thus does not have courage to confront the world of *becoming*. Nietzsche jeers at Plato's lack of courage, writing that "*Courage* in face of reality ultimately distinguishes such natures as Thucydides and Plato: Plato is the coward in face of reality" (*TI IX: 2*).

#### 4. *Agon*, Authenticity, and Nietzsche's Ethics

Although Nietzsche strongly criticizes the objectivity of ethics as well as many other ethical theories, Nietzsche, in fact, has a virtue ethics theory in his philosophy. Contrary to the prevailing act-oriented model of ethical theories during his time, Nietzsche's ethical theory is person-oriented. Nietzsche's ethics focus on the development of individuals, and thus the underlying value of Nietzsche's virtue ethics, as I have implied in the last section, is *authenticity*, *truthfulness*, or *honest integrity* of individuals; indeed, one of Nietzsche's main philosophical projects is to overcome nihilism, providing a philosophy of life that could address the crisis of modernity. Nietzsche believes that his ethics -- the ethics of individual authenticity -- could overcome the cultural crisis of nihilism. In this section, I will demonstrate Nietzsche's ethics of authenticity and show the connection between Nietzsche's ethics and his notion of *agon*. But first, I will provide the historical background to Nietzsche's ethical project and layout of Nietzsche's main criticisms of Plato's ethics and normative ethics in general, as this is essential to the development of Nietzsche's own philosophy and helpful to the understanding of Nietzsche's ethics.

Historically and philosophically, Nietzsche's goal behind his ethical project is twofold: to refute nihilism and to redeem the fading glory of the nineteenth-century culture. As I have previously mentioned, modern nihilism was deeply rooted in the Platonic view of the world; Nietzsche thus engaged in an *agon* with Plato and Platonism in order to uproot modern nihilism. On the other hand, however, Nietzsche's ethical project also seems to undo the rise of bourgeois, socialist, and democratic culture. Frequently criticizing the "lifelessness, fragmentation, and conformity of modern life," Nietzsche's ethical project aims at building a culture that allows the strongest creative spirit to flourish (Kellner, unpaginated). Nietzsche's ethical project, thus, is not only toward eradicating a certain type of philosophical attitude but also toward undoing certain cultural phenomena (Solomon, 120).

One of Nietzsche's first criticisms of Plato's ethics is Plato's eudaimonism. Plato eudaimonism approximates the best state of life by taking all things considered; thus, Plato's eudaimonism suggests that it is possible to assign an objective value or offer an objective assessment to our state of life. Contrary to Plato, Nietzsche argues

that values are essentially the will to power and contain no objective facts, so every evaluation is an interpretation and could not offer an objective assessment (Reginster 58-59). Every fact and value, in Nietzsche's own words, is only an "interpretation" (*WP*: 481). According to Nietzsche, Plato's eudaimonism is also a will to power and thus fails its descriptive objectivism as an overall assessment of life. Nietzsche explicates this view in *Twilight of Idols*. As Nietzsche explains it,

Judgments, values judgments concerning life, for or against, can in the last resort never be true: they possess value only as symptoms, they come into consideration only as symptoms -- in themselves such judgments are stupidities. One must reach out and try to grasp this astonishing *finesse*, that the value of life cannot be estimated [*dass der Werth des Lebens nicht abgeschätzt werden kann*] (*TI* I: 2).

In this passage, Nietzsche does not only suggest Platonic judgments regarding as flawed but also reveal one of the key themes of his philosophy: *the value of life*. Problematizing the judgments about the value of life as one's physio-psychological symptoms, Nietzsche suggests that valuations about life are only for those who are weary of life and so place a value on life (Elgat 10-11). Thus, an individual who attempts to evaluate her life or lacks the fortitude to confront and embrace her life. In this way, Plato's eudaimonism not only fails its descriptive objectivism but also fails to confront the reality of life and embrace it.

Plato's flawed valuation of life -- the manifestation of his psychological weakness and also a failure to embrace life -- allows Nietzsche to advance his critique of Plato's eudaimonism: Plato's ethics results in the negation of life. Criticizing the Platonic vision of life, Nietzsche writes:

Nothing actual has escaped from [philosophers'] hands alive. [...] Death, change, age as well as procreation and growth, are for them objections -- refutations even. What is, does not *become*; what becomes, is not...Now they all believe, even to the point of despair, in that which is. But since they cannot get hold of it, they look for reasons why it is being withheld from them. 'It must be an illusion, a deception which prevents us from perceiving that which is: where is the deceiver to be found?' -- 'We've got it, they cry in delight, 'it is the senses! These senses, *which are so immoral as well*, it is they which deceive us about the *real* world (*TI* III: 1).

To Nietzsche, Plato's condemnation of the world of *becoming* leads us to nihilistic despair, as everything we discovered is transient and becoming; according to Plato and his fellow philosophers, the world is essentially illusionary and deceptive, so there is nothing to cherish in this world of becoming. The Platonic vision of life and the subsequent theories derived from Platonism, such as Christianity, in turn, assumed that our life in this world is meaningless and led us into the abyss of nihilism (Reginster 49). What we should do, according to these philosophers, is to *negate* and transcend our life in this world for the sake of life in *another* world, in the world of *being*. The 'meaning' of life, thus, lies in negating and denying our life in *this* world - - the meaning of our life, paradoxically, is its *meaninglessness* and the hatred toward our life. Indeed, Plato's insistence on, and pursuit of, the fictional world of *being* directly causes the beginning of nihilism and leads people to hate their life and negate the value of life: "*That the highest values devalue themselves*" (*WP* I: 2).

To overcome Plato's nihilistic version of life and offer a direction to our life, Nietzsche proposes authenticity as the underlying pursuit of life.<sup>10</sup> Throughout Nietzsche's writings, Nietzsche views his philosophy as essentially a therapy to modern maladies, and this therapeutic role is centered on his promotion and protection of individual authenticity (Ansell-Pearson, 1). Authenticity, according to Nietzsche's definition, is to be *truthful* and *honest* with oneself, to be truthful to truth, and to display their characters. In this way, Plato and the pursuers of the world of *being* are inauthentic to themselves, for they negate the values and meanings of their life; they are nihilistic to the world, to themselves, and to their lives. They blind themselves through the pursuit of the world of *being*. As Nietzsche criticizes it, "We are unknown to ourselves, we men of knowledge -- and with good reason. We have never sought ourselves -- how could it happen that we should ever find ourselves" (*GM*, Pref. 1). The insistence on the world of *becoming* leads us foreign to ourselves; we are ourselves indeed, but we do not have the knowledge of ourselves, as we are constantly preoccupied with the contemplation of the Forms and the pursuit of the world of *being*. Indeed, it is the greatest irony that philosophers, as the lovers of wisdom, do not even or incline to have the knowledge of themselves.

To Nietzsche, furthermore, morality as the negation of life also causes us to be inauthentic to ourselves due to its denial of history. Morality, according to Nietzsche, lacks the historical sense of individuals. What is "moral," according to morality, is to "escape from sense-deception, from becoming, from history, from falsehood -- history is nothing but belief in the senses, belief in falsehood" (*TI* II: 1). The philosophers reject history, the past, and the experience. They are, once again, inauthentic to themselves, as they attempt to ignore one of the necessary constituents of life -- history. Against the philosopher's rejection of history, Nietzsche praises highly of the use of history and historical sense, as the embrace of history leads us to authenticity. As Nietzsche writes, "As men of the "historical sense" we also have our virtues; that cannot be denied: we are unpretentious, selfless, modest, courageous, full of self-overcoming, full of devotion, very grateful, very patient, very accommodating" (*BGE*: 224). Our embrace and consciousness of the "historical sense" allows us to know ourselves and understand our life as a continuity of events. Our capacity to be the agents of our history, Nietzsche argues, makes us something to be valued, something valuable (Ansell-Pearson, 37).

Morality does not only lead us to be inauthentic but also suppress the growth of characters and the physiological needs of individuals. In *Beyond Good and Evil*, Nietzsche explains this view,

The power of moral prejudices has penetrated deeply into the most spiritual world [...] If, however, a person should regard even the affects of hatred, envy, covetousness, and the lust to rule as conditions of life, as factors which, fundamentally and essentially, must be present in the general economy of life (and must, therefore, be further enhanced if life is to be enhanced) -- he will suffer from such a view of things as from seasickness (*BGE*: 23).

The power of moral prejudices has deeply controlled our conduct and limited the growth of life; if a person's necessary conditions of life are those characters opposed by moral prejudices, then this person is doomed to negate himself and reject his own style of living and suppress his own physio-psychology. Thus, moral prejudices in turn lead this person into a state of inauthenticity, as this person will not be truthful and honest to his own physiology after the acceptance of moral prejudices. However,

the problem of morality is more than the suppression of the physio-psychological needs of individuals; it also suppresses the growth and development of our characters. Suggesting "morality as a problem," Nietzsche writes, "The lack of personality always takes its revenge: A weakened, thin, extinguished personality that denies itself is no longer fit for anything good -- least of all for philosophy" (GS: 345). Morality, in this way, leads us to submission and docility. It renders us obedient, so we can not develop our own personality and to be fully ourselves. When an individual lacks her own unique personality, she ceases to be an individual and loses her authenticity, as she is not different from the others. To remedy the problems caused by Christian morality and the legalistic model of ethical theories, Nietzsche put the emphasis of his ethics on the development of individuals rather than the motivational or consequential nature of actions, so the Nietzschean ethics concentrates on personality and character (Brobjer, 66). Nietzsche's ethics is an *ethics of character*, in short.

However, Nietzsche's ethics seem to face a challenge posed by his theory of truth, as Nietzsche argues that every philosophy is will to power and denies the universality of principles. Does Nietzsche's virtue ethics contradict his theory of truth? The answer is no. Nietzsche never intends to construct a legalistic model of an ethical system, nor does Nietzsche intend to establish *telos* or terminus in his virtue ethics; instead, the most appropriate interpretation for Nietzsche's ethics is that Nietzsche has a contextualist model of virtue ethics. The theory of ethical contextualism is summarized by Robert Holmes:

Contextualism emphasizes the immediate situation in which judgments are made, as does particularism. And like legalism, it gives a role to rules and principles. But whereas legalism and particularism are theories about actual rightness, contextualism is a theory about *actionable* rightness. It contends that you simply cannot know what is actually right much of the time, so you must make the best judgments you can from a perspective of limited knowledge. And this means judging each situation on its merits (Holmes, 226).

Applying Holmes's description to Nietzsche's theory of virtue ethics, we could find that Nietzsche's virtue ethics is inherent in contextualism. As Nietzsche's critique of Plato's eudaimonism and Christianity shows, Nietzsche is often critical of the prescriptive aspect of normative ethics, as this might suppress the growth of an individual's character or personality; while Nietzsche's ethics lack the prescriptive nature of the traditional virtue ethics, Nietzsche's ethics nevertheless contains a value system. The essential foundation of Nietzsche's value system is authenticity or an individual's individuality. It is the *self* that Nietzsche emphasizes here. Therefore, the aim of, and the meaning of, Nietzsche's virtue ethics is toward the improvement of individuals and human beings. What Nietzsche truly intends in his virtue ethics is the attainment of an individual authenticity: to find a way to allow the individual to be herself. What matters to Nietzsche is the *actionable* rightness. The individual acts rightly as long as she pursues her authenticity and tries to be truthful to herself. In short, it is a private virtue ethics rather than a public virtue ethics.

Through the contextualist emphasis on the characters and development of *each individual*, Nietzsche develops the theory of master-slave morality to counter the degeneration of society caused by nihilism, the bourgeois culture, and Christian morality. As I have previously suggested, the mediocre banality of modern culture

and legalistic ethics suppress the characters and potentials of individuals by treating everyone through the same rules and values; what Nietzsche intends to do is to expose the problems caused by the culture of modernity and undo the 'slave morality.' As Nietzsche illustrates his view:

[Master morality] is different with the second type of morality, *slave morality*. Suppose the violated, oppressed, suffering, unfree, who are uncertain of themselves and weary, moralize: what will their moral valuations have in common? Probably, a pessimistic suspicion about the whole condition of man along with his condition. The slave's eye is not favorable to the virtues of the powerful: he is skeptical and suspicious, subtly suspicious, of all the good that is honored there -- he would like to persuade himself that even their happiness is not genuine (BGE: 260).

When the 'slaves,' or weak human beings, come to power, they replace the master morality with their own invention: a morality that suppresses the necessary conditions of life -- the will to power and excellence. They treat everyone -- both the strong and the weak -- as equals, universalizing the moral codes. By legislation such moral codes, however, they limit the potentials, power, and characters of the strong. The consequence of such a morality, as Nietzsche suggests, is "a denial of life, a principle of disintegration and decay" (BGE: 259). Slave morality does not lead to the improvement of humankind or the excellence of society; it only limits the potentials within noble individuals and makes the noble human beings ashamed of their excellence, leading to the mediocre bourgeois culture. The master morality, on the other hand, leads to the enhancement of individuals and the flourishing of characters, as, in such a society, the noble is the one that everyone craves to be and to compete with the noble human beings. What an excellent society aims at is "the production of individual great men, [...], the individual higher exemplar, the more uncommon, more powerful, more complex, more fruitful" (UT 162). What Nietzsche's ethics aims at, thus, is to reverse this phenomenon, to undo this social atavism, and to establish a society in which individuals truly flourish. Indeed, similar to Aristotle's virtue ethics, Nietzsche's ethics emphasizes on actualizing the potential within each individual and developing individual excellence or *arete*.

Nietzsche's ethics, as master morality, allows individuals not only to actualize their potentials but also to retain individual authenticity. Christian morality treats every individual through the same principle and thus suppresses the excellent characters of individuals. The problem behind this morality is that it does not allow every individual to actualize their potentials and thus does not allow individuals to be truthful to their characters; consequently, it prevents excellent individuals from being *authentic* and truthful to their characters and themselves. Unlike Christian morality, however, Nietzsche's ethics does not suppress and limit the growth of characters or force one to negate one's life. It is not a life-negating ethics but a life-affirming ethics. It encourages individuals to actualize their potentials, to achieve personal excellence, and to affirm their characters. Thus, authenticity is achieved through the Nietzschean affirmation of life.

Indeed, Nietzsche's ethics of character is essentially an ethic of affirmation and personal excellence. The underlying pursuit of Nietzsche's ethics, as I have already mentioned, is authenticity and an individual's character. A deeper implication in Nietzsche's ethics of characters, however, is the constant self-creation of individuals. As Nietzsche illustrates this view,



We, however, want to *become who we are*— human beings who are new, unique, incomparable, who give themselves laws, who create themselves. To that end we must become the best learners and discoverers of everything that is lawful and necessary in the world: we must become physicists in order to be able to be creators in this sense -- while hitherto all valuations and ideals have been based on the ignorance of physics or were constructed so as to contradict it (GS: 335).

Here, Nietzsche emphasizes that the process of self-creation requires us to create new values and laws; the essence of self-creation, thus, lies in creating new values for oneself and affirming these values. It is through creating values and laws that we "*become who we are*" rather than adhering to certain legalist ethics or certain moral codes.<sup>11</sup> By creating new values and laws, we also follow our true conscience and attain our individual authenticity: "*What does your conscience say?* – 'You shall become the person you are'" (GS: 270). In this *becoming* of the self, thus, we are also being truthful, honest, and authentic to ourselves. This *becoming of the self* is what Nietzsche emphasizes in his philosophy and seems to be Nietzsche's solution to the problem of nihilism. Indeed, God is dead, and we have killed him, so we have to replace the role of God in the world: to create our own values, to discover our own truth, and to find our own *ethos*.<sup>12</sup>

Indeed, Nietzsche's ethics of characters essentially is an ethic unique to each individual; it fits every individual's potential properly. However, Nietzsche's virtue ethics is far more complicated than the preservation of individual authenticity or the flourishing of individual excellence. Nietzsche's ethics is not only an ethic of authenticity and excellence but also an ethic of *agon*. Frequently, Nietzsche emphasizes the themes of individual growth, personal improvement, and self-overcoming. For example, Nietzsche writes, "And life itself confided this secret to me: "Behold," it said, " I am *that which must always overcome itself*" (Z II: 12). Even the most excellent individuals, according to Nietzsche, "there is something that must be overcome" (Z III: "Old and New" 14). Nietzsche seems to establish self-overcoming as a perpetual enterprise for individuals. Every individual needs to overcome something, to improve themselves, to compete with themselves. Individuals should always be in *agon* with themselves.

The essence of life, according to Nietzsche, is not only to preserve individuality, authenticity, and excellence but also to overcome one's *excellence*, to create new *arete*. Through the *agon* with oneself, the self could always be in a state of becoming that is never complete, being, and finished but always growing, striving, and improving. *To contest with our life, this seems to be Nietzsche's quintessential doctrine of life*. We have to constantly compete with ourselves. The Nietzschean sense of life, therefore, is always a contest or a form of *agon* with oneself. The *agonal* spirit allows us to be in a perpetual state of self-growth, self-creation, and becoming.

Nietzsche encourages us to infuse the *agonal* spirit in our life: Nietzsche urges us to view our life as a form of *agon*, a contest. Thus, Nietzsche proposes the notion of the eternal recurrence of the same. As Nietzsche describes it:

*The greatest weight:* -- What, if some day or night a demon were to steal after you into your loneliest loneliness and say to you: "This life as you now live it and have lived it, you will have to live once more and innumerable times more; and there will be nothing new in it, but every pain and every joy and every thought and sigh and everything

unutterably small or great in your life will have to return to you, in the same succession and sequence--even this spider and this moonlight between the trees, and even this moment and I myself. The eternal hourglass of existence is turned upside down again and again, and you with it, speck of dust! Would you not throw yourself down and gnash your teeth and curse the demon who spoke thus? Or have you once experienced a tremendous moment when you would have answered him: "You are a god and never have I heard anything more divine!" If this thought gained possession of you, it would change you as you are or perhaps crush you. The question in each and every thing, "Do you desire this once more and innumerable times more?" would lie upon your actions as the greatest weight. Or how well disposed would you have to become to yourself and to life *to crave nothing more fervently* than this ultimate eternal confirmation and seal? (*GS*: 341).

Nietzsche's notion of eternal recurrence seems to pose this thought experiment to us: What should you do if the life you now live and have lived repeats itself over and over again? Should you negate your own life? Or should you affirm it? Indeed, Nietzsche's eternal recurrence is essentially a thought experiment for the problem of modern nihilism: both in the eternal recurrence and modern nihilism, life has no inherent value or meaning in it. Should we negate our life and be inauthentic to ourselves because of this? The answer that Nietzsche provides is, No. We should, as Nietzsche advocates, say "Yes to life" (*TI IX*: 4). That is, Nietzsche asks us to affirm and love our life even if it is completely meaningless. Instead of submitting ourselves to the nihilistic despair or the meaninglessness of eternal recurrence, we should be in constant *agon* with "the greatest weight" of life. We should infuse the *agonal spirit* into our life in order to overcome ourselves, our fate, and the eternal recurrence.

The Nietzschean sense of life, as a constant *agon* with oneself, is similar to the *agonal* relationships between gods and heroes and between gods or heroes and the struggles in the *agonal* age (Tuncel 44). But what is the ultimate result of the *agon* with oneself, if the nature of this *agon* is productive rather than destructive? Is it a new cycle of self-creation? Or is it the transcendence of the *self*? Or is it the creation of new values and the end of nihilism? Nietzsche's answer seems to be this: the birth of *Übermensch*. As Nietzsche writes in *Thus Spoke Zarathustra*,

Behold, I teach you the *Übermensch*. The *Übermensch* is the meaning of the earth. Let your will say: the *Übermensch shall be* the meaning of the earth! I beseech you, my brothers, *remain faithful to the earth*, and do not believe those who speak to you of otherworldly hopes! Poison-mixers are they, whether they know it or not. Despisers of life are they, decaying and poisoned themselves, of whom the earth is weary: so let them go. (*Z* "Prologue": 3).

*Übermensch* is the one who overcomes the eternal-recurrence. It is the one who overcomes modern nihilism, overcomes one's excellence and produces new excellence, and overcomes the worldly values and generates new values: *Übermensch* is the victor in the *agon* with life. *Übermensch's* victory in the *agon* allows her to replace the meanings and values created by god. Similar to the Greek gods who become gods after their *agon*, *Übermensch* transforms into a godly figure that defines meanings, truths, and values at her own discretion.<sup>13</sup>In short, *Übermensch* is the

master of the future who purses the will to life, creates and continues transfiguring herself, and defines her own existence through the creative *agonal* spirit. *Übermensch* is the result of Nietzsche's *agon* and the type of human that Nietzsche wants the posterity to be.

## Conclusion

In the book *Nietzsche*, Martin Heidegger suggests that Nietzsche's philosophy is an "inverted Platonism," and during the last years of his life, Nietzsche "labors at nothing else than the overturning of Platonism" (*N<sub>I</sub>* 154). Indeed, Nietzsche's entire philosophy is centered on Plato's philosophy: Nietzsche both speaks admiringly of Plato and repudiates Plato vehemently. Through overturning Platonism, Nietzsche intends to resolve the cultural crisis of modernity and give birth to a 'philosophy of future' that is free from nihilism and mediocrity, but, in this process of destroying Platonism, Nietzsche also preserves the core of Plato's philosophy that he considers as fundamental to the philosophers of future and a vision of flourishing culture: *agon*.

As I have presented in this article, the notion of *agon* is central to the entirety of Nietzsche's philosophy. In his critique of the Platonic model of truth, Nietzsche criticizes Plato for being unable to confront the *becoming* nature of the world and encourages us to confront the world of *becoming*: the way to confront the world of *becoming* is to embrace the *agonal* spirits in ourselves, to be in *agon* with the world of becoming. In his critique of the Platonic-moral interpretation of the world and Christian morality, Nietzsche criticizes them for suppressing personal growth, excellence, and individual authenticity; while criticizing these ethical theories and moralities, Nietzsche constructs his own vision of ethics: an ethic of *becoming*, personal excellence, life-affirmation, and *agon*. The central doctrine of Nietzsche's ethics is to be constantly in *agon* with oneself and others. Humans are always working toward a process of *becoming* that is never complete and finished. Nietzsche's ethics is the one that encourages perpetual personal growth and self-overcoming. This doctrine of *agon* seems to be Nietzsche's key message for posterity.

This is also Nietzsche's message to the 'philosophers of the future.' It is important for philosophers of the future to carry this *agonal* spirit in their philosophical enterprise: philosophers of the future are the ones who always compete for excellence, who always aim toward creativity, and who philosophize with a hammer. The enterprise of philosophy, to Nietzsche and myself, is less about discovering truth than about creating truths and concepts: philosophy is also an artistic practice. Indeed, philosophers of the future need to be always creative and finally find their own philosophical ethos. As Nietzsche writes in *Thus Spoke Zarathustra*, "This is *my* way; where is yours?" (*Z* III: "Spirit of Gravity" 2).

## Abbreviations

All works by Nietzsche are cited in the text by abbreviation and by volume or section number. Occasionally, I refer to *Digitale Kritische Gesamtausgabe Werke und Briefe* to ensure the accuracy of the translations. Plato's works are cited by title abbreviation and line numbers. Other works in philosophical cannon are also cited in conventional abbreviations and by line number or page number.

Nietzsche's works:

*BGE*     *Beyond Good and Evil*

<i>BT</i>	<i>The Birth of Tragedy</i>
<i>D</i>	<i>Daybreak</i>
<i>EH</i>	<i>Ecce Homo</i>
<i>GM</i>	<i>On the Genealogy of Morals</i>
<i>GS</i>	<i>The Gay Science</i>
<i>HC</i>	“Homer’s Contest” in Walter Kaufmann’s <i>The Portable Nietzsche</i>
<i>HH</i>	<i>Human, All-Too Human</i>
<i>PTG</i>	<i>Philosophy in the Tragic Age of the Greeks</i>
<i>TI</i>	<i>Twilight of Idols</i>
<i>TL</i>	“On Truth and Lies in a Nonmoral Sense” in Walter Kaufmann’s <i>The Portable Nietzsche</i>
<i>UT</i>	<i>Untimely Meditations</i>
<i>WLT</i>	<i>Writings from the Late Notebooks</i>
<i>WP</i>	<i>The Will to Power</i>
<i>Z</i>	<i>Thus Spoke Zarathustra</i> in Walter Kaufmann’s <i>The Portable Nietzsche</i>

Plato’s works:

<i>Eut.</i>	<i>Euthyphro</i>
<i>R.</i>	<i>Republic</i>
<i>Tht.</i>	<i>Theaetetus</i>

Other works:

<i>PA</i>	Aristotle: <i>Posterior Analytics</i>
<i>EHU</i>	Hume: <i>An Enquiry Concerning Human Understanding</i>
<i>N<sub>1</sub></i>	Heidegger: <i>Nietzsche: The Will to Power as Art</i>
<i>OP</i>	Empiricus: <i>Outlines of Pyrrhonism</i>

## End Notes

1. The culture of *agon* [ἀγών] was one of the most salient features of the Ancient Greeks, and the Greek culture of *agon* greatly influenced Nietzsche during his early years as a student, a classist, a philologist, and a philosopher. Given the broadness of this topic and the scope of this paper, I shall restrain myself from extensively discussing this topic here; however, I still would like to offer a brief introduction to the culture of *agon* in order to help readers better understand the historical context in which Nietzsche developed his philosophy and the notion of *agon*.

The practice of *agon* first originated from the polytheistic nature of the Greek religion: in Greek mythology, gods become gods through constant struggles, contests, and fights; several prominent examples are the fight between Zeus and his father, the contest between Athena and Arachne, and the competition between Hermes and Apollo. The Greeks also acknowledged this *agonal* nature of gods, and an example of this is the crowd’s plea for other gods to exile the war god Ares in Sophocles’s *Oedipus Rex*. This *agonal* practice among the Greek gods then transformed into the Greek culture, and it was developed into two kinds: physical *agon* (e.g., sports, wars, and the Olympic Games) and mental or spiritual *agon* (e.g., contests of drama, poetry, and philosophy). Nietzsche praises highly of the latter one in his essay “Homer’s Contest.”

However, the agonal practice, as Nietzsche suggests, is not toward annihilation and destruction. Instead, the essence of *agon* is its productive nature, as the transfiguration of gods through *agon* shows. The eminent classist and Nietzsche’s

colleague Jacob Burckhardt also points out the productive aspect of *agon*. As Jacob Burckhardt writes in *The Greeks and Greek Civilization*,

Thus after the decline of heroic kingship all higher life among the Greeks, active as well as spiritual, took on the character of the agon. Here excellence (*arete*) and natural superiority were displayed, and victory in the agon, that is noble victory without enmity, appears to have been the ancient expression of the peaceful victory of an individual. Many different aspects of life came to bear the marks of this form of competitiveness. We see it in the conversations and round-songs of the guests in the symposium, in philosophy and legal procedure, down to cock- and quail-fighting or the gargantuan feats of eating (Burckhardt, 165-166).

Burckhardt's interpretation that *agon* was a means to produce individual *arete* in Greek culture has a great influence on Nietzsche's intellectual formation and on his philosophy as a whole, and indeed the concepts of *agon* and *arete* are manifested throughout Nietzsche's philosophy; in particular, Nietzsche's interpretation of *agon* in his essay "Homer's Contest" was greatly influenced by Burckhardt's historiography. In the later sections of this article, I will explicitly point out the relationship between *agon* and *arete*, and I will suggest how this relationship connects to Nietzsche's philosophy as a whole.

For further information about the culture of *agon*, see Wang Daqing's "On the Ancient Greek ἀγών"; Yunus Tuncel's *Agon in Nietzsche*; Elton T. E.'s *Entering the Agon: Dissent and Authority in Homer, Historiography and Tragedy*; Jacob Burckhardt's *The Greeks and Greek Civilization*; Heather L. Reid, John Serrati, and Tim Sorg's *Conflict and Competition: Agon in Western Greece: Selected Essays from the 2019 Symposium on the Heritage of Western Greece*, and part five of this book offers two scholarly essays on the relationship between Nietzsche's philosophy and *agon*.

2. It is noteworthy that Nietzsche's understanding of ἀγών also encompasses the meanings of *challenging* or the *urge to contest* with others. Frequently, Nietzsche uses the word *Wettkampf*, which is the closest German translation for *agon*, to substitute the Greek term *agon* [ἀγών] in his writings; for example, Nietzsche's most important writing on the Greek *agon* is entitled "Homer's *Wettkampf*," which translates into English is "Homer's Contest." Different in its etymological sense from the Greek term ἀγών, the German term *Wettkampf* connotes both the meaning of challenging [*wetten*] and fighting/struggling [*kämpfen*] (Tuncel 11-12). Therefore, it is reasonable to assume that Nietzsche's notion of *agon* not only suggests contest as a quotidian activity but also suggests an individual's zest for contesting with others.

3. Of course, this is not to suggest that Nietzsche only considers Socrates as the turning point of philosophy due to his discovery of *agon*. I am aware that Nietzsche frequently remarks Socrates as the "decadent" or "the turning point of philosophy into the age of decadence." What I want to emphasize here, however, is that Socrates invents a new spiritual form of *agon* (i.e., philosophy); that is, Socrates seems to infuse the *agonal* spirit of the Greeks into philosophy. Nietzsche acknowledges Socrates's invention of philosophy as a form of *agon*, and he definitely admires Socrates for this. As Nietzsche writes,

*The dying Socrates.* I admire the courage and wisdom of Socrates in everything he did, said -- and did not say. This mocking and enamored monster and pied piper of Athens, who made the most arrogant youths tremble and sob, was not only the wisest talker who ever lived: he was just as great in his silence... (GS: 340).

Although he does not explicitly use the word *agon* in the passage, Nietzsche apparently admires the *agonal* spirit of Socrates. Nietzsche might mock Socrates for his pessimism and hatred toward life, but Nietzsche definitely admires Socrates's *agonal* spirit and intends to instantiate this spirit in his philosophy.

However, I do not intend to pretend to be an expert at this complex relationship between Nietzsche and Socrates; the point I want to make is simple: Nietzsche admires Socrates for his *agonal* spirit. For an extensive discussion on the relationship between Nietzsche and Socrates, see Walter Kaufmann's *Nietzsche*, pp 391-411; Alexander Nehemas's *The Art of Living*, pp 128-157; Alexander Nehemas's *Nietzsche: Life as Literature*, pp 14-41; Werner Dannhauser's *Nietzsche's View of Socrates*.

4. Although Nietzsche contends that Socrates is the transitional figure in Greek intellectual history, his understanding of *agon* as the essence of Greek culture originates from Homer, Heraclitus, and other pre-Socratic figures. The influence of the culture of *agon* on Nietzsche can be best found in his essay "Homer's Contest." In the essay, Nietzsche subverts the modern understanding of the Greeks as humane and instead portrays the Greeks as *agonal* and competitive. The relationship between Nietzsche and Heraclitus is more complicated than his relationship to Homer. Nietzsche seems to derive a cosmology of *agon* and *becoming* from Heraclitus's philosophy. As Nietzsche explains his cosmological view:

[T]he world itself is a mixed drink which must constantly be stirred. The strife of the opposites gives birth to all that comes-to-be; the definite qualities which look permanent to us express but the momentary ascendancy of one partner. But this by no means signifies the end of the war; the contest endures in all eternity. Everything that happens, happens in accordance with this strife, and it is just in the strife that eternal justice is revealed. It is a wonderful idea, welling up from the purest strings of Hellenism, the idea that strife embodies the everlasting sovereignty of strict justice, bound to everlasting laws. Only a Greek was capable of finding such an idea to be the fundament of a cosmology; it is Hesiod's good. *Eris* transformed into the cosmic principle; it is the contest-idea of the Greek individual and the Greek state, taken from the gymnasium and the palaestra, from the artist's *agon*, from the contest between political parties and between cities--all transformed into universal application so that now the wheels of the cosmos turn on it (PTG 5).

For a more extensive discussion on Heraclitus's influence on Nietzsche's development of the notion of *agon*, see Yunus Tuncel's *Agon in Nietzsche*, pp 69-88; Daw-Nay N.R. Evans Jr., *Nietzsche and classical greek philosophy: Essays on Socrates, Plato, Aristotle, and Heraclitus*, pp 138-175.

5. Although Nietzsche uses the words "*krieg*" and "*kriegerisch*" to express his *agonal* spirit in this passage, we should never understand Nietzsche's *krieg* in its literal sense; instead, we should understand Nietzsche's "*krieg*" as a form of contest, competition, or an *agon*. Nietzsche's *krieg* and *agon* is not about destruction or annihilation; rather, it is about creation. In the essay "Homer's Contest," Nietzsche writes that "[a]nd not only Aristotle but the whole of Greek antiquity thinks differently from us about hatred and envy, and judges with Hesiod, who is one place calls one *Eris* evil -- namely, the one that leads men into hostile fights of annihilation against one another -- while praising another *Eris* as good -- the one that, as jealousy, hatred, and envy, spurs men to activity; not to the activity of fights of annihilation but to the activity of fights which are contests" (HC 35). Similar to the Ancient Greeks, Nietzsche only praises *contests* and condemns complete

destruction and annihilation; therefore, what Nietzsche emphasizes in “*krieg*” is an individual's engagement with the “*personal fight*” rather than *hor concours*.

For further discussion on Nietzsche's notion of war, see Walter Kaufmann's *Nietzsche: Philosopher, Psychologist, Antichrist*, pp 386 - 390; H. Siemens's “Nietzsche's Hammer: Philosophy, Destruction, or the Art of Limited Warfare”; Yunus Tuncel's *Agon in Nietzsche*, pp 89-105.

6. By the capitalized truth, “Truth,” I mean that it is metaphysical, transcendental, absolute, and outside the human world. It is a truth that belongs only to Plato's *noesis* or Kant's *noumena*. When truth is not capitalized, I mean it in a Nietzschean sense: a ‘truth’ that dwells in the world of becoming or an interpretation of the world.

7. I am not unaware of the scholarly controversy on the distinction between Plato the writer and Socrates the philosopher, but this does not affect my argument, for Plato in either way embodies the Socratic style in his philosophy.

8. Another aspect of Plato's philosophy that Nietzsche is criticizing here is Plato's theory of language. See Plato's *Cratylus*.

9. Nietzsche's psychological analysis of the pursuit of Truth seems to resonate with the Pyrronhian skeptics, as they both reject this idea of the attainment of certainty through the discovery of Truth. As Sextus Empiricus writes,

For the man who opines that anything is by nature good or bad is for every being disquieted: when he is without the things which he seems good he believes himself to be tormented by things naturally bad and he pursues after the things which are, as he thinks, good; which when he has obtained he keeps falling into still more perturbations because of his irrational and immoderate elation, and in his dread of a change of fortune he uses every endeavor to avoid losing the things which he deems good. On the other hand, the man who determines nothing as to what is naturally good or bad neither sun nor pursues anything eagerly; and, in consequence, he is unperturbed (*OP*, I. 27)

Indeed, in a way, Nietzsche's critique of truth as *being* could be understood through the Pyrronhian critique of Platonic rationalism: we could never acquire quietude and certainty through the pursuit of truth, as we have no warrant to claim we discover the truth. Does the pursuit of truth offer us quietude? Or does it disturb us more? Nietzsche makes the same point here: Nietzsche seems to suggest that, paradoxically, the state of becoming might be a better way for us to acquire quietude.

10. Although Nietzsche never explicitly uses the term authenticity in his writings, Nietzsche's *Wahrhaftigkeit* (truthfulness) seems to be the term he uses for authenticity, as he frequently distinguishes *Wahrhaftigkeit* (truthfulness) from *Wahrheit* (truth), and the former is also a synonym for the Heideggerian term *eigentlich* that later developed by existentialist as *authentisch* and *authentique* (Golomb 243). In the rest of my essay, therefore, I will treat Nietzsche's *Wahrhaftigkeit* as a synonym or substitute for authenticity. For further discussion on Nietzsche's authenticity, see Jacob Golomb's “Nietzsche on Authenticity.”

11. Although this subject is often overlooked by Nietzsche scholars and commentators, it is worth noting that Nietzsche's concept of the self does not denote a preexisting and unified self. Nietzsche's self does not aim at attaining ‘true nature’ within us or self-realization but rather aims at creating the self – the Nietzschean selfhood a self of

*becoming* rather than a self of *being*. This self of becoming is central to Nietzsche's philosophy and has a deep connection to Nietzsche's *agon*, which I will discuss later in the essay. For further interest in the discussion on Nietzsche's concept of selfhood, see Paul Franco's "Becoming Who You are: Nietzsche on Self-Creation"; Alexander Nehamas's *Nietzsche: Life As Literature*, pp 170-199.

12. It is worth noting the concept of *ethos*, as *ethos* is one of the most important themes in Nietzsche's philosophy. The word *ethos* comes from the Greek word *ἦθος*, meaning "character, habitual way of life," or, as Heraclitus defines, *ethos* is a person's "*daimon*" (Peters 66). Throughout his philosophy, Nietzsche constantly emphasizes on the importance of *ethos*. As Nietzsche writes, "*One thing is needful -- To "give style" to one's character -- a great and rare art! It is practiced by those who survey all the strengths and weaknesses of their nature and then fit them into an artistic plan until every one of them appears as an art and reason and even weakness delight the eye*" (GS: 290). Indeed, Nietzsche believes that every individual needs to find their own *ethos* or style: we need to cultivate our own *ethos*. This cultivation of *ethos* ultimately transforms to *ethikos* [ἠθικός]. To Nietzsche, thus, the cultivation of individual characters is the main principle of his ethics: the ethics of characters.

13. For the discussion on the transformation of *Übermensch*, see my second endnote.

## Bibliography

- Anderson, Mark. *Plato and Nietzsche: Their Philosophical Art*. Bloomsbury, 2014.
- Ansell-Pearson, Keith. "In Search of Authenticity and Personality: Nietzsche on the Purifications of Philosophy," *American Catholic Philosophical Quarterly*, Vol.84. No.2, pp. 283-312.
- Aristotle. *Aristotle: Introductory Readings*. Translated and edited by Terence Irwin and Gail Fine. Hackett, 1996.
- Barker, Elton T. E. *Entering the Agon: Dissent and Authority in Homer, Historiography and Tragedy*. Oxford University Press, 2009
- Bishop, Paul. *Nietzsche and Antiquity: His Reaction and Response to the Classical Tradition*. Boydell & Brewer, 2004.
- Brobjer, Thomas H. "Nietzsche's Affirmative Morality: An Ethics of Virtue." *Journal of Nietzsche Studies*, no. 26, 2003, pp. 64-78. JSTOR, [www.jstor.org/stable/20717820](http://www.jstor.org/stable/20717820).
- Burckhardt, Jacob, and Oswyn Murray. *The Greeks and Greek Civilization*. St. Martin's Press, New York, 1998
- Burnham, Douglas. *The Nietzsche Dictionary*. Bloomsbury Academic, 2015.
- Clark, Maudemarie. *Nietzsche on Truth and Philosophy*. Cambridge University Press, 1990.
- Daigle, Christine. "Nietzsche: Virtue Ethics... Virtue Politics?" *Journal of Nietzsche Studies*, no. 32, 2006, pp. 1-21. JSTOR, [www.jstor.org/stable/20717883](http://www.jstor.org/stable/20717883).
- Dannhauser, Werner J. *Nietzsche's View of Socrates*. Cornell University Press, 1974.
- Daw-Nay N.R. Evans Jr. *Nietzsche and classical greek philosophy: Essays on Socrates, Plato, Aristotle, and Heraclitus*. 2012. DePaul University, PhD Dissertation.
- Elgat, Guy. "Judgments That Have Value 'Only as Symptoms': Nietzsche on the Denial of Life in *Twilight of the Idols*." *Journal of Nietzsche Studies*, vol. 48, no. 1, 2017, pp. 4-16. JSTOR, [www.jstor.org/stable/10.5325/jnietstud.48.1.0004](http://www.jstor.org/stable/10.5325/jnietstud.48.1.0004).



- Empiricus, Sextus. *Outlines of Pyrrhonism*. Translated by R. G. Bury. Prometheus Books, 1990.
- Finnigan, Maureen. "Nietzsche's Perspective: Beyond Truth as an Ideal." in: *Topics in Feminism, History and Philosophy, IWM Junior Visiting Fellows Conferences, Vol. 6*, edited by Rogers, Dorothy, Joshua Wheeler, Marina Zavacká, and Shawna Casebier. Vienna: IWM 2000
- Franco, Paul. "Becoming Who You Are: Nietzsche on Self-Creation." *Journal of Nietzsche Studies*, vol. 49, no. 1, 2018, pp 52–77 JSTOR, [www.jstor.org/stable/10.5325/jnietstud.49.1.0052](http://www.jstor.org/stable/10.5325/jnietstud.49.1.0052).
- Golomb, Jacob, and DePaul University. "Nietzsche on Authenticity." *Philosophy Today (Celina)*, vol. 34, no. 3, 1990, pp. 243-258
- Harper, Aaron. "Nietzsche's Thumbscrew: Honesty as Virtue and Value Standard." *Journal Nietzsche Studies*, vol. 46, no. 3, 2015, pp. 367–390.
- Harris, Daniel I. "Nietzsche on Honesty and the Will to Truth." *JBSP. Journal of the British Society for Phenomenology*, vol. 51, no. 3, 2020, pp. 247-258.
- Heidegger, Martin. *Nietzsche: Volumes One and Two*. Translated by David Farrell Krell. Harper Collins, 1991.
- Holmes, Robert L. *Basic Moral Philosophy*. Wadsworth, 2007.
- Hume, David. *An Enquiry Concerning Human Understanding*. Edited by Tom L. Beauchamp. Oxford University Press, 2010.
- Kaufmann, Walter. *Nietzsche: Philosopher, Psychologist, and Antichrist*. Princeton University Press, 2013.
- Kellner, Douglass. "Modernity and Its Discontents: Nietzsche's Critique." UCLA. Unpublished chapter from a book. [https://pages.gseis.ucla.edu/faculty/kellner/papers/FNmod.htm#\\_edn1](https://pages.gseis.ucla.edu/faculty/kellner/papers/FNmod.htm#_edn1)
- Leiter, Brian. *Routledge Philosophy Guidebook to Nietzsche on Morality*. Routledge Books, 2002.
- Nehemas, Alexander. *Nietzsche: Life as Literature*. Harvard University Press, 2002.
- Nehemas, Alexander. *The Art of Living: Socratic Reflections from Plato to Foucault*. University of California Press, 2000.
- Nietzsche, Friedrich. *Twilight of the Idols and The Anti-Christ*. Translated by R. J. Hollingdale. Penguin Books, 1990.
- . *Digitale Kritische Gesamtausgabe Werke und Briefe*. Edited by Paolo D'Iorio. Paris: Nietzsche Source, 2009. <http://www.nietzschesource.org/#eKGWB>
- . *Beyond Good and Evil: Prelude to a Philosophy of the Future*. Translated by Walter Kaufmann. Vintage Books, 1989.
- . *The Birth of Tragedy and the Case of Wagner*. Translated by Walter Kaufmann. Vintage Books, 1967.
- . *On the Genealogy of Morals and Ecce Homo*. Translated by Walter Kaufmann. Vintage Books, 1989.
- . *The Gay Science: With a Prelude in Rhymes and an Appendix of Songs*. Translated by Walter Kaufmann. Vintage Books, 1974.
- . *Untimely Meditations*. Edited by Daniel Breazeale. Cambridge University Press, 2007.
- . *The Portable Nietzsche*. Edited and Translated by Walter Kaufmann. Penguin Books, 1976.
- . *Writings from the Late Notebooks*. Edited by Rudiger Bittner. Cambridge University Press, 2003.
- . *Philosophy in the Tragic Age of the Greeks*. Translated by Marianne Cowan. Regnery Publishing, 1998.

- . *Daybreak: Thoughts on the Prejudice of Morality*. Translate by R. J. Hollingdale. Edited by Brian Leiter and Maudemarie Clark. Cambridge University Press, 1997.
- . *Will to Power*. Translated by Walter Kaufmann and R. J. Hollingdale. Vintage Books, 1968.
- . "Homer's Contest." Translated by Christa Davis Acampora. Nietzscheana, 1996. <http://www.northamericannietzschesociety.com/uploads/7/3/2/5/73251013/nietzscheana5.pdf>
- Peters, F. E. *Greek Philosophical Terms: A Historical Lexicon*. New York University Press, 1967.
- Pippin, Robert B. "Nietzsche and the Melancholy of Modernity." *Social Research*, vol. 66, no. 2, 1999, pp. 495–520. *JSTOR*, [www.jstor.org/stable/40971334](http://www.jstor.org/stable/40971334).
- Plato. *Five Dialogues: Euthyphro, Apology, Crito, Meno, Phaedo*. Translated by G. M. A. Grube. Hackett, 2002.
- . *Republic*. Translated by G. M. A. Grube. Hackett, 1992.
- . *The Theaetetus of Plato*. Translated by M. J. Levett. Hackett, 1990.
- Reid, Heather L., et al., editors. *Conflict and Competition: Agon in Western Greece: Selected Essays from the 2019 Symposium on the Heritage of Western Greece*. Vol. 5, Parnassos Press – Fonte Aretusa, 2020. *JSTOR*, [www.jstor.org/stable/j.ctv15tt78p](http://www.jstor.org/stable/j.ctv15tt78p).
- Reginster, Bernard. *The Affirmation of Life: Nietzsche on Overcoming Nihilism*. Harvard University Press, 2006.
- Siemens, H. "Nietzsche's Hammer: Philosophy, Deconstruction, or the Art of Limited Warfare." *Tijdschrift Voor Filosofie*, vol. 60, no. 2, 1998, pp. 321–347.
- Solomon, Robert C. *Living with Nietzsche: What the Great "Immoralist" has to Teach Us*. Oxford University Press, 2004.
- Tuncel, Yunus. *Agon in Nietzsche*. Marquette University Press, 2013.
- Wang, Daqing. "On the Ancient Greek  $\alpha\gamma\omega\upsilon$ ." *Procedia Social and Behavioral Sciences* 2, 2010, pp 6805–6812, Science Direct, <https://doi.org/10.1016/j.sbspro.2010.05.031>.
- Zuckert, Catherine. *Postmodern Platos: Nietzsche, Heidegger, Gadamer, Strauss, Derrida*. The University of Chicago Press, 1996.



# The Roles of Extracellular Vesicles in Cancer and the Investigation of Exosomal PD-L1 in Immunotherapy of Colorectal Carcinoma

---

*Yuanmuhuang Long*

**Author background:** *Yuanmuhuang Long grew up in China and currently attends the High School Affiliated to Shanghai Jiao Tong University, in Shanghai, China. Her Pioneer research concentration was in the field of biology and titled “Cell Biology.”*

## Abstract

Extracellular vesicles are small membrane vesicles which can be released by various types of cells. There are three types of these small membrane sacs: exosomes, microvesicles and apoptotic bodies (Stahl et al., 2019, p.11). Two of them, exosomes and microvesicles, are believed to play important roles in the process of tumor progression in several ways: helping the transfer of aggressive phenotypes, angiogenesis, the development of a favorable microenvironment for cancer invasion, multi-drug resistance and immune escape (Muralidharan-Chari et al., 2010, p.1606).

Programmed death-ligand 1 (PD-L1), the major ligand of immune checkpoint programmed death-1 (PD-1), is expressed in the tumor exosomes. The bonding between PD-L1 and PD-1 is able to reduce the activity of human T-cells, which helps the tumor escape the immune surveillance (Iwai et al., 2002, p.12993; Taube et al., 2014, p.5065). In normal tumor patients, the tumor-specific immune response is usually very weak, not strong enough to eliminate the tumor. Thus, the blockade of the PD-1/PD-L1 pathway, which can reasonably increase the activity of human T cells and the tumor-targeted immune response, has become a possible route for cancer immunotherapy. It has already shown clinical benefits in several types of cancer, including non-small-cell lung cancer, head and neck cancer, etc. (Theodoraki et al., 2018, p.896; Kim et al., 2019, p.1; Rajan & Gulley, 2014, p.403). However, for colorectal carcinoma, the function and mechanism of PD-L1 exosomes in tumor development are not thoroughly understood and the effectiveness of PD-1/PD-L1 as a therapy pathway is not fully determined either. This paper discusses the functions of exosomes in cancer progression and proposes an experiment to test the function and mechanism of PD-L1 exosomes in colorectal carcinoma.

## 1. Introduction

Extracellular vesicles are small membrane sacs which are produced by various cell types. These extracellular vesicles are divided into three categories: exosomes, microvesicles and apoptotic bodies which are different in size and methods of formation (Ståhl et al., 2019, p.11). The exosomes have a diameter between 30 nm and 200 nm, while the microvesicles have a size ranges from 100 nm to 1000 nm. The apoptotic bodies are even larger, with a size ranges from 1µm to 4 µm (Turturici et al., 2014, p.622). Microvesicles are derived directly from the outward budding of cell membranes and the following fission (Muralidharan-Chari et al., 2010, p.1064). Exosomes are produced by the fusion of multivesicular bodies with the cell membranes and the following release of intraluminal vesicles (ILVs) into the extracellular environment (Schorey & Bhatnagar, 2008, p.871). The apoptotic bodies are formed at the later stage of cell apoptosis (Turturici et al., 2014, p.622). Microvesicles and exosomes are believed to be actively released by the cells and contain different kinds of cargoes, including molecules, nucleic acids, proteins, microRNAs and mRNAs (Ståhl et al., 2019, p.12). According to the cargoes, microvesicles and exosomes have many functions, and some of them are able to affect and promote tumor progression in a variety of ways.

One of the ways microvesicles and exosomes help tumor development is to help the tumor escape the immune surveillance by downregulating the activity of immune cells (Muralidharan-Chari et al., 2010, p.1068). Thus, microvesicles and exosomes have become possible targets for cancer immunotherapy. Actually, there are already successful examples. Cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) is a negative costimulatory receptor of the immunoglobulin family expressed in tumor microvesicles, which can restrict the activity of T cells. Ipilimumab, an inhibitor of CTLA-4, has been proven to have reliable responses in about 20% of patients and increase the overall survival of metastatic melanoma patients (Taube et al., 2014, p.5064-5065; Stephen & Hajjar, 2018, p.18). PD-1 is another negative costimulatory receptor and the PD-1/PD-L1 pathway has become a popular target for cancer immunotherapy. According to previous research, PD-1 is an immune checkpoint with 2 ligands: PD-L1 and PD-L2 (Taube et al., 2014, p.6054; Stephen & Hajjar, 2018, p.12). The binding of PD-1 and PD-L1 is able to decrease the activity of the T-cells and sometimes induces apoptosis (Kim et al., 2019, p.1-2). PD-1 and PD-L1 are found in exosomes derived from plasma samples of patients with various types of cancer (Taube et al., 2014, p.5065). Thus, the blockage of the PD-1/PD-L1 pathway is another attractive immunotherapy trail.

Colorectal cancer projected to cause 53,200 deaths in 2020 in the United States. It ranks in the second place for the number of deaths caused by cancer each year for men and women combined, with a 64% 5-year survival rate (Siegel et al., 2020, p.9; American Society of Clinical Oncology (ASCO), n.d.). In order to determine whether the blockage of the PD-1/PD-L1 pathway is a possible immunotherapy for colorectal cancer, several experiments are needed. Because the expression of the protein PD-L1 varies according to cancer types (Taube et al., 2014, p.5064), it is necessary to determine whether protein PD-L1 is expressed in colorectal cancer and its expression level. The effect of PD-L1 exosomes on the T cell activity and its effect on tumor growth of mice should also be tested.

## 2. Background Knowledge

### 2.1 Formation of microvesicles

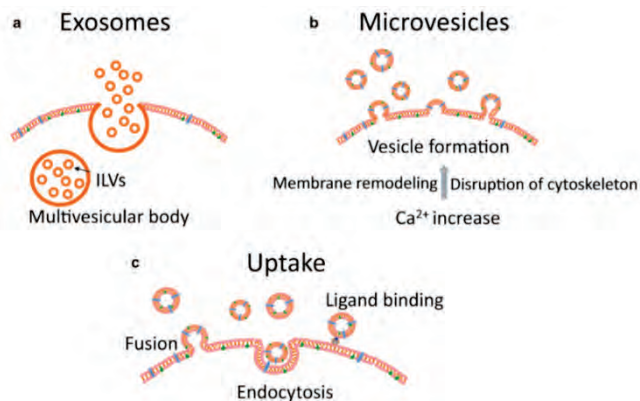
The formation of microvesicles is divided into two stages: the budding of the cell membrane and the following fission process (Muralidharan-Chari et al., 2010, p.1064). Different kinds of proteins and the composition of the cell membrane play a role. Contractile proteins can add force to one leaflet of the membrane, so that a curvature is created (Muralidharan-Chari et al., 2010, p.1064). Floppases and flippases can control the movement of phospholipids. Ca<sup>2+</sup> dependent scramblase can randomize the phospholipid of the cell membrane and remove the curvature restriction (Muralidharan-Chari et al., 2010, p.1065). Sphingomyelinase, a protein which hydrolyzes sphingomyelin into ceramide, causes the bending of the plasmid membrane by producing a ceramide gradient (Taniguchi & Okazaki, 2014, p.692). The function of sphingomyelin as cargo is described later. Phosphatidylserine is another important protein with many functions: it is located at the extracellular leaflet of the vesicles, changes the shape of the membrane, and promotes cytoskeleton detachment (Muralidharan-Chari et al., 2010, p.1065). In addition, the composition of the plasmid membrane also plays a role: microvesicles usually contain a high degree of cholesterol. Reducing the amount of cholesterol in the cell membrane can reduce the production of vesicles (Muralidharan-Chari et al., 2010, p.1065).

Contractile proteins are responsible for the fission of the budding vesicles. Phosphorylated myosin light chain kinase (MLCK) and myosin 1A (MYO1A) pinch the neck of the vesicles from inside the microvesicles. The GTP-binding protein ADP-ribosylation factor 6 (ARF6) acts as a regulator: it regulates MLCK, the subsequent myosin light chain phosphorylation and, finally, the release of microvesicles (Muralidharan-Chari et al., 2010, p.1065).

### 2.2 Formation of exosomes

Exosomes are formed by the fusion of multivesicular bodies with the cell membrane and the following release of intraluminal vesicles (ILV) into the extracellular environment. The exosomes can carry a variety of cargoes, including proteins, mRNA, microRNA, and lipids (Muralidharan-Chari et al., 2010, p.1603; Ståhl et al., 2019, p.12). The process of sorting specific cargoes into multivesicular bodies is very complex and is regulated by many proteins. The ubiquitination of targets is one of the sorting mechanisms. It has been found that monoubiquitination of endosomal proteins can act as a sign for the sorting process. In addition, the oligoubiquitination can act as a sorting signal, which increases the sorting efficiency of multivesicular bodies (Taylor & Gercel-Taylor, 2011, p.443). Hetero-oligomeric protein complexes also play an important role in the sorting: the endosomal sorting complex required for transport I, II and III (ESCRT-I, ESCRT -II and ESCRT -III) are responsible for the recognition of monoubiquitinated cargoes. They also send the monoubiquitinated cargoes into multivesicular bodies (Taylor & Gercel-Taylor, 2011, p.454). The adenosine triphosphatase vacuolar protein sorting 4 (Vps4) helps the dissociation of ESCRTs from multivesicular bodies once the sorting process is finished (Schorey & Bhatnagar, 2008, p.872). The hepatocyte growth factor exhibits many functions in exosome production: it regulates the ubiquitinated cargo which

binds with tyrosine kinase substrate (HRS), gathers the ESCRTs, and sends the cargo for inclusion into multivesicular bodies (Schorey & Bhatnagar, 2008, p.872). Non-ubiquitinated protein also exists in the sorting process. Transferrin receptor proteins, for example, can interact with ESCRTs without ubiquitination (Schorey & Bhatnagar, 2008, p.872). Protein aggregation is also known to affect the trafficking of protein to the multivesicular bodies (Schorey & Bhatnagar, 2008, p.872). In general, there are multiple mechanisms in the process of sorting cargoes into multivesicular bodies.



**Figure 1.** The formation of microrvesicles and exosomes (Ståhl et al., 2019)

### 2.3 Acquisition of aggressive cancerous phenotypes

The first way extracellular vesicles help tumor progression is by transferring aggressive phenotypes. It has been discovered that tumor cells are not equally aggressive and the more aggressive tumor cells have a higher tendency to spread, which may lead to secondary tumors (Muralidharan-Chari et al., 2010, p.1606). The aggressive tumor cell can transfer oncogenic receptors to non-aggressive cells through extracellular vesicles. It has been found that the epidermal growth factor receptor variant III (EGFRvIII), an oncogenic receptor only found in aggressive glioma, can be horizontally transferred into normal tumor cells through extracellular vesicles, changing the expression of original non-aggressive cells into an aggressive phenotype (Muralidharan-Chari et al., 2010, p.1607).

### 2.4 Extracellular vesicles and tumor angiogenesis

Extracellular vesicles with proangiogenic factors can stimulate the growth and invasion of endothelial cells. These endothelial cells turn into blood vessels which enable the tumor to get nutrients and remove waste more efficiently, promoting the progression of the tumor (Muralidharan-Chari et al., 2010, p.1607). In addition, extracellular vesicles can carry mRNA, which encodes growth factors like vascular endothelial growth factor (VEGF), thus stimulating the growth of endothelial cells. MicroRNA and mRNA in the extracellular vesicles can also change the translational profile of endothelial cells (Stephen & Hajjar, 2018, p.3; Muralidharan-Chari et al., 2010, p.1607; Pap, 2011, p.187).

## **2.5 The impact of extracellular vesicles on the microenvironment can help the invasion of tumors**

The degradation of the extracellular matrix and the changing of the microenvironment into a tumor-favorable one can significantly promote the metastasis of a tumor. The extracellular vesicles are able to change the microenvironment in favor of tumor progression by increasing the resistance and motility to apoptosis of surrounding fibroblasts (Peinado, 2011, p142). In addition, extracellular vesicles with proteases matrix metalloproteinase 2 and 9 (MMP2 and MMP9), membrane type 1 matrix metalloproteinase (MT1-MMP) and extracellular matrix metalloproteinase inducer (EMMPRIN), as well as their corresponding zymogens urokinase-type plasminogen activator (uPA) provide another pathway for the degradation of the matrix. The uPA catalyzes the conversion of plasminogen into plasmin, which then activates MMP zymogens (D'Souza-Schorey & Clancy, 2012, p.1290). The Cathepsin B, which is active specifically in the low PH environments produced by tumors, turns MMP zymogens into the activation form which degrade the basement collagens (Pap, 2011, p.192). Sphingomyelin, which is normally a constituent of membrane microdomains, is another major cargo in extracellular vesicles. It can provide endothelial cells with migratory and angiogenesis-inducing properties (Taniguchi & Okazaki, 2014, p.692; Schorey & Bhatnagar, 2008, p.874; Muralidharan-Chari et al., 2010, p.1607). There is also a positive correlation between the invasiveness of the tumor and the number of extracellular vesicles, though the invasive ability of vesicles does not change with the stage of the cancer (Muralidharan-Chari et al., 2010, p.1606). A study of malignant ovarian ascites showed that the patients with more advanced cancer contained more vesicles than those with early-stage cancer. Similar results were also discovered with breast cancer (Muralidharan-Chari et al., 2010, p.1606).

## **2.6 Extracellular vesicles help the tumor to escape apoptosis**

Extracellular vesicles help the tumor cell escape the cell apoptosis. It has been found that after chemotherapy, the tumors are able to shed the chemotherapy drugs through microvesicles, which leads to the development of drug resistance and the enhancement of tumor survival. In one study, doxyrubicin, a drug used to treat cancer, was found in the microvesicles. The drug-efflux mechanism through extracellular vesicles explained how a tumor develops multi-drug resistance and how chemotherapy becomes useless (Muralidharan-Chari et al., 2010, p.1609). The extracellular vesicles also act as a protective mechanism against intracellular stress. The tumor cells can avoid the accumulation of caspase-3, a protein which can lead to cell apoptosis, by releasing it through extracellular vesicles (Turturici et al., 2013, p.625).

## **2.7 Extracellular vesicles help to escape immune surveillance**

In normal cases, the immune system of the host spontaneously attacks tumor cells once the tumor antigen is presented. However, when a tumor progresses, multiple immune escape mechanisms are developed to promote tumor progression. Microvesicles and exosomes released by the tumor can help it escape the attack from immune system by modifying the antigen-presenting cells or eliminating the anti-tumor effect of the cells (Muralidharan-Chari et al., 2010, p.1608).

Fusion of microvesicles with monocytes can stop the normal differentiation of monocytes. Instead, the monocytes differentiate into immunosuppressive cells with a CD14<sup>+</sup> HLA-DR<sup>-</sup> phenotype (Baj-Krzyworzeka et al., 2007, p.76). These CD14<sup>+</sup> HLA-DR<sup>-</sup> cells are able to produce immunosuppressive cytokines. The transforming growth factor beta (TGFβ) for example, can inhibit the functions of T lymphocyte and thus help the tumor escape the immune response. It has also been found that the level of cells with a CD14<sup>+</sup> HLA-DR<sup>-</sup> phenotype is higher in patients with advanced melanoma (Valenti et al., 2007, p.2912).

Tumor microvesicles (TMVs) can also malfunction to induce apoptosis in anti-tumor T cells, promoting the progression of the tumor. A study of melanosomes discovered that TMVs are able to induce Fas ligand (FasL) mediated and dose-dependent apoptosis on Jurkat cells. The FasL pathway is confirmed by the inhibition effect of anti-Fas monoclonal antibodies clone ZB4 (anti-Fas mAb ZB4) (Andreola et al., 2002, p.1303-1306). TMVs with tumor necrosis factor related apoptosis inducing ligand (TRAIL) are another pathway for tumor cells to induce T cell apoptosis (Taylor & Gercel-Taylor, 2011, p.447-448; Huber et al., 2005, p.1797). The expression of PD-1 and its ligands can also diminish the activity of T cells and induce apoptosis. PD-1 is an immune checkpoint with 2 ligands: PD-L1 and PD-L2. In normal conditions, PD-1 is expressed by both the T cells and the B cells. When PD-L1, which is usually expressed by macrophages, binds with PD-1, the function of the T-cell is diminished and apoptosis may happen. PD-L2 is another ligand, which is expressed by macrophages and dendritic cells. The role of PD-L2 in immunomodulation is still not fully understood (Stephen & Hajjar, 2018, p.12-13). In many studies, researchers found that the level of PD-L1 protein in exosomes is different in different types of cancer. The study done by Taube J.M. (2014, p.5065) covered non-small-cell lung cancer, head and neck cancer, melanoma and renal cell carcinoma. These exosomes extracted from patients' plasma were able to decrease the production of interferon gamma (INF-γ) of T cells (Kim et al., 2019, p.5). Tumor-derived vesicles can also promote the production of myeloid-derived suppressor cells (MDSCs), and thus suppress T cell immune functions (Noman et al., 2014, p.781).

Immune cells other than monocytes and T cells can be influenced by tumor microvesicles as well. The function of natural killer cells is also influenced by the expression of major histocompatibility complex class I-related chain protein A and B (MICA and MICB) (Taylor & Gercel-Taylor, 2011, p.447). These not only block the recognition of tumor cells by the immune system, but also inhibit the natural killer group 2 member D (NKG2D) expression of various immune cells, including gamma delta T cells and circulating CD8<sup>+</sup> T cells (cytotoxic T cells which express cluster of differentiation 8) (Taylor & Gercel-Taylor, 2011, p.447). The gathering of immune cells is also inhibited by TMVs through the expression of membrane bound intercellular adhesion molecule 1 (ICAM-1), which blocks the interaction between endothelial cells and lymphocytes (Taylor & Gercel-Taylor, 2011, p.448). Microvesicles with latent membrane protein 1 (LMP-1) have the ability to suppress the proliferation of leukocyte cells (Muralidharan-Chari et al., 2010, p.1608). There is also a hypothesis that microvesicles from normal cells can fuse with cancer cells and act as a camouflage. According to one study, Mucin 1 (MUC1), a cancer-cell marker, and glycoprotein IIIa, a protein only expressed in platelets, are found expressed simultaneously in some exosomes (Muralidharan-Chari et al., 2010, p.1608). However, there is still no solid evidence for this hypothesis. The expression of the costimulatory molecules in macrophage can also be inhibited by tumor-derived vesicles according to Taylor & Black (1985, p.864) and cause multiple



myeloid cell dysfunctions.

## **2.8 Tumor derived vesicles as a target for immunotherapy**

Immunotherapy for cancer uses the patient's own immune system to fight against the cancer. Some immunotherapy trials try to increase the immune response of patients, while others try to overcome the immunosuppression effect of the tumor (Stephen & Hajjar, 2018, p.16). Since tumor-derived vesicles play an important role in immunosuppression, this has become a possible target for immunotherapy. The PD-1/PD-L1 pathway is one popular target. When PD-L1 binds with PD-1, the function of T-cells is diminished and apoptosis may happen (Stephen & Hajjar, 2018, p.13). Thus, the blockage of the PD-1/ PD-L1 pathway by anti-PD-1 or anti-PD-L1 antibodies could possibly reverse the immune suppression of T cells. Actually, treatments focusing on the PD-1/ PD-L1 pathway have already shown effects on melanoma, renal cell carcinoma and non-small cell lung carcinoma (Taube et al., 2014, p.5064). It is therefore an attractive candidate for this colorectal cancer immunotherapy trial.

## **3. Experiment Design and Methodology**

Peripheral blood specimens and the clinicopathologic characteristics data will be obtained from 40 randomly selected patients with colorectal carcinoma. Mini-size exclusion chromatography (Mini-SEC) will be used to isolate the exosomes from the plasma samples. The bicinchoninic acid (BCA) protein assay will be used to test the overall protein expression of exosomes. The size and the concentration of exosomes will be assessed by using tunable resistive pulse sensing (TRPS). A transmission electron microscope will be used to detect the morphology and size of the exosomes. Western blotting will be used to detect the presence of different kinds of protein markers. Then, the level of protein PD-L1 and PD-1 expression will be measured by flow cytometry. CD8<sup>+</sup> T cells will be taken from healthy donors' buffer. A coincubation experiment of the CD8<sup>+</sup> T cells and the derived exosomes will be done to test the hypothesis that PD-L1 has a suppression effect on the T cells. Anti-PD-1 antibodies will be used to confirm whether this suppression is mediated by the PD-L1 pathway. Mice experiments will be done to test the effect of PD-L1 exosomes on vivo tumor growth (Theodoraki et al., 2018, p.897; Kim et al., 2019, p.2).

### **3.1 Isolate exosomes from blood samples of colorectal carcinoma patients by mini-SEC**

Blood samples from the 40 patients with colorectal carcinoma will be collected and then centrifuged for 10 minutes at 1,000×g to get the plasma samples. The samples should be stored frozen before exosome isolation. To isolate the exosomes from the plasma samples, the frozen samples will be thawed and then centrifuged at 4 °C for 10 min at 2,000×g, then 30 minutes at 14,000×g at 4°C. Next, a 0.22 mm-pore Millipore filter will be used to ultrafilter the samples. The exosomes will then be isolated from the samples by using 1.5 cm×12 cm mini-SEC columns with a bed volume of 10 ml. The columns will be washed by phosphate-buffered saline (PBS) in advance. Five 1 ml fractions will be collected. According to previous studies, the third 1 ml fraction contains a very small number of exosomes, the fourth 1 ml fraction is

rich in exosomes with few other kinds of plasma proteins, and the fifth fraction contains exosomes as well as other plasma proteins, including immunoglobulins and albumin. Thus, we will use the fourth fraction in the following experiments (Theodoraki et al., 2018, p.897; Hong et al., 2016, p.2).

### **3.2 Isolate CD8<sup>+</sup> T cells from the blood samples of healthy donors**

Peripheral blood samples will be taken from healthy donors. The peripheral blood mononuclear cells (PBMCs) will then be isolated by using Ficoll-Hypaque gradients. CD8<sup>+</sup> T cells will next be selected by using CD8 immunobeads (Saze et al., 2013, p.10; Ludwig et al., 2017, p.3).

### **3.3 Characteristics of exosomes**

Several experiments will be done to test the characteristics of the exosomes derived from the fourth fraction of mini-SEC.

A BCA protein assay (Pierce Biotechnology, USA) will be used to measure the overall concentration of the proteins contained in the exosomes from the fourth fraction (Ludwig et al., 2017, p.4).

Tunable resistive pulse sensing (TRPS) will be done to assess the size and the concentration of exosomes from the samples. In order to prevent the binding between the proteins and the pores, an Izon reagent kit (1:1, 200 EV) and 0.03% Tween/PBS will be added to the samples. Each sample will then be placed in the nanopore (NP150) and measured at 43.3 mm stretch at 2 pressure levels, with a voltage of 0.74 V for 4 and 8 mbar. Each short pulsate of current means a blockade from a particle. The software Izon (version 3.2) will be used to collect the data of sizes and concentrations (Ludwig et al., 2017, p.4).

A transmission electron microscope will be used to measure the size of the exosomes and assess their morphology. The exosomes from the fourth fraction of mini-SEC will be placed on a copper grid with chloroform of 0.125% formvar, then stained by 1% uranyl acetate (Ludwig et al., 2017, p.4).

### **3.4 Western blotting**

Western blotting will be used to detect the presence of protein markers on the exosomes. First, the exosomes from the fourth fraction of mini-SEC will be concentrated by using VivaSpin 500. Next, the exosomes will be lysed by using Laemmli sample buffer and then separated on 7–15% SDS/PAGE gels. The samples will later be transferred onto a polyvinylidene flouride (PVDF) membrane for Western Blotting. The membrane will be incubated with specific antibodies at 4°C overnight. Next, the horseradish peroxidase (HRP)-conjugated secondary antibody will be used at room temperature for 1 hour. Enhanced chemiluminescent detection reagents will be used to develop blots. The density of protein bands will be measured by using the software Image (Ludwig et al., 2017, p.4-5; Hong et al., 2016, p.3; Kim et al., 2019, p.3).

The antibodies to be used include:

- **From company Abcam:** PD-L1 (1:400,ab58810), FasL(1:500,ab68338), Fas(1:1000,ab133619), TSG101(1:500,ab30871), CD80 (1:1000,ab134120), CD70 (1:1000,ab96323), TRAIL (1:500,ab2056), CTLA-4(1:5000,ab134090), CD9 (1:500, ab97999), CD34 (1:2,000, ab81289), CD44 (1:1,000, ab41478), CD96 (1:500, ab56653), HSP70 (1:5,000, ab9920);
- **From company R&D:** PD-1(1:500, MAB1086), CLL-1 (1:2,000, AF2946);
- **From company Santa Cruz:** CD39 (1:400, sc-33558), CD63 (1:200, sc-15363), CD73 (1:400, sc-25603), COX-2 (1:500, sc-1745CD81);  
From company Thermo Fisher: CD123 (1:200, PA5-13582).

### 3.5 Flow cytometry

Exosomes are usually small in size, about 30 to 200 nm, and are therefore hard to study by normal flow cytometers. One resolution of this problem would be to combine exosomes with larger beads so that flow cytometry could work well. It is a three-step process: isolate the exosomes from the plasma sample, bind the exosomes with the beads, and stain the exosome-bead complex for flow cytometric analysis. Since the isolation process is described earlier, the following protocol is for the next two steps (Morales-Kastresana & Jones, 2017, p.215-225).

#### 3.5.1 Bind exosomes with the beads

In this step, the exosomes will be bound with magnetic beads so that the complex will be large enough for the flow cytometry to work. In the first step, 10  $\mu$ l streptavidin-coated magnetic beads will be transferred into a 2 ml microfuge tube and washed with 250  $\mu$ l T-TBS buffer gently 3 times by using a magnet. Then, 0.5  $\mu$ g biotin-labeled anti-CD63 monoclonal antibodies (mAb) and T-TBS buffer will be added to get a final volume of 50  $\mu$ l. The exact antibody used should be determined by the result of western blotting in the previous step. Since CD63 has been discovered to be expressed in colorectal carcinoma-derived exosomes, anti-CD63 mAb will be used in this protocol according to Ruiz-López et al. (2018, p.793). The solution will be incubated at room temperature for one hour with agitation so that the beads will bind with the antibody. Next, the unbound antibodies will be washed away by washing this solution with 250  $\mu$ l T-TBS buffer and a magnetic stand three times. The anti-CD63Ab-coupled beads will be resuspended in 50  $\mu$ l of PBS and then co-incubated with 10  $\mu$ g exosomes derived from the fourth fraction of mini-SEC overnight at room temperature. The optimal ratio between exosome protein, anti-CD63 mAb and beads has been determined by titration in a previous study (Theodoraki et al., 2018, p.898; Morales-Kastresana & Jones, 2017, p.219-221).

#### 3.5.2 Stain the bead antibody complex

The bead antibody complex will be co-incubated with anti-PD-L1 phycoerythrin protein (PE) or anti-PD-1 PE for the detection of PD-L1 by flow cytometry. The bead antibody complex with PBS will be first incubated with the Casein blocking buffer and Fc Block for 10 minutes at 4°C. Then, the solution will be co-incubated with 0.5  $\mu$ g of anti-PD-L1 PE or anti-PD-1 PE at room temperature for an hour. The labeled isotope will be used as the control (Theodoraki et al., 2018, p.898; Morales-

Kastresana & Jones, 2017, p.221-222).

### **3.5.3 Detect the PD-L1 and PD1-1 level by flow cytometry**

The sample will be analyzed by a flow cytometer with the software Kaluza 1.0. According to Theodoraki (2018), when at least 2% of the isotope control shows a positive result, the sample is considered “positive” in the relative fluorescence values (RFV) calculation. The mean fluorescence intensity (MFI) of the samples and the isotope controls will also be measured. The  $\Delta$ MFI is defined as subtracting the MFI of the isotope control from the MFI of the sample. The RFV is the percentage of “positive” beads times the  $\Delta$ MFI of the exosomes from the fourth fraction of mini-SEC (Theodoraki et al., 2018, p.898-899).

## **3.6 Functional Assay**

### **3.6.1 Determine the effect of exosomes on the activity of CD8<sup>+</sup> cells**

In order to determine the effect of the PD-1/PD-L1 pathway on the activity of T cells, CD8<sup>+</sup> T cells will be obtained from the plasma samples from healthy donors and isolated by using Ficoll-Hypaque gradients as described earlier. The T cells will be activated by coincubation with anti-CD3/anti-CD-28 antibody-coated beads and interleukin-2 (IL2) at 37°C for 5 hours. Then, the activated T cells will be co-incubated with PD-L1 high exosomes and PD-L1 low exosomes taken from the mini-SEC for 16 hours with PBS added as a control. In order to get the expression levels of protein CD69 of the T cells, the exosomes will be stained with anti-human CD8 antibodies and anti-human CD69 antibodies. The stained cells will then be washed twice with permeabilization buffer and once with PBS/1% fetal bovine serum (FBS). Finally, the expression level of protein CD69 will be assessed by flow cytometry (Hong et al., 2016, p.3; Theodoraki et al., 2018, p.899).

In order to confirm that the effect of the exosomes' coincubation on T cells is mediated by the PD-1/PD-L1 pathway, anti-PD-1 antibody will be used to block the pathway. After the activation of T cells by coincubation with anti-CD3/anti-CD-28 antibody-coated beads and IL2 for 5 hours, anti-PD-1 antibody or isotype control will be added and co-incubated for one hour before the addition of colorectal carcinoma-derived exosomes and antibodies. The expression levels of protein CD69 will be assessed by flow cytometry as described before (Hong et al., 2016, p.3; Theodoraki et al., 2018, p.899; Kim et al., 2019, p.5).

### **3.6.2 Determine the effect of exosomes on the apoptosis of CD8<sup>+</sup> cells**

The CD8<sup>+</sup> cells isolated from plasma samples from healthy donors will be stained by using 2.5  $\mu$ M carboxyfluorescein succinimidyl ester (CFSE) at 37°C in serum-free medium for 15 minutes. The CFSE-labeled CD8<sup>+</sup> T cells will then be activated by coincubation with anti-CD3/anti-CD-28 antibody-coated beads and IL2 at 37°C for 5 hours. Next, the activated T cells will co-incubate with PD-L1 high exosomes and PD-L1 low exosomes derived from the colorectal carcinoma patients. An annexin V-fluorescein isothiocyanate (FITC)/propidium iodide apoptosis kit will be used to assess the apoptosis: the cells will be co-incubated with FITC-conjugated annexin V and propidium iodide in annexin V binding buffer at room temperature for 30 minutes in the dark. Flow cytometry and the software FlowJo will be used to get the results. Anti-PD-1 antibody and isotype control will be used as the control group (Kim et al., 2019, p.11; Ludwig et al., 2017, p.4845; Hong, 2016, p.3).

### 3.7 Mouse experiment

In order to determine the effect of colorectal carcinoma-derived exosomes on tumor growth, mouse experiments will be done. In this proposed experiment, the mice will be 6 weeks old with a weight between 18-20g. Tumors will be implanted into the mice by injecting colorectal carcinoma cell line with 50% Matrigel to the right flanks of the mice. For the experiment group, PD-L1 high exosomes and PD-L1 low exosomes extracted from the fourth fraction of mini-SEC in the previous step will be injected 3 times before and after the tumor implantation. Normal saline will be used as the control group. To estimate the tumor size, calipers will be used to measure the length (L) and width (W) of the tumor dump. The volume of the tumor will be defined as  $(L \times W^2)/2$  (Kim et al., 2019, p.5).

### 3.8 Multiplex immunofluorescence

To further explore the effect of exosomes on vivo tumor growth, the expression levels of gene CD3 and Ki-67 of the mice experiment will be tested by using multiplex immunofluorescence. The Opal™ Solid Tumor Immunology kit will be used to stain the sample with antibodies against CD3 and Ki-67. The CD3 and Ki-67 will then be labeled with the fluorescent marker Opal 520 and Opal 690. Next, the slides will be stained with DAPI so that nuclei in the sample become visible and can be mounted. By using the Opal™ method, one slide will show the result of the staining of two primary antibodies simultaneously. Finally, a Vectra slide scanner will be used to scan the slides under fluorescence (Kim et al., 2019, p.4).

## 4. Hypothesized Result

### 4.1 Clinicopathologic data of patients

The clinicopathologic data of 40 patients who did not receive immunotherapy will be collected. The patients will be divided into two groups: patients who have not received any therapy and patients who have completed therapy. The data collected will include age, gender, disease status (with active disease or no evidence of disease), tumor stage (T1, T2, T3 or T4), nodal status (N0, N1, N2 or N3), distant metastasis (M0 or M1), and UICC stage (I, II, III or IV) (ASCO, n.d.; Rosen & Sapa, 2020).

The tumor stages (T) refer to the size of the primary tumor and how deeply the primary tumor has grown. The larger the number, the bigger the size and the deeper the tumor.

**Table 1.** *The definitions of tumor stages (T) (ASCO, n.d.)*

<b>T1</b>	The tumor has grown into the submucosa, the tissue beneath the lining of the colon.
<b>T2</b>	The tumor has grown into the muscularis propria, a deeper muscle layer of the intestines.
<b>T3</b>	The tumor has grown into the subserosa, the connective tissue underneath the large intestine; or it has grown into the surrounding tissues of the colon or rectum.
<b>T4</b>	The tumor has grown through all the colon; or it has attached to other organs.

The node stages refer to the degree of invasion of the tumor of local or distant lymph nodes. The larger the number, the more advanced the invasion of the lymph nodes.

**Table 2.** *The definitions of node stages (N) (ASCO, n.d.)*

<b>No</b>	The tumor has not spread to regional lymph nodes
<b>N1</b>	The tumor has spread to 1-3 regional nodes
<b>N2</b>	The tumor has spread to 4-6 regional nodes
<b>N3</b>	The tumor has spread to more than 7 regional nodes

The metastasis stages (M) refer to whether there is distant tumor metastasis or not.

**Table 3.** *The definitions of metastasis stages (M) (Rosen & Sapra, 2020)*

<b>Mo</b>	The tumor has not spread to other parts of body
<b>M1</b>	The tumor has spread to other parts of body

UICC stages are more general. The stages are described according to the T M N stages.

**Table 4.** *The definitions of UICC stages (Rosen & Sapra, 2020)*

<b>Stage I</b>	Generally small, localized cancer (T1-T2, No, Mo)
<b>Stage II</b>	Generally large in size, localized cancer (T2-T4, No, Mo)
<b>Stage III</b>	Cancer has invaded regional nodes (T1-T4, N1-N3, Mo)
<b>Stage IV</b>	Cancer has invaded other parts of the body (T1-T4, N1-N3, M1)

#### **4.2 The correlation between exosome protein level and clinicopathologic data**

The overall protein concentration of exosomes derived from the fourth fraction of mini-SEC will be assessed by using the BCA protein assay kit. According to previous studies, there is a positive correlation between total protein concentration and disease activity. Thus, the hypothesis of this study is that the patients with active disease (AD) will show a higher overall protein concentration than patients with no evidence of disease (NED) (Theodoraki et al., 2018, p.899; Ludwig et al., 2017, p.4846).

#### **4.3 The correlation between PD-L1, PD-1 level and the clinicopathologic data**

The PD-L1 and PD-1 expression level of exosomes will be measured by flow cytometry. The RFV of PD-L1 and PD-1 will be compared with the disease activity, tumor stage, nodal status and the UICC stage of the patients. The possible result is that the AD patients will have a higher RFV of PD-L1 than the NED patients. The patients with tumors of higher nodal status stage (N1 or above) will show higher PD-L1 RFV than those who are NO. Also, the patients with III or IV UICC stage disease are expected to show a higher RFV than those who are at UICC stage I or II. A similar result is hypothesized for the T stages: patients with tumors of III or IV T stage will show higher PD-L1 expression. This hypothesis is based on the results of Theodoraki (2018, p.899-900).

To further investigate the relationship between exosomes PD-L1 level and the clinicopathologic data, the patients will be divided into 4 groups: AD patients with high UICC stage, NED patients with high UICC stage, AD patients with low UICC stage, and NED patients with low UICC stage. The expected result is that the AD patients with high UICC stage will have the highest mean RFV. The RFV levels of AD patients with low UICC stages and the NED patients with high UICC stages are hypothesized to be similar. The NED patients with I or II UICC stage will show the lowest RFV level (Theodoraki et al., 2018, p.900).

According to the previous study on head and neck cancer by Theodoraki (2018, p.900), the RFV value of PD-1 does not have a correlation with the clinicopathologic data. Thus, the expected result of the PD-1 RFV level is that all the patients will show a high PD-1 expression with little difference.

#### **4.4 The effect of colorectal carcinoma-derived exosomes on T cells**

The CD8<sup>+</sup> T cells will be derived from normal healthy people and then activated. The activated T cells will then be co-incubated with PD-L1 high exosomes and PD-L1 low exosomes derived from colorectal carcinoma patients. The expression level of CD69 protein and apoptosis will be assessed by flow cytometry. The expected result is that the T cells co-incubated with the PD-L1 high exosomes from the fourth fraction of mini-SEC will show less expression of CD69 protein than the PD-L1 low exosomes or the PBS control group. The expression level of protein of CD8<sup>+</sup> T cells will recover when the T cells are co-incubated with both the colorectal carcinoma-derived exosomes and the anti-PD-1 antibody (Theodoraki et al., 2018, p.901).

To test the effect of exosomes derived from colorectal carcinoma patients on the apoptosis of T cells, an annexin V-fluorescein isothiocyanate (FITC)/propidium

iodide apoptosis kit will be used. The hypothesized result is that the exosomes derived from colorectal carcinoma patients will induce apoptosis on T cells and the apoptosis can be reversed by co-incubation with anti-PD-1 antibody (Kim et al., 2019, p.8).

#### **4.5 The effect of colorectal carcinoma-derived exosomes on tumor growth**

To determine the effect of exosomes with high PD-L1 level on tumor growth in vivo, mice will be injected with colorectal carcinoma cell line, PD-L1 high exosomes, PD-L1 low exosomes derived from colorectal carcinoma patients, or normal saline solution as a control. The possible result is that the mice with PD-L1 high exosomes injection show a higher tumor growth rate than those with PD-L1 low exosomes or normal saline solution (Kim et al., 2019, p.5-6).

The effects of PD-L1 exosomes derived from colorectal carcinoma patients on the tumor microenvironment will be tested by using multiplex immunofluorescence against the CD3 and Ki-67. The expected result is that the tissue from the mice with PD-L1 high exosomes injection will show a lower CD3 level than that from the ones with PD-L1 low exosomes or normal saline control injections, which would indicate a decrease in the number of tumor-infiltrating T cells. In addition, the Ki-67 level from the tumor is expected to be higher in mice with PD-L1 high exosomes, indicating a high proliferation rate of the tumor. The expectations come from the results of experiments done by Kim et al. (2019, p.6&9).

### **5. Discussion**

This study concerns the expression of PD-1 and PD-L1, an immune checkpoint and its ligand, on exosomes derived from colorectal carcinoma patients. The experiment described can help us understand the relationship between the overall protein level, PD-1 level and PD-L1 level of colorectal carcinoma derived exosomes and the clinicopathologic data of patients. It also investigates the effect of colorectal carcinoma derived exosomes on T cell activity and apoptosis. The effect of PD-L1 exosomes on tumor growth will also be tested in vivo. If all the hypotheses described above are proved to be correct, then we can make the following conclusions: the expression level of protein PD-L1 on colorectal carcinoma derived exosomes has a positive correlation with tumor stages and activity (AD or NED). In addition, the tumor exosomes with PD-L1 can promote tumor growth in vivo. The PD-1/PD-L1 pathway is able to suppress the activity of T cells and induce T cell apoptosis, which helps the tumor escape the tumor-specific immune response. The blockade of the PD-1/PD-L1 pathway could suppress tumor growth and the PD-1/PD-L1 inhibitor could be a possible new path for immunotherapy of colorectal carcinoma.

In previous studies, the expression levels of protein PD-L1 by a variety of cancer cells were assessed (Taube et al., 2014). The effects of PD-L1 exosomes on head and neck cancer as well as non-small-cell lung cancer have been investigated. Theodoraki and his group (2018) proved that in head and neck cancer, there is a correlation between the level of PD-L1 protein in exosomes and the tumor progression. However, soluble PD-L1 did not affect the tumor growth. The exosomes with PD-L1 were also proved to be able to decrease the activity of T cells. Dong Ha



Kim and his group (2019) investigated the effect of PD-L1 in non-small cell lung cancer. They found that Jurkat cells treated by PD-L1 exosomes showed a decrease in the production of INF- $\gamma$  in a dose-dependent pattern. There was no decrease in groups with an anti-PD-L1 neutralizing antibody and PD-L1 gene knockout cells, which further confirms the suppression effect of PD-1/PD-L1 pathway. The team also proved that the PD-L1-containing exosomes can promote lung tumor growth in vivo. However, in vitro, the expression PD-L1 did not show an effect on tumor growth. In their study, the tumor cells in mice treated with PD-L1 exosomes show a decrease in proliferation, which is not consistent with the expectation.

The effect of PD-L1 exosomes on additional tumor types could be a good topic for future studies. Investigation of the effect of soluble PD-L1 on tumor progression is also necessary. The effect of PD-L1 on in vitro and in vivo tumor growth should also be assessed again. Future studies can also focus on the effect of PD-L1 exosomes on immune cells other than T cells, including dendritic cells, monocyte cells and natural killer cells. In addition, the effect and mechanism of molecule PD-L2, another ligand of the immune checkpoint PD-1, on tumor progression is still not clear.

## References

- Andreola, G., Rivoltini, L., Castelli, C., Huber, V., Perego, P., Deho, P., Squarcina, P., Accornero, P., Lozupone, F., Lugini, L., Stringaro, A., Molinari, A., Arancia, G., Gentile, M., Parmiani, G., & Fais, S. (2002). Induction of lymphocyte apoptosis by tumor cell secretion of FasL-bearing microvesicles. *The Journal of Experimental Medicine*, 195(10), 1303–1316. <https://doi.org/10.1084/jem.20011624>
- Baj-Krzyworzeka, M., Szatanek, R., Weglarczyk, K., Baran, J., & Zembala, M. (2007). Tumour-derived microvesicles modulate biological activity of human monocytes. *Immunology Letters*, 113(2), 76–82. <https://doi.org/10.1016/j.imlet.2007.07.014>
- American Society of Clinical Oncology (ASCO). (n.d.). *Colorectal Cancer: Stages*. Cancer. Net. <https://www.cancer.net/cancer-types/colorectal-cancer/stages>
- American Society of Clinical Oncology (ASCO). (n.d.). *Colorectal Cancer: Statistics*. Cancer. Net <https://www.cancer.net/cancer-types/colorectal-cancer/statistics>
- D'Souza-Schorey, C., & Clancy, J. W. (2012). Tumor-derived microvesicles: shedding light on novel microenvironment modulators and prospective cancer biomarkers. *Genes & Development*, 26(12), 1287–1299. <https://doi.org/10.1101/gad.192351.112>
- Henry, B., Ziobro, R., Becker, K. A., Kolesnick, R., & Gulbins, E. (2013). Acid sphingomyelinase. *Handbook of Experimental Pharmacology*, (215), 77–88. [https://doi.org/10.1007/978-3-7091-1368-4\\_4](https://doi.org/10.1007/978-3-7091-1368-4_4)
- Hong, C. S., Funk, S., Muller, L., Boyiadzis, M., & Whiteside, T. L. (2016). Isolation of biologically active and morphologically intact exosomes from plasma of patients with cancer. *Journal of Extracellular Vesicles*, 5, 29289. <https://doi.org/10.3402/jev.v5.29289>

- Huber, V., Fais, S., Iero, M., Lugini, L., Canese, P., Squarcina, P., Zaccheddu, A., Colone, M., Arancia, G., Gentile, M., Seregini, E., Valenti, R., Ballabio, G., Belli, F., Leo, E., Parmiani, G., & Rivoltini, L. (2005). Human colorectal cancer cells induce T-cell death through release of proapoptotic microvesicles: role in immune escape. *Gastroenterology*, 128(7), 1796–1804. <https://doi.org/10.1053/j.gastro.2005.03.045>
- Iwai, Y., Ishida, M., Tanaka, Y., Okazaki, T., Honjo, T., & Minato, N. (2002). Involvement of PD-L1 on tumor cells in the escape from host immune system and tumor immunotherapy by PD-L1 blockade. *Proceedings of the National Academy of Sciences of the United States of America*, 99(19), 12293–12297. <https://doi.org/10.1073/pnas.192461099>
- Kim, D. H., Kim, H., Choi, Y. J., Kim, S. Y., Lee, J. E., Sung, K. J., Sung, Y. H., Pack, C. G., Jung, M. K., Han, B., Kim, K., Kim, W. S., Nam, S. J., Choi, C. M., Yun, M., Lee, J. C., & Rho, J. K. (2019). Exosomal PD-L1 promotes tumor growth through immune escape in non-small-cell lung cancer. *Experimental & Molecular Medicine*, 51(8), 1–13. <https://doi.org/10.1038/s12276-019-0295-2>
- Lee, C. W., Ren, Y. J., Marella, M., Wang, M., Hartke, J., & Couto, S. S. (2020). Multiplex immunofluorescence staining and image analysis assay for diffuse large B cell lymphoma. *Journal of Immunological Methods*, 478, 112714. <https://doi.org/10.1016/j.jim.2019.112714>
- Ludwig, S., Floros, T., Theodoraki, M. N., Hong, C. S., Jackson, E. K., Lang, S., & Whiteside, T. L. (2017). Suppression of Lymphocyte Functions by Plasma Exosomes Correlates with Disease Activity in Patients with Head and Neck Cancer. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 23(16), 4843–4854. <https://doi.org/10.1158/1078-0432.CCR-16-2819>
- Morales-Kastresana, A., & Jones, J. C. (2017). Flow Cytometric Analysis of Extracellular Vesicles. *Methods in Molecular Biology (Clifton, N.J.)*, 1545, 215–225. [https://doi.org/10.1007/978-1-4939-6728-5\\_16](https://doi.org/10.1007/978-1-4939-6728-5_16)
- Muralidharan-Chari, V., Clancy, J. W., Sedgwick, A., & D'Souza-Schorey, C. (2010). Microvesicles: mediators of extracellular communication during cancer progression. *Journal of Cell Science*, 123(Pt 10), 1603–1611. <https://doi.org/10.1242/jcs.064386>
- Noman, M. Z., Desantis, G., Janji, B., Hasmim, M., Karray, S., Dessen, P., Bronte, V., & Chouaib, S. (2014). PD-L1 is a novel direct target of HIF-1 $\alpha$ , and its blockade under hypoxia enhanced MDSC-mediated T cell activation. *The Journal of Experimental Medicine*, 211(5), 781–790. <https://doi.org/10.1084/jem.20131916>
- Pap E. (2011). The role of microvesicles in malignancies. *Advances in Experimental Medicine and Biology*, 714, 183–199. [https://doi.org/10.1007/978-94-007-0782-5\\_10](https://doi.org/10.1007/978-94-007-0782-5_10)
- Peinado, H., Lavotshkin, S., & Lyden, D. (2011). The secreted factors responsible for pre-metastatic niche formation: old sayings and new thoughts. *Seminars in Cancer Biology*, 21(2), 139–146. <https://doi.org/10.1016/j.semcancer.2011.01.002>
- Rajan, A., & Gulley, J. L. (2014). Nivolumab (anti-PD-1, BMS-936558, ONO-4538) in patients with advanced non-small cell lung cancer. *Translational Lung Cancer Research*, 3(6), 403–405. <https://doi.org/10.3978/j.issn.2218-6751.2014.09.02>
- Rosen RD, & Sapra A. TNM Classification. (2020). *Treasure Island: StatPearls*. <https://www.ncbi.nlm.nih.gov/books/NBK553187/>

- Ruiz-López, L., Blancas, I., Garrido, J. M., Mut-Salud, N., Moya-Jódar, M., Osuna, A., & Rodríguez-Serrano, F. (2018). The role of exosomes on colorectal cancer: A review. *Journal of Gastroenterology and Hepatology*, 33(4), 792–799. <https://doi.org/10.1111/jgh.14049>
- Saze, Z., Schuler, P. J., Hong, C. S., Cheng, D., Jackson, E. K., & Whiteside, T. L. (2013). Adenosine production by human B cells and B cell-mediated suppression of activated T cells. *Blood*, 122(1), 9–18. <https://doi.org/10.1182/blood-2013-02-482406>
- Schorey, J. S., & Bhatnagar, S. (2008). Exosome function: from tumor immunology to pathogen biology. *Traffic (Copenhagen, Denmark)*, 9(6), 871–881. <https://doi.org/10.1111/j.1600-0854.2008.00734.x>
- Siegel, R. L., Miller, K. D., & Jemal, A. (2020). Cancer statistics, 2020. *CA: a cancer journal for clinicians*, 70(1), 7–30. <https://doi.org/10.3322/caac.21590>
- Ståhl, A. L., Johansson, K., Mossberg, M., Kahn, R., & Karpman, D. (2019). Exosomes and microvesicles in normal physiology, pathophysiology, and renal diseases. *Pediatric nephrology (Berlin, Germany)*, 34(1), 11–30. <https://doi.org/10.1007/s00467-017-3816-z>
- Stephen, B., & Hajjar, J. (2018). Overview of Basic Immunology and Translational Relevance for Clinical Investigators. *Advances in experimental medicine and biology*, 995, 1–41. [https://doi.org/10.1007/978-3-030-02505-2\\_1](https://doi.org/10.1007/978-3-030-02505-2_1)
- Taniguchi, M., & Okazaki, T. (2014). The role of sphingomyelin and sphingomyelin synthases in cell death, proliferation and migration-from cell and animal models to human disorders. *Biochimica et biophysica acta*, 1841(5), 692–703. <https://doi.org/10.1016/j.bbali.2013.12.003>
- Taube, J. M., Klein, A., Brahmer, J. R., Xu, H., Pan, X., Kim, J. H., Chen, L., Pardoll, D. M., Topalian, S. L., & Anders, R. A. (2014). Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 20(19), 5064–5074. <https://doi.org/10.1158/1078-0432.CCR-13-3271>
- Taylor, D. D., & Gercel-Taylor, C. (2011). Exosomes/microvesicles: mediators of cancer-associated immunosuppressive microenvironments. *Seminars in Immunopathology*, 33(5), 441–454. <https://doi.org/10.1007/s00281-010-0234-8>
- Taylor, D. D., & Black, P. H. (1985). Inhibition of macrophage Ia antigen expression by shed plasma membrane vesicles from metastatic murine melanoma lines. *Journal of the National Cancer Institute*, 74(4), 859–867. <https://doi.org/10.1093/jnci/74.4.859>
- Tesselaar, M. E., Romijn, F. P., Van Der Linden, I. K., Prins, F. A., Bertina, R. M., & Osanto, S. (2007). Microparticle-associated tissue factor activity: a link between cancer and thrombosis. *Journal of Thrombosis and Haemostasis: JTH*, 5(3), 520–527. <https://doi.org/10.1111/j.1538-7836.2007.02369.x>
- Theodoraki, M. N., Yerneni, S. S., Hoffmann, T. K., Gooding, W. E., & Whiteside, T. L. (2018). Clinical Significance of PD-L1+ Exosomes in Plasma of Head and Neck Cancer Patients. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 24(4), 896–905. <https://doi.org/10.1158/1078-0432.CCR-17-2664>
- Turturici, G., Tinnirello, R., Sconzo, G., & Geraci, F. (2014). Extracellular membrane vesicles as a mechanism of cell-to-cell communication: advantages and disadvantages. *American journal of physiology. Cell physiology*, 306(7), C621–C633. <https://doi.org/10.1152/ajpcell.00228.2013>

- Valcz, G., Galamb, O., Krenács, T., Spisák, S., Kalmár, A., Patai, Á. V., Wichmann, B., Dede, K., Tulassay, Z., & Molnár, B. (2016). Exosomes in colorectal carcinoma formation: ALIX under the magnifying glass. *Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc*, 29(8), 928–938. <https://doi.org/10.1038/modpathol.2016.72>
- Valenti, R., Huber, V., Iero, M., Filipazzi, P., Parmiani, G., & Rivoltini, L. (2007). Tumor-released microvesicles as vehicles of immunosuppression. *Cancer Research*, 67(7), 2912–2915. <https://doi.org/10.1158/0008-5472.CAN-07-0520>
- Whiteside T. L. (2005). Tumour-derived exosomes or microvesicles: another mechanism of tumour escape from the host immune system. *British Journal of Cancer*, 92(2), 209–211. <https://doi.org/10.1038/sj.bjc.6602360>



# Types of Graphs Generated by Pascal's Simplices and a Problem Proposed by Tegua and Godbole

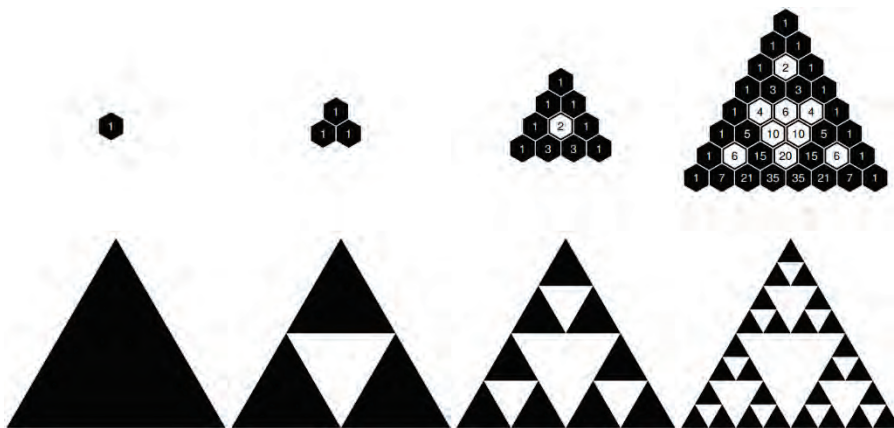
*Zizheng Fang*

**Author Background:** *Zizheng Fang grew up in China and currently attends Hangzhou Foreign Languages School, in Hangzhou, China. His Pioneer research concentration was in the field of mathematics and titled "Combinatorics and Graph Theory."*

## 1. Introduction

Pascal's triangle, also known as the arithmetic triangle, is one of the most well-known mathematical objects. A triangular array of binomial coefficients, it is a favorite of mathematical expositors [2], and has many real-world applications. For example, it has been used to model the Tower of Hanoi game [7], design a cryptosystem [4], and solve PDEs [14]. It has even proved to be useful in physics [1] [8].

One of the most striking properties of Pascal's triangle is that the odd numbers in it form a pattern similar to the Sierpiński gasket, a fractal object obtained from an equilateral triangle by removing smaller triangles from it recursively.



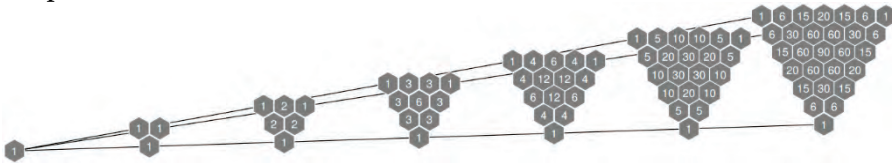
**Figure 1.** Pascal's triangle modulo 2 approximates the Sierpiński gasket.

Other patterns arise when the modulo  $m$  is not 2. When  $m$  is a prime, the pattern is similar to the case where  $m = 2$ . When  $m$  is a prime power, the pattern becomes much more complicated. And when  $m$  has multiple different prime divisors, the pattern is more complicated still, as it is the union of the patterns generated by the prime powers that divide  $m$ . These patterns were considered by Holter et al. [8], who called them the Pascal-Sierpiński gaskets.



**Figure 2.** Numbers not divisible by  $m$  in Pascal's triangle,  $m = 2, 3, 4, 6$ .

It is natural to consider the generalizations of the arithmetic triangle into the  $t$ -dimensional Euclidean space. For example, the 3-dimensional generalization is called Pascal's pyramid [16], in which trinomial coefficients are placed on a tetrahedral lattice.



**Figure 3.** Pascal's pyramid.

In geometry, the  $t$ -dimensional generalization of the triangle is called the  $t$ -simplex. For example, a 2-simplex is a triangle, and a 3-simplex is a tetrahedron. Similarly, the  $t$ -dimensional generalization of Pascal's triangle is called Pascal's simplex [19]. It is rather unsurprising that the numbers not divisible by  $m$  in Pascal's simplex also yield fractal-like patterns [15].

In this paper, we consider the graphs  $A_t(m, n)$  generated by Pascal's  $t$ -simplex modulo  $m$ . The vertices of  $A_t(m, n)$  are the multinomial coefficients in the first  $n + 1$  components (when  $n = 2$ , a component is a row) of Pascal's  $t$ -simplex that are not divisible by  $m$ , and edges are drawn between geometrically adjacent vertices. For prime  $p$  and positive integer  $k$ ,  $A_{t,p}^k$  denotes  $A_t(p, p^k - 1)$ . In Sections 2 to 4, we present a rigorous

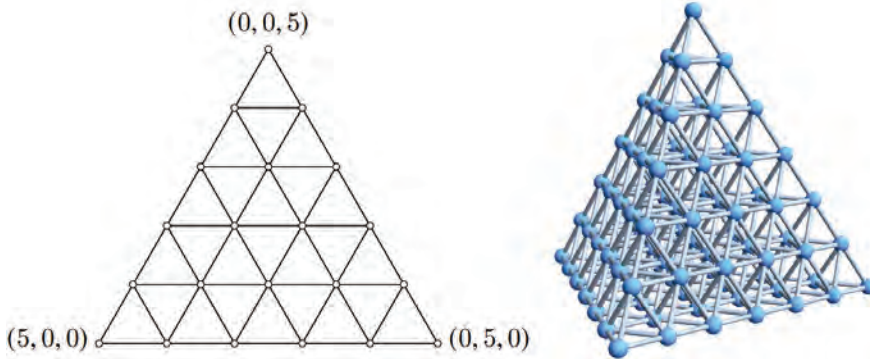
definition of the graphs  $A_t(m, n)$ , and discuss some of their properties, such as symmetry, self-similarity and the number of their vertices and edges.

Hinz [7] considered the graphs  $A_{2_2}^k$ , and proved that it is isomorphic to  $TH_k$ , the Hanoi graph with  $k$  discs. The vertices of  $TH_k$  are the legal states of the discs in the Tower of Hanoi, with two vertices adjacent if one can be obtained from the other with a legal move. Klavžar and Milutinović [12] proposed the Sierpiński graphs  $S_n^k$ , which were motivated by a topological construction called the Lipscomb space [13], and proved that  $S_3^k$  is isomorphic to  $TH_k$  (and therefore isomorphic to  $A_{2_2}^k$ ). In Section 5, we prove the more general statement that the graph  $A_{t_2}^k$  is isomorphic to the Sierpiński graph  $S_{t+1}^k$ . In Sections 6 and 7, we investigate the diameters and Hamiltonicity of  $A_t(m, n)$ . Tegua and Godbole [17] studied the Sierpiński gasket graphs  $S_k$  (not to be confused with the Sierpiński graphs  $S_n^k$ ), which are generated by the iterative process shown in Figure 1 that defines the Sierpiński gasket. They also proposed some open problems. One such problem concerns the properties of a generalization of the Sierpiński gasket graph generated by the numbers not divisible by  $p$  in Pascal's triangle, where  $p$  is a prime other than 2. In Section 8, we make use of our previous results to answer this problem. In the Appendix, we present an algorithm to find Hamiltonian cycles of the graphs  $A_{t_p}^k$ .

## 2. Definitions and basic properties

In Pascal's simplex, the multinomial coefficients are placed on a simplex grid. Therefore, it is convenient to first introduce the simplex grid graphs, a natural generalization of the triangular grid graphs [18], and then define the graphs  $A_t(m, n)$  as their induced subgraphs.

**Definition 2.1.** The  $t$ -simplex grid graph is denoted by  $G_t(n)$ . We assume that  $t \geq 2$  and  $n \geq 1$  unless otherwise stated. Its vertices are the ordered tuples of nonnegative integers  $(x_0, x_1, \dots, x_t)$  that sum to  $n$ , and two vertices  $(x_0, x_1, \dots, x_t)$  and  $(y_0, y_1, \dots, y_t)$  are adjacent if  $\frac{1}{2} \sum_{i=0}^t |x_i - y_i| = 1$ .



**Figure 4.** Simplex grid graphs  $G_2(5)$  and  $G_3(6)$ .

*Remark.* Note that  $\frac{1}{2} \sum_{i=0}^t |x_i - y_i| \equiv \frac{1}{2} \sum_{i=0}^t (x_i - y_i) = 0 \pmod{2}$ , so  $\frac{1}{2} \sum_{i=0}^t |x_i - y_i|$  is always a nonnegative integer, and 1 is therefore its least possible nonzero value.

**Proposition 2.1.** The condition  $\frac{1}{2} \sum_{i=0}^t |x_i - y_i| = 1$  holds if and only if there exist integers  $a, b \in \{0, 1, \dots, t\}$ ,  $a \neq b$  such that  $x_a = y_a + 1$ ,  $y_b = x_b + 1$ , and  $x_c = y_c$  for  $c \in \{0, 1, \dots, t\} - \{a, b\}$ .

*Proof.* First observe that for  $a \in \{0, 1, \dots, t\}$ ,

$$\begin{aligned} |x_a - y_a| &= \frac{1}{2} (|x_a - y_a| + |(n - x_a) - (n - y_a)|) \\ &= \frac{1}{2} \left( |x_a - y_a| + \left| \sum_{i \in \{0, 1, \dots, t\} - \{a\}} x_i - \sum_{i \in \{0, 1, \dots, t\} - \{a\}} y_i \right| \right) \\ &\leq \frac{1}{2} \sum_{i=0}^t |x_i - y_i| = 1. \end{aligned}$$

Therefore, there exist integers  $a, b \in \{0, 1, \dots, t\}$ ,  $a \neq b$  such that  $|x_a - y_a| = |x_b - y_b| = 1$  and  $x_c = y_c$  for  $c \in \{0, 1, \dots, t\} - \{a, b\}$ . Now observe that  $(x_a - y_a) + (x_b - y_b) = \sum_{i=0}^t (x_i - y_i) = n - n = 0$ , so either  $x_a = y_a + 1$ ,  $y_b = x_b + 1$  or  $y_a = x_a + 1$ ,  $x_b = y_b + 1$ . Either way, the proposition holds. ■



Recall the definition of the sign function that  $\text{sgn}(x) = 0$  when  $x = 0$  and  $\text{sgn}(x) = 1$  when  $x > 0$ . In the following proposition, we represent the degree of a vertex in  $G_t(n)$  using the sign function. Note that such a representation may seem complicated, but will make it easier for us to find the number of edges in  $G_t(n)$  in the proof of Proposition 4.1.

**Proposition 2.2.** The degree of vertex  $(x_0, x_1, \dots, x_t)$  in  $G_t(n)$  is

$$\deg((x_0, x_1, \dots, x_t)) = t \sum_{i=0}^t \text{sgn}(x_i).$$

*Proof.* The neighbors of  $(x_0, x_1, \dots, x_t)$  are the ordered tuples  $(y_0, y_1, \dots, y_t)$  of integers satisfying two conditions: 1)  $y_i$  is nonnegative for  $0 \leq i \leq t$  and 2) there exist integers  $a, b \in \{0, 1, \dots, t\}$ ,  $a \neq b$  such that  $x_a = y_a + 1$ ,  $y_b = x_b + 1$ , and  $x_c = y_c$  for  $c \in \{0, 1, \dots, t\} - \{a, b\}$ . Now for every  $a, b \in \{0, 1, \dots, t\}$ ,  $a \neq b$ , there is exactly one such tuple if  $x_a > 0$  and there is none if  $x_a = 0$ . Therefore, for every  $0 \leq i \leq t$  satisfying  $x_i > 0$ , there are  $t$  such tuples. ■

Now we can see that in  $G_t(n)$ , the vertices with the least degrees are the ones containing the most zeros. Geometrically, they are the corners of a simplex.

**Definition 2.2.** The corners of  $G_t(n)$  are  $c_0 = (n, 0, \dots, 0)$ ,  $c_1 = (0, n, \dots, 0)$ , ...,  $c_t = (0, 0, \dots, n)$ .

Now we present a rigorous definition of the graphs  $A_t(m, n)$ . The multinomial coefficient  $\binom{x_1+x_2+\dots+x_t}{x_1, x_2, \dots, x_t}$  is equal to  $\frac{(x_1+x_2+\dots+x_t)!}{x_1!x_2!\dots x_t!}$ . It generalizes the binomial coefficient, which is its special case when  $t = 2$ .

**Definition 2.3.** The graph  $A_t(m, n)$  with integer  $m \geq 2$  is the subgraph of  $G_t(n)$  induced by the set of every vertex  $(x_0, x_1, \dots, x_t)$  satisfying  $m \nmid \binom{x_1+x_2+\dots+x_t}{x_1, x_2, \dots, x_t}$ . A corner of  $G_t(n)$  is also a corner of  $A_t(m, n)$ .

**Proposition 2.3.** Every corner of  $G_t(n)$  is in  $A_t(m, n)$ .

*Proof.* If  $(x_0, x_1, \dots, x_t) = c_0$ , then  $\binom{x_1+x_2+\dots+x_t}{x_1, x_2, \dots, x_t} = \frac{0!}{(0!)^t} = 1$ . If  $(x_0, x_1, \dots, x_t) =$

$c_i$  where  $i \geq 1$ , then  $\binom{x_1+x_2+\dots+x_t}{x_1, x_2, \dots, x_t} = \frac{n!}{n!(0!)^{t-1}} = 1$ . Either way, it is not divisible by  $m$ . ■

**Proposition 2.4.** If  $(x_0, x_1, \dots, x_t) \neq c_0$  is in  $A_t(m, n)$ , then it is adjacent to some  $(y_0, y_1, \dots, y_t)$  where  $y_0 = x_0 + 1$ .

*Proof.* Let  $y_{ij}$  denote  $\begin{cases} x_i & i \neq j \\ x_i - 1 & i = j \end{cases}$ . Notice that  $\binom{x_1+x_2+\dots+x_t}{x_1, x_2, \dots, x_t}$  is equal to  $\sum_{1 \leq i \leq t, x_i > 0} \binom{y_{i1}+y_{i2}+\dots+y_{it}}{y_{i1}, y_{i2}, \dots, y_{it}}$ , so if  $m \nmid \binom{x_1+x_2+\dots+x_t}{x_1, x_2, \dots, x_t}$ , then there exists  $1 \leq i \leq t$  such that  $m \nmid \binom{y_{i1}+y_{i2}+\dots+y_{it}}{y_{i1}, y_{i2}, \dots, y_{it}}$ . Consequently,  $(y_0, y_1, \dots, y_t)$ , where  $y_0 = x_0 + 1$ ,  $y_i = x_i - 1$ , and  $y_c = x_c$  for  $c \in \{0, 1, \dots, t\} - \{0, i\}$ , is also in  $A_t(m, n)$ . By definition  $(x_0, x_1, \dots, x_t)$  is adjacent to  $(y_0, y_1, \dots, y_t)$ . ■

We can apply this proposition to get an  $x - c_0$  path for every vertex  $x = (x_0, x_1, \dots, x_t) \neq c_0$  in  $A_t(m, n)$ . It follows immediately that the graph is connected.

**Corollary**  $A_t(m, n)$  is connected for all  $t$ ,  $m$  and  $n$ .

Figure 2 suggests that the pattern is the most regular when  $m$  is prime and  $n$  is one less than a power of  $m$ . We give a special notation for this type of graphs.

**Definition 2.4.** For prime  $p$  and positive integer  $k$ ,  $A_{t,p}^k$  denotes  $A_t(p, p^k - 1)$ .

**Proposition 2.5.** The graph  $A_{t,p}^1$  is isomorphic to  $G_t(p - 1)$ .

*Proof.* For any vertex  $(x_0, x_1, \dots, x_t)$  in  $G_t(p - 1)$ ,  $x_1 + x_2 + \dots + x_t \leq p - 1$ , so the factor  $p$  will not appear in the numerator of  $\binom{x_1+x_2+\dots+x_t}{x_1, x_2, \dots, x_t} = \frac{(x_1+x_2+\dots+x_t)!}{x_1!x_2!\dots x_t!}$ , so the vertex is in  $A_{t,p}^1$ . Therefore,  $A_{t,p}^1$  is the subgraph of  $G_t(p - 1)$  induced by its own vertex set, so it is isomorphic to  $G_t(p - 1)$ . ■

**Definition 5.** We omit  $t$  when it is equal to 2. Therefore,  $G(n)$ ,  $A(m, n)$ , and  $A_p^k$  denote  $G_2(n)$ ,  $A_2(m, n)$ , and  $A_{2p}^k$  respectively.

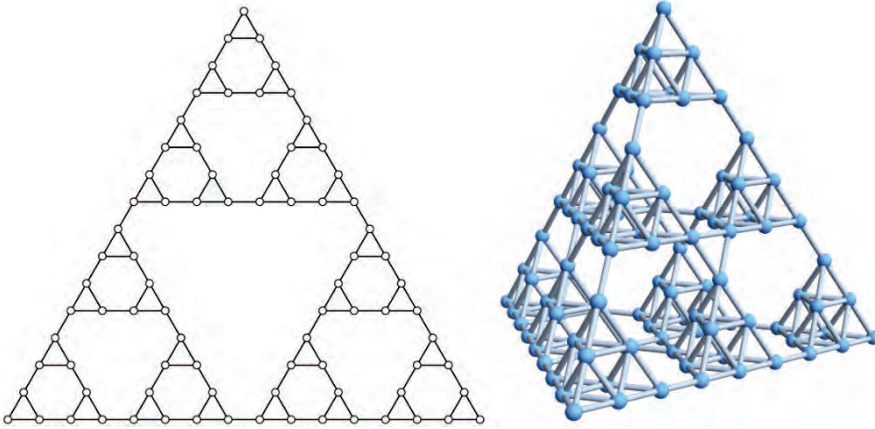


Figure 5. Graphs  $A_{2_2}^4$  and  $A_{3_3}^2$ .

### 3. Symmetry and self-similarity

Since the inclusion of vertices in  $A_t(m, n)$  depends on the divisibility of a multinomial coefficient by a prime number, it is useful to introduce a relevant result by Dickson [6].

**Theorem 3.1.** For nonnegative integers  $x_0 + x_1 + \dots + x_t = s$ ,  $\binom{s}{x_0, x_1, \dots, x_t}$  is not divisible by prime  $p$  if and only if  $s_j = \sum_{i=0}^t x_{i+j}$  for  $0 \leq j < r$ , where  $s = \sum_{j=0}^r s_j p^j$ ,  $x_i = \sum_{j=0}^r x_{i+j} p^j$  and  $0 \leq s_j, x_{i+j} \leq p-1$  for  $0 \leq j < r$ ,  $0 \leq i \leq t$ .

Applying Theorem 3.1 to  $A_{tp}^k$ , we obtain the following proposition.

**Proposition 3.2.** Vertex  $(x_0, x_1, \dots, x_t)$  of  $G_t(p^k - 1)$  is in  $A_{tp}^k$  if and only if  $\sum_{i=0}^t x_{i+j} = p-1$  for  $0 \leq j < k-1$ , where  $x_i = \sum_{j=0}^{k-1} x_{i+j} p^j$ , and  $0 \leq x_{i+j} \leq p-1$  for  $0 \leq j < k-1$ ,  $0 \leq i \leq t$ .

*Proof.* Notice that  $\binom{p^k-1}{x_0, x_1, x_2, \dots, x_t} = \binom{p^k-1}{x_0} \binom{x_1+x_2+\dots+x_t}{x_1, x_2, \dots, x_t}$ . According to Theorem 3.1,  $p \nmid \binom{p^k-1}{x_0}$ . Therefore, vertex  $(x_0, x_1, \dots, x_t)$  of  $G_t(p^k - 1)$  is in  $A_{tp}^k$  if

and only if  $p \nmid \binom{p^k-1}{x_0, x_1, x_2, \dots, x_t}$ . ■

We can see from Figure 5 that the visual representations of  $A_{t_p}^k$  are highly symmetric. The following proposition shows this fact by naming all the automorphisms of  $A_{t_p}^k$ , which are isomorphisms from  $A_{t_p}^k$  to itself. (We refrain from applying the word “symmetric” to the graphs  $A_{t_p}^k$  so as not to cause confusion with the type of graphs called symmetric graphs.)

**Proposition 3.3.** By changing the order of  $x_0, x_1, \dots, x_t$ , we can obtain  $(t + 1)!$  automorphisms of  $A_{t_p}^k$ . No other automorphisms exist.

*Proof.* Let a permutation of  $0, 1, \dots, t$  be  $\alpha$ . There are  $(t + 1)!$  such permutations, each of which corresponds to an isomorphism

$$f((x_0, x_1, \dots, x_t)) = (x_{\alpha(0)}, x_{\alpha(1)}, \dots, x_{\alpha(t)}).$$

Now since the corners are the vertices with the least possible degrees, an isomorphism must map every corner to a corner. Therefore, the number of isomorphisms is no more than the number of permutations of corners, which is  $(t + 1)!$ . This completes the proof. ■

The following corollary is useful when we prove the Hamiltonicity of  $A_{t_p}^k$  in Section 7.

**Corollary** For any two pairs of corners  $(c_i, c_j)$  and  $(c_{i'}, c_{j'})$  such that  $i \neq j$  and  $i' \neq j'$ , there is an isomorphism that maps  $c_i$  to  $c_{i'}$  and  $c_j$  to  $c_{j'}$ .

Self-similarity is what defines a fractal. The Sierpiński triangle is a fractal because it can be divided into three smaller triangles, each of which is similar to the original one. The drawings of  $A_2^k$  approximate the Sierpiński triangle, and the ( $t$ -dimensional) geometrical representations of the graphs  $A_{t_p}^k$  also show similar structures. Therefore, it is reasonable to suggest that, although a finite, discrete object cannot be partitioned into parts that are identical to itself,  $A_{t_p}^k$  can still possess some form of self-similarity. In the following definition, we propose a partition of the vertex set of  $A_{t_p}^k$  such that

every part of such a partition induces a subgraph that is isomorphic to  $A_{t_p}^{k-1}$ .

Such a partition can prove useful when investigating other problems, as we will show in Sections 4, 6 and 7.

**Definition 3.1.** For nonnegative integers  $x'_0 + x'_1 + \dots + x'_t = p - 1$ , a block  $B_{x'_0, x'_1, \dots, x'_t}$  of  $P = A_{t_p}^k$  is the subgraph of  $P$  induced by the set of every vertex  $(x_0, x_1, \dots, x_t)$  such that  $\lfloor \frac{x_i}{p^{k-1}} \rfloor = x'_i$  for  $0 \leq i \leq t$ .

**Proposition 3.4.** Every vertex in  $A_{t_p}^k$  is in exactly one block, and every block of  $A_{t_p}^k$  is isomorphic to  $A_{t_p}^{k-1}$ .

*Proof.* According to Proposition 3.2, vertex  $(x_0, x_1, \dots, x_t)$  of  $G_t(p^k - 1)$  is in  $A_{t_p}^k$  if and only if  $\sum_{i=0}^t \lfloor \frac{x_i}{p^{k-1}} \rfloor = p - 1$  and  $(x_0 \bmod p^{k-1}, x_1 \bmod p^{k-1}, \dots, x_t \bmod p^{k-1})$  is in  $A_{t_p}^{k-1}$ .

By uniqueness of the division algorithm, no vertex in  $A_{t_p}^k$  can be in two different blocks. And the first condition implies that every vertex  $(x_0, x_1, \dots, x_t)$  in  $A_{t_p}^k$  is in block  $B_{\lfloor \frac{x_0}{p^{k-1}} \rfloor, \lfloor \frac{x_1}{p^{k-1}} \rfloor, \dots, \lfloor \frac{x_t}{p^{k-1}} \rfloor}$ .

By the second condition, the vertex set of  $B_{x'_0, x'_1, \dots, x'_t}$  is  $\{(x'_0 p^{k-1} + r_0, x'_1 p^{k-1} + r_1, \dots, x'_t p^{k-1} + r_t) \mid (r_0, r_1, \dots, r_t) \in A_{t_p}^1\}$ . Also notice that  $(x'_0 p^{k-1} + r_0, x'_1 p^{k-1} + r_1, \dots, x'_t p^{k-1} + r_t)$  and  $(x'_0 p^{k-1} + s_0, x'_1 p^{k-1} + s_1, \dots, x'_t p^{k-1} + s_t)$  are adjacent if and only if  $(r_0, r_1, \dots, r_t)$  and  $(s_0, s_1, \dots, s_t)$  are adjacent. Therefore,  $B_{x'_0, x'_1, \dots, x'_t}$  is isomorphic to  $B_{p-1, 0, \dots, 0}$ , which is isomorphic to  $A_{t_p}^{k-1}$ . ■

**Definition 3.2.** For  $G = A_{t_p}^k$ , two different blocks  $B_1$  and  $B_2$  are b-adjacent when there exists edge  $uv \in G$  such that  $u \in B_1$  and  $v \in B_2$ . Such an edge is called a *bridge* between  $B_1$  and  $B_2$ .

**Proposition 3.5.** Two different blocks  $B_{x'_0, x'_1, \dots, x'_t}$  and  $B_{y'_0, y'_1, \dots, y'_t}$  are b-adjacent if and only if  $\frac{1}{2} \sum_{i=0}^t |x'_i - y'_i| = 1$ . There is exactly one bridge between two such blocks, and its endpoints are corners of each block.

*Proof.* Suppose that there exist adjacent vertices  $(x_0, x_1, \dots, x_t) \in B_{x'_0, x'_1, \dots, x'_t}$  and  $(y_0, y_1, \dots, y_t) \in B_{y'_0, y'_1, \dots, y'_t}$ . Therefore, there exists integers  $a, b \in \{0, 1, \dots, t\}$ ,  $a \neq b$  such that  $x_a = y_a + 1$ ,  $y_b = x_b + 1$ , and  $x_c = y_c$  for  $c \in \{0, 1, \dots, t\} - \{a, b\}$ . Therefore,  $\frac{1}{2} \sum_{i=0}^t |x'_i - y'_i| = \frac{1}{2} \sum_{i=0}^t \left| \left\lfloor \frac{x_i}{p^{k-1}} \right\rfloor - \left\lfloor \frac{y_i}{p^{k-1}} \right\rfloor \right| \leq \frac{1}{2} \sum_{i=0}^t |x_i - y_i| = 1$ . Since  $B_{x'_0, x'_1, \dots, x'_t}$  and  $B_{y'_0, y'_1, \dots, y'_t}$  are different blocks, they can only be b-adjacent.

Now suppose that  $B_{x'_0, x'_1, \dots, x'_t}$  and  $B_{y'_0, y'_1, \dots, y'_t}$  are b-adjacent blocks of  $A_{tp}^k$ .

Therefore, there exists integers  $a, b \in \{0, 1, \dots, t\}$ ,  $a \neq b$  such that  $x'_a = y'_a + 1$ ,  $y'_b = x'_b + 1$ , and  $x'_c = y'_c$  for  $c \in \{0, 1, \dots, t\} - \{a, b\}$ . Therefore, for vertices  $(x_0, x_1, \dots, x_t) \in B_{x'_0, x'_1, \dots, x'_t}$  and  $(y_0, y_1, \dots, y_t) \in B_{y'_0, y'_1, \dots, y'_t}$ ,  $x_a > y_a$ ,  $y_b > x_b$ . Suppose they are adjacent. Then  $x_a = y_a + 1$ ,  $y_b = x_b + 1$ . Now  $(y'_a + 1)p^{k-1} + (x_a \bmod p^{k-1}) = y'_a p^{k-1} + (y_a \bmod p^{k-1}) + 1$ , so  $x_a \bmod p^{k-1} = 0$ ,  $y_a \bmod p^{k-1} = p^{k-1} - 1$ . Similarly,  $y_b \bmod p^{k-1} = 0$ ,  $x_b \bmod p^{k-1} = p^{k-1} - 1$ . By Proposition 3.2,  $\sum_{i=0}^t (x_i \bmod p^{k-1}) = \sum_{i=0}^t (y_i \bmod p^{k-1}) = p^{k-1} - 1$ . So  $x_c = x'_c p^{k-1}$  for  $0 \leq c \leq t$ ,  $c \neq b$ ,  $x_b = x'_b p^{k-1} + p^{k-1} - 1$ , and  $y_c = y'_c p^{k-1}$  for  $0 \leq c \leq t$ ,  $c \neq a$ ,  $y_a = y'_a p^{k-1} + p^{k-1} - 1$ . We can see that  $(x_0, x_1, \dots, x_t)$  and  $(y_0, y_1, \dots, y_t)$  are indeed adjacent. Thus, there is exactly one bridge between  $B_{x'_0, x'_1, \dots, x'_t}$  and  $B_{y'_0, y'_1, \dots, y'_t}$ , and its endpoints are corners of each block. ■

**Definition 3.3.** Let the block-graph of  $G = A_{tp}^k$  be the graph  $G'$  whose vertices are the blocks of  $G$ , and two vertices are adjacent if they are b-adjacent.

**Proposition 3.6.** For  $G = A_{tp}^k$ ,  $G'$  is isomorphic to  $G_t(p-1)$ , and thus to  $A_{tp}^1$ .

*Proof.* By Propositions 2.5 and 3.5,  $f(B_{x'_0, x'_1, \dots, x'_t}) = (x'_0, x'_1, \dots, x'_t)$  is an isomorphism from  $G'$  to  $A_{t,p}^1$ . ■

#### 4. Number of vertices and edges

**Proposition 4.1.** The graph  $G_t(n)$  has  $\binom{n+t}{t}$  vertices and  $\frac{t(t+1)}{2} \binom{n+t-1}{t}$  edges.

*Proof.* The number of vertices of  $G_t(n)$  is equal to the number of ways to distribute  $n$  identical objects into  $t + 1$  distinguishable containers, which is well understood to be  $\binom{n+t}{t}$ .

By degree,  $\deg((x_0, x_1, \dots, x_t)) = t \sum_{i=0}^t \text{sgn}(x_i)$ . Therefore,

$$\begin{aligned} |E(G_t(n))| &= \frac{1}{2} \sum_{(x_0, x_1, \dots, x_t) \in G_t(n)} \left( t \sum_{i=0}^t \text{sgn}(x_i) \right) \\ &= \frac{t(t+1)}{2} \sum_{(x_0, x_1, \dots, x_t) \in G_t(n)} \text{sgn}(x_0) \\ &= \frac{t(t+1)}{2} \binom{n+t-1}{t} \end{aligned}$$

The last step of the equation above derives from the fact that the number of vertices  $(x_0, x_1, \dots, x_t)$  in  $G_t(n)$  where  $x_0$  is nonzero is equal to the number of ways to distribute  $n - 1$  identical objects (the missing one is reserved for  $x_0$ ) into  $t + 1$  distinguishable containers. ■

**Proposition 4.2.** The graph  $A_{t,p}^k$  has  $\binom{p+t-1}{t}^k$  vertices and

$$\frac{t(t+1) \binom{p+t-2}{t} \left( \binom{p+t-1}{t} \right)^{k-1}}{2 \left( \binom{p+t-1}{t} - 1 \right)} \text{ edges.}$$

*Proof.* By Proposition 4.1, the proposition is true for  $k = 1$ . Assume that it is true for some  $k$ . Then by Propositions 3.4, 3.6 and 4.1,  $|V(A_{t,p}^{k+1})| =$

$$\binom{p+t-1}{t} |V(A_{t,p}^k)| = \binom{p+t-1}{t}^{k+1}, \quad \text{and} \quad |E(A_{t,p}^{k+1})| = \binom{p+t-1}{t} |E(A_{t,p}^k)| +$$

$\frac{t(t+1)}{2} \binom{p+t-2}{t} = \frac{t(t+1) \binom{p+t-2}{t} \left( \binom{p+t-1}{t}^{k+1} - 1 \right)}{2 \left( \binom{p+t-1}{t} - 1 \right)}$ . By induction, this proposition holds

for every positive integer  $k$ . ■

## 5. Graphs $A_{t_2}^k$ and the Sierpiński graphs

The Sierpiński graphs, introduced by Klavžar and Milutinović in [12], were motivated by a topological construction called the Lipscomb space [13], which is connected to the Sierpiński gasket. It has been proven in [12] that the Sierpiński graph  $S_3^k$  is isomorphic to the Hanoi graph  $TH_k$ , and is therefore isomorphic to  $A_{2_2}^k$ . In this section, we generalize this result to prove the following theorem.

**Theorem 5.1.** The graph  $A_{t_2}^k$  is isomorphic to the Sierpiński graph  $S_{t+1}^k$ .

*Proof.* By Proposition 3.2, vertex  $(x_0, x_1, \dots, x_t)$  of  $G_t(2^k - 1)$  is in  $A_{t_2}^k$  if and only if  $\sum_{i=0}^t x_{i_j} = 1$  for  $0 \leq j \leq k-1$ , where  $x_i = \sum_{j=0}^{k-1} x_{i_j} 2^j$  and  $0 \leq x_{i_j} \leq 1$  for all  $0 \leq j \leq k-1$ ,  $0 \leq i \leq t$ . Therefore, let  $f((x_0, x_1, \dots, x_t)) = (s_1, s_2, \dots, s_k)$ , such that  $x_{s_{d-1}} = 1$  for all  $1 \leq d \leq k$ , and  $f$  is a bijection between  $V(A_{t_2}^k)$  and  $V(S_{t+1}^k)$ . Now we prove that  $f$  is adjacency-preserving.

By Proposition 2.1, two vertices  $x = (x_0, x_1, \dots, x_t)$  and  $y = (y_0, y_1, \dots, y_t)$  of  $A_{t_2}^k$  are adjacent if and only if there exist integers  $a, b \in \{0, 1, \dots, t\}$ ,  $a \neq b$  such that  $x_a = y_a + 1$ ,  $y_b = x_b + 1$ , and  $x_c = y_c$  for  $c \in \{0, 1, \dots, t\} - \{a, b\}$ . Now  $x_a > 0$ , so let  $\delta$  be the least integer such that  $2^\delta \nmid x_a$ , and let  $s = (s_1, s_2, \dots, s_k) = f(x)$  and  $t = (t_1, t_2, \dots, t_k) = f(y)$ . Then  $s_\delta = a$ ,  $t_\delta \neq a$ , and  $s_d \neq a$ ,  $t_d = a$  for all  $1 \leq d \leq \delta - 1$ . Now for any  $c \in \{0, 1, \dots, t\} - \{a, b\}$ ,  $x_c = y_c$ , so for any  $1 \leq d \leq k$ ,  $s_d = c$  if and only if  $t_d = c$ . Therefore,  $t_\delta = b$  and  $s_d = b$  for  $1 \leq d \leq \delta - 1$ , so  $s$  and  $t$  are adjacent in  $S_{t+1}^k$ . Finally, according to [11],  $|E(S_{t+1}^0)| = 0$  and  $|E(S_{t+1}^k)| = (t+1)|E(S_{t+1}^k)| + \binom{t+1}{2}$ ,

and by Proposition 4.2,  $|E(A_{t_2}^k)| = \frac{(t+1)((t+1)^k - 1)}{2}$ , so  $S_{t+1}^k$  and  $A_{t_2}^k$

have the same number of edges, and our proof is complete. ■



## 6. Diameter

Chen and Duh [3] proved that  $\text{diam}(S_n^k) = 2^k - 1$ . In this section, we will extend this result by proving that  $\text{diam}(A(m, n)) = n$  if  $n = rp^{k'} - 1$  ( $1 \leq r \leq p^k - 1$ ) where prime  $p$  divides  $m$ , and that  $\text{diam}(A_{tp}^k) = p^k - 1$ . The latter contains the result in [3] while the former contains the  $n = 3$  case.

**Lemma 6.1.** In the graph  $G_t(n)$ ,  $d(x, y) = \frac{1}{2} \sum_{i=0}^t |x_i - y_i|$ , where  $x = (x_0, x_1, \dots, x_t)$  and  $y = (y_0, y_1, \dots, y_t)$ .

*Proof.* By definition, if  $\frac{1}{2} \sum_{i=0}^t |x_i - y_i| = 1$ , then  $d(x, y) = 1$ . Assume that if

$\frac{1}{2} \sum_{i=0}^t |x_i - y_i| = k$ , then  $d(x, y) = k$ . Then if  $\frac{1}{2} \sum_{i=0}^t |x_i - y_i| = k + 1$ , then

there exists  $a, b \in \{0, 1, \dots, t\}$ ,  $a \neq b$ , such that  $x_a > y_a$ ,  $x_b < y_b$ . Consider  $z = (z_0, z_1, \dots, z_t)$  where  $z_a = x_a - 1$ ,  $z_b = x_b + 1$ ,  $z_c = x_c$  for  $c \in \{0, 1, \dots, t\} - \{a, b\}$ .

Observe that  $\frac{1}{2} \sum_{i=0}^t |z_i - y_i| = \frac{1}{2} \sum_{i=0}^t |x_i - y_i| + \frac{1}{2} (|z_a - y_a| - |x_a - y_a| + |z_b - y_b| - |x_b - y_b|) =$

$$\frac{1}{2} \sum_{i=0}^t |x_i - y_i| + \frac{1}{2} ((z_a - x_a) - (z_b - x_b)) =$$

$(k + 1) - 1 = k$ . By the induction hypothesis  $d(z, y) = k$ . Therefore,  $d(x, y) \leq d(x, z) + d(z, y) = k + 1$ .

Then let  $d = d(x, y)$  and let a shortest path between  $x$  and  $y$  be  $x^0, x^1, \dots, x^d$  where  $x^0 = x$ ,  $x^d = y$ , and  $x^j = (x_0^j, x_1^j, \dots, x_t^j)$  for  $0 \leq j \leq d$ .

Then  $d(x, y) = d = \sum_{j=0}^{d-1} \left( \frac{1}{2} \sum_{i=0}^t |x_i^j - x_i^{j+1}| \right) = \frac{1}{2} \sum_{i=0}^t \sum_{j=0}^{d-1} |x_i^j - x_i^{j+1}| \geq$

$$\frac{1}{2} \sum_{i=0}^t |x_i^0 - x_i^d| = \frac{1}{2} \sum_{i=0}^t |x_i - y_i| = k + 1.$$

Therefore,  $d(x, y) = k + 1$ , and by induction the proof is complete. ■

**Lemma 6.2.** The diameter of  $G_t(n)$  is  $n$ .

*Proof.* By Lemma 6.1, for any two vertices  $x, y \in G_t(n)$ ,

$$d(x, y) = \frac{1}{2} \sum_{i=0}^t |x_i - y_i| \leq \frac{1}{2} \sum_{i=0}^t (x_i + y_i) = n.$$

Also, the distance between two different corners is equal to  $n$ . Therefore,  $\text{diam}(G_t(n)) = n$ . ■

**Theorem 6.3.** The diameter of  $A_{t_p}^k$  is  $p^k - 1$ .

*Proof.* When  $k = 1$ , the theorem is true by Lemma 6.2. Assume that it is true for some  $k$ . Now we consider the case of  $G = A_{t_p}^{k+1}$ .

Let  $x \in B_{x'}$  and  $y \in B_{y'}$  be two vertices. Let  $P = B_0, B_1, \dots, B_l$  where  $B_0 = B_{x'}$  and  $B_l = B_{y'}$  be a shortest path in  $G$  between  $B_{x'}$  and  $B_{y'}$ . By Lemma 6.2 and Proposition 3.6,  $l \leq p - 1$ . Let  $v_0 v_1, v_2 v_3, \dots, v_{2l-2} v_{2l-1}$  be the bridges between consecutive blocks in  $P$ , such that  $v_{2i} \in B_i$  and  $v_{2i+1} \in B_{i+1}$  for  $0 \leq i \leq l - 1$ . Now observe that

$$\begin{aligned} d(x, y) &\leq d(x, v_0) + d(v_0, v_1) + d(v_1, v_2) + d(v_2, v_3) + \dots + d(v_{2l-2}, v_{2l-1}) \\ &\quad + d(v_{2l-1}, y) \\ &= d(x, v_0) + d(v_1, v_2) + \dots + d(v_{2l-1}, y) + l \\ &\leq (l + 1) \text{diam}(A_{t_p}^k) + l \\ &\leq p(p^k - 1) + p - 1 \\ &= p^{k+1} - 1 \end{aligned}$$

Finally, by Lemma 6.2,  $\text{diam}(A_{t_p}^{k+1}) \geq p^{k+1} - 1$ . This completes the proof.

■

Now when  $t = 2$ ,  $A_{t_p}^k = A_p^k$  is planar, and we can exploit its planarity to prove a more general result. We start with a lemma.

**Lemma 6.4.** Of vertices  $(x_0, x_1, x_2 - 1)$ ,  $(x_0, x_1 - 1, x_2)$ ,  $(x_0 - 1, x_1, x_2)$  in  $G(n)$ , if one of them is in  $A(m, n)$ , then at least one of the other two is in  $A(m, n)$ .

*Proof.* Consider  $(x_0, x_1 - 1, x_2)$  and  $(x_0, x_1, x_2 - 1)$ . Since  $\binom{(x_1-1)+x_2}{x_1-1} +$

$\binom{x_1+(x_2-1)}{x_1} = \binom{x_1+x_2}{x_1}$ , if neither of them is in  $A(m, n)$ , then  $(x_0 - 1, x_1, x_2)$  is

not in  $A(m, n)$ . If exactly one of them is in  $A(m, n)$ , then  $(x_0 - 1, x_1, x_2)$  is in  $A(m, n)$ . If both of them are in  $A(m, n)$ , then the lemma holds regardless of whether  $(x_0 - 1, x_1, x_2)$  is in  $A(m, n)$  or not. ■

**Theorem 6.5.** The diameter of  $G = A(m, n)$  is equal to  $n$  if  $(0, n, 0), (0, n - 1, 1), \dots, (0, 0, n) \in G$ .

*Proof.* By Lemma 6.2,  $\text{diam}(A(m, n)) \geq n$ , so we only need to prove that  $d((x_0, x_1, x_2), (y_0, y_1, y_2)) \leq n$  for every  $(x_0, x_1, x_2), (y_0, y_1, y_2) \in G$ . Notice that when  $x_1 = 0$  or  $x_2 = 0$ ,  $(x_0, x_1, x_2) \in G$ , and  $(0, n, 0), (0, n - 1, 1), \dots, (0, 0, n) \in G$ , so every vertex containing 0 is in  $G$ .

Let  $P_{0_x}$  be a maximal path  $(x_0^0, x_1^0, x_2^0), (x_0^1, x_1^1, x_2^1), \dots, (x_0^l, x_1^l, x_2^l)$  constructed with the following rules:

- 1)  $(x_0^0, x_1^0, x_2^0) = (x_0, x_1, x_2)$ ;
- 2)  $x_0^{i+1} = x_0^i + 1$  for  $0 \leq i \leq l - 1$ .

By Lemma 6.4, 2) can be achieved when both  $(x_0^i + 1, x_1^i - 1, x_2^i)$  and  $(x_0^i + 1, x_1^i, x_2^i - 1)$  are in  $G(n)$ . If exactly one of them is in  $G(n)$ , it means that one of  $x_1^i$  or  $x_2^i$  is 0, so the one candidate that is in  $G(n)$  also contains 0, and is therefore in  $G$ . Therefore, the process stops only when neither candidate is in  $G(n)$ , so  $(x_0^l, x_1^l, x_2^l) = c_0$ .

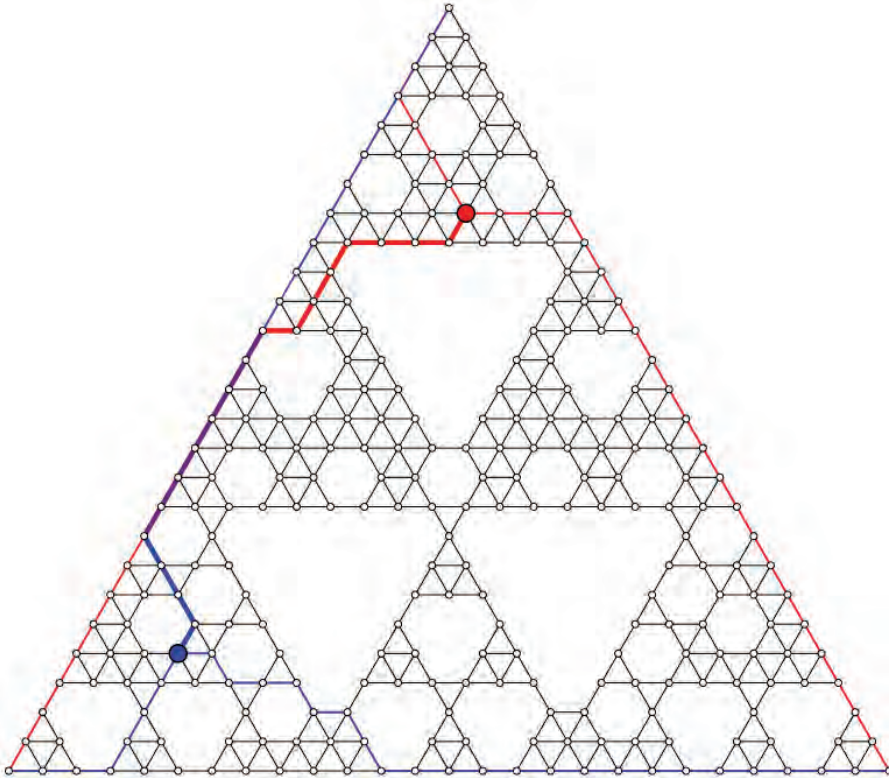
Similarly, we construct paths  $P_{1_x}$  and  $P_{2_x}$ , which end in  $c_1$  and  $c_2$  respectively. And similarly, we start from  $(y_0, y_1, y_2)$  to construct paths  $P_{0_y}$ ,  $P_{1_y}$  and  $P_{2_y}$ .

Since every vertex  $(x_0^i, x_1^i, x_2^i)$  in  $P_{0_x}$  satisfies  $x_0^i = x_0 + i$ , and every vertex  $(x_0^j, x_1^j, x_2^j)$  in  $P_{1_x}$  satisfies  $x_0^j \leq x_0$ ,  $P_{0_x}$  and  $P_{1_x}$  do not intersect except at the common starting vertex. Similarly, no two of  $P_{0_x}$ ,  $P_{1_x}$  and  $P_{2_x}$  intersect except at the common starting vertex. Therefore, these three paths partition  $G$  into three non-overlapping regions, such that every vertex in  $G$  is either in one of these regions or on one of these paths.

The length of any of these paths is less than or equal to  $n$ , so if  $(y_0, y_1, y_2)$  is on one of these paths then the theorem holds. If not, then WLOG let  $(y_0, y_1, y_2)$  be in the region confined by  $P_{1_x}$  and  $P_{2_x}$ . Then  $P_{0_y}$  intersects with one of  $P_{1_x}$  and  $P_{2_x}$  (or with both, if it passes through  $(x_0, x_1, x_2)$ ). WLOG let  $P_{0_y}$  intersect with  $P_{1_x}$ . Let the first vertex in  $P_{0_y}$  that is also in  $P_{1_x}$  be  $(y_0^j, y_1^j, y_2^j) = (x_0^i, x_1^i, x_2^i)$ . Since  $(x_0^i, x_1^i, x_2^i) = (y_0^j, y_1^j, y_2^j) \in G$ ,  $d((x_0, x_1, x_2), (y_0, y_1, y_2)) \leq d((x_0, x_1, x_2), (x_0^i, x_1^i, x_2^i)) + d((y_0^j, y_1^j, y_2^j), (y_0, y_1, y_2)) = i + j \leq x_1 + i + y_0 + j = x_1^i + y_0^j = x_1^i + x_0^i \leq x_0^i + x_1^i + x_2^i = n$ . ■

The process described in the proof of Theorem 6.5 is illustrated by Figure 6. The two arbitrarily chosen vertices and the paths they generate are highlighted with red and blue respectively. Overlapping edges are drawn in

purple. The desired path is marked by the thick edges.



**Figure 6.** Illustration for proof of Theorem 6.5. The graph used here is  $A(6, 26)$ .

Given  $m$ , Proposition 6.7 can be used to find infinitely many values of  $n$  for which  $\text{diam}(A(m, n)) = n$ . Let us first prove the following lemma.

**Lemma 6.6.** For prime  $p$ ,  $\max_{0 \leq x \leq n} v_p \left( \binom{n}{x} \right) = \lfloor \log_p(n+1) \rfloor - v_p(n+1)$ ,

where  $v_p(n)$  denotes the number  $N$  such that  $p^N \parallel n$ .

*Proof.* Let  $n$  be represented as  $rp^{k'} - 1$  where  $p \nmid r$ .

If  $r < p$ , then by Theorem 3.1  $\max_{0 \leq x \leq n} v_p \left( \binom{n}{x} \right) = 0$ , so the lemma holds.

If  $r > p$ , then by Kummer's theorem [6],  $v_p \left( \binom{n}{x} \right)$  is equal to the number of carries that occur when adding  $x$  and  $n - x$  in base- $p$ . Now let  $x'$  be the greatest integer less than  $n$  of the form  $p^{k''} - 1$ . Let  $y$  denote  $n - x$ . Let  $0 \leq x_i \leq p - 1$  denote the  $(i + 1)$ -th digit of  $x$ , so that  $x = \sum_{i=0}^{k''-1} x_i p^i$ . Let  $y_i$  and  $n_i$  be defined similarly. Notice that the base- $p$  representation of  $n$

with exactly  $k'$  digits of  $p - 1$ , so carries cannot occur on the last  $k' + 1$  digits. Now  $x_{k'} > n_{k'}$ , so a carry occurs on the  $k' + 1$ -th digit. Since  $x_i + 1 > n_i$  for each  $k' + 1 \leq i \leq k'' - 1$ , it follows by induction that a carry occurs on the  $i$ -th digit for each  $k' + 2 \leq i \leq k''$ . Therefore,  $x = p^{k''} - 1$  gives the greatest possible number of carries, that is  $k'' - k' = \lfloor \log_p r \rfloor = \lfloor \log_p(n + 1) \rfloor - v_p(n + 1)$ . ■

In this paper, we use the notation  $p^k \parallel n$  to denote  $p^k \mid n$  and  $p^{k+1} \nmid n$ .

**Proposition 6.7.** For prime  $p \mid m$  and  $n = rp^{k'} - 1$  ( $1 \leq r \leq p^k - 1$ ,  $p^k \parallel m$ ),  $\text{diam}(A(m, n)) = n$ .

*Proof.* By Lemma 6.6, for  $0 \leq x \leq rp^{k'} - 1$ , we have  $p^{k'} \nmid \binom{rp^{k'} - 1}{x}$ , so  $m \nmid \binom{rp^{k'} - 1}{x}$ . Therefore, we have  $(0, n, 0), (0, n - 1, 1), \dots, (0, 0, n) \in G$ , and by Theorem 6.5,  $\text{diam}(A(m, n)) = n$ . ■

It should be noted that Proposition 6.7 does not necessarily find all values of  $n$  satisfying  $m \nmid \binom{n}{x} \forall 0 \leq x \leq n$ . An efficient method to generate them would improve the upper bound for  $\text{diam}(A(m, n))$  that we can obtain from Proposition 6.8.

**Proposition 6.8.** If  $\text{diam}(A(m, n')) = n'$ , then  $\text{diam}(A(m, n)) \leq 2n - n'$ .

*Proof.* Let  $V$  be  $V(A(m, n))$ , and let  $V'$  be the set of vertices  $(z_0, z_1, z_2)$  in  $V$  satisfying  $z_1 + z_2 \leq n'$ . We know that the subgraph of  $A(m, n)$  induced by  $V'$  is isomorphic to  $A(m, n')$ , so the distance between two vertices in  $V'$  is at most  $n'$ . Let  $x = (x_0, x_1, x_2)$  and  $y = (y_0, y_1, y_2)$  be two vertices arbitrarily chosen from  $V$ . By Proposition 2.4, there exists a vertex  $x' \in V'$  such that  $d(x, x') = \max(0, x_0 - n') \leq n - n'$ . Similarly, there exists a vertex  $y' \in V'$  such that  $d(y, y') \leq n - n'$ . Therefore,  $d(x, y) \leq d(x, x') + d(y, y') + d(x', y') \leq 2n - n'$ . ■

The analogue of Proposition 6.5 for  $t \geq 3$  does not hold. Indeed, we make the following conjecture.

**Conjecture** For any  $t \geq 3$  and non-prime-power  $m$  (that is, it has multiple different prime divisors)  $m$ , then there exists a positive integer  $N(m)$  such that  $\text{diam}(A_t(m, n)) > n$  for all  $n > N(m)$ .

Little is yet known about the structures of Pascal's simplices modulo prime powers. Indeed, at least 3 papers ([5], [9], [10]) with titles of the form

“Pascal’s triangle modulo  $m$ ”, where  $m$  is a prime power, have been published. We believe that our current state of knowledge is not sufficient for us to conjecture that the analogue of Theorem 6.5 for  $t \geq 3$  holds when  $m$  is a prime power or that it does not.

## 7. Hamiltonicity

In [17], Tegua and Godbole proved that the Sierpiński gasket graphs are Hamiltonian and proposed as an open problem the Hamiltonicity of the generalized Sierpiński gasket graphs generated by Pascal’s triangle modulo a prime greater than 2. In this section, we will prove the existence of Hamiltonian paths and cycles that satisfy a certain property in  $A_{t_p}^k$ , and in the next section, we will apply our result to prove that these generalized Sierpiński gasket graphs are indeed Hamiltonian. By the isomorphism between  $A_{t_2}^k$  and  $S_{t+1}^k$ , our result in this section also contains the result by Klavžar and Milutinović [12] that the Sierpiński graphs  $S_n^k$  are Hamiltonian.

**Definition 7.1.** For tuple of nonnegative integers  $x' = (x'_0, x'_1, \dots, x'_t)$ ,  $F(x')$  is defined to be the set  $\{(x'_0 - 1, x'_1, \dots, x'_t), (x'_0, x'_1 - 1, \dots, x'_t), \dots, (x'_0, x'_1, \dots, x'_t - 1)\}$ .

A set  $S$  is called a funnel if there exists some  $x' = (x'_0, x'_1, \dots, x'_t)$  where  $t \geq 2$  such that  $S = F(x')$ .

For some funnel  $S = F((x'_0, x'_1, \dots, x'_t))$ , the element  $(x_0, x_1, \dots, x_t) \in S$  where  $x_a = x'_a - 1$  and  $x_i = x'_i$  for all  $i \in \{0, 1, \dots, t\} - \{a\}$  is denoted as  $S_a$ .

A set of blocks in some  $G = A_{t_p}^k$  is a block-funnel if it is mapped to a funnel by an isomorphism from  $G'$  to  $A_{t_p}^1$ .

The definition of a funnel is introduced to address a potential flaw in Section 8, in which we generate Hamiltonian cycles of a type of graphs introduced by Tegua and Godbole [17] from funnel-free Hamiltonian cycles of  $A_p^k$ . The proof will fail if the original cycle contains two pairs of consecutive vertices whose union is the subset of a funnel.

We will consider the intersections between funnels and the vertex sets of the graphs  $G_t(n)$  and  $A_{t_p}^k$ . Therefore, it is useful to note the following lemmas.

**Lemma 7.1.** If  $x$  and  $y$  are two adjacent vertices of  $G_t(n)$ , then there is exactly one funnel containing both  $x$  and  $y$ .

*Proof.* Let  $x = (x_0, x_1, \dots, x_t)$  and  $y = (y_0, y_1, \dots, y_t)$ . Assume that  $x_a = y_a + 1$ ,  $y_b = x_b + 1$ , and  $x_c = y_c$  for  $c \in \{0, 1, \dots, t\} - \{a, b\}$ . Therefore, the funnel  $S = F((x'_0, x'_1, \dots, x'_t))$  where  $x'_a = x_a = y_a + 1$ ,  $x'_b = y_b = x_b + 1$ , and  $x'_c = x_c = y_c$  for  $c \in \{0, 1, \dots, t\} - \{a, b\}$  contains both  $x$  and  $y$ , since  $x = S_b$  and  $y = S_a$ .

Now suppose for contradiction that funnel  $T = F((y'_0, y'_1, \dots, y'_t))$  also includes both  $x$  and  $y$ , and that  $T \neq S$ . Assume that  $x = T_b$  and  $y = T_a$ . Since  $T \neq S$ , we have  $a' \neq a$ . Therefore,  $y_a + 1 = x_a \leq y'_a = y_a$ , which is a contradiction. This completes the proof. ■

**Lemma 7.2.** For the corner  $c_i$  of  $G_t(n)$ , any funnel that contains it must contain at most one other vertex of  $G_t(n)$ .

*Proof.* Assume that funnel  $S = F(x')$  contains  $c_i$ . First observe that the number of elements of  $S$  that are also vertices in  $G_t(n)$  is equal to the number of positive integers in  $x'$ , which has at most one more positive integer than of  $c_i$ , which has only one positive integer. Therefore,  $|S \cap G_t(n)| \leq 2$ . ■

**Lemma 7.3.** If  $x = (x_0, x_1, \dots, x_t)$ ,  $y = (y_0, y_1, \dots, y_t)$  and  $z = (z_0, z_1, \dots, z_t)$  are different elements of  $S = F((x'_0, x'_1, \dots, x'_t))$  and  $c$  is an integer such that  $x_c = y_c$ , then  $z_c \leq x_c = y_c$ .

*Proof.* Suppose for contradiction that  $z_c > x_c$ . Since  $x'_c \geq z_c$ , we have  $x'_c > x_c$  and  $x'_c > y_c$ , which means that  $x = y = S_c$ , which is a contradiction. ■

The funnel is thus named because when  $t = 2$ , it (or rather the subgraph of  $G_t(n)$  it induces) looks like a funnel. But note that by the automorphic properties of  $G_t(n)$  and  $A_{tp}^k$ , the "bottom" of a funnel (that is,  $S_0$ ) does not distinguish itself from the other elements of the funnel.

**Definition 7.2.** A path or cycle in  $G_t(n)$  or  $A_{tp}^k$  (or in the block-graph  $G'$ ) is funnel-free if it does not contain two different unordered pairs of consecutive vertices such that their union is the subset of a funnel (or block-funnel). (The two pairs need not be disjoint - for example,  $(u, v)$  and  $(v, w)$  also count.)

Note that a funnel is mapped to a new funnel when the numbers in every one of its elements are permuted in a certain way. A subset of a funnel is mapped to the subset of a new funnel when a certain index with the same value in every one of its elements is deleted.

**Example 7.1.** Let  $S = F((x_0, x_1, x_2, x_3))$ , then  $\{(x_2, x_0, x_1, x_3 - 1), (x_2 - 1, x_0, x_1, x_3), (x_2, x_0, x_1 - 1, x_3), (x_2, x_0 - 1, x_1, x_3)\}$  is a funnel. Also,  $\{(x_0, x_1, x_2, x_3 - 1), (x_0, x_1, x_2 - 1, x_3), (x_0 - 1, x_1, x_2, x_3)\} \subset S$  can be mapped to  $\{(x_0, x_2, x_3 - 1), (x_0, x_2 - 1, x_3), (x_0 - 1, x_2, x_3)\}$ , which is also the subset of a funnel (though not a proper subset).

**Lemma 7.4.** In the graph  $G_t(n)$ , there exist funnel-free Hamiltonian paths between  $c_i$  and  $c_j$  for  $0 \leq i < j \leq t$ .

*Proof.* When  $t = 1$ , the graph is itself a path, so its only Hamiltonian path is funnel-free. When  $n = 0$ , the graph has a Hamiltonian path with a length of 0, which is funnel-free. (These degenerate cases are discussed here only to simplify the base case proof.)

Therefore, the lemma holds when  $t + n = 1$ . Suppose that it holds when  $t + n = s$ . Then we consider the case when  $t + n = s + 1$ . If  $t = 1$  or  $n = 0$ , then the lemma holds. Therefore, assume that  $t > 1$  and  $n > 0$ .

Let  $G_1$  and  $G_2$  be subgraphs of  $G_t(n)$  induced respectively by the set of all vertices  $(x_0, x_1, \dots, x_t)$  where  $x_0 > 0$  and the set of all vertices  $(y_0, y_1, \dots, y_t)$  where  $y_0 = 0$ . Isomorphisms between  $G_1$  and  $G_t(n - 1)$  and between  $G_2$  and  $G_{t-1}(n)$  are given by  $f((x_0, x_1, \dots, x_t)) = (x_0 - 1, x_1, \dots, x_t)$  and  $g((y_0, y_1, \dots, y_t)) = (y_1, \dots, y_t)$ . Notice that the intersection of a funnel with the vertex set of one of the two lesser simplex grid graphs is mapped to the intersection of a funnel with the vertex set of the corresponding subgraph by these isomorphisms.

Since the lemma holds when  $t + n = s$ , both  $G_t(n - 1)$  and  $G_{t-1}(n)$  have funnel-free Hamiltonian paths between any two corners. Let  $P_1$  be a funnel-free Hamiltonian path of  $G_1$  between  $(n, 0, \dots, 0)$  and  $(1, n - 1, \dots, 0)$  and let  $P_2$  be a funnel-free Hamiltonian path of  $G_2$  between  $(0, n, \dots, 0)$  and  $(0, 0, \dots, n)$ . We concatenate these two paths with the edge  $(1, n - 1, \dots, 0) - (0, n, \dots, 0)$  to get  $P$ , a Hamiltonian path of  $G_t(n)$  between  $(n, 0, \dots, 0)$  and  $(0, 0, \dots, n)$ .

Suppose for contradiction that  $P$  is not funnel-free. Then let  $(u_1, u_2)$



and  $(v_1, v_2)$  be the two pairs of consecutive vertices in  $P$  whose union is the subset of a funnel. By Lemma 7.2, no such pair contains  $(0, n, \dots, 0)$ , so each is either a subset of  $V(G_1)$  or a subset of  $V(G_2)$ . By our hypothesis, the two pairs cannot both be subsets of  $V(G_1)$  or of  $V(G_2)$ . Therefore, one such pair is a subset of  $V(G_1)$  and the other pair a subset of  $V(G_2)$ . However, according to Lemma 7.3, if two vertices in a funnel are in  $V(G_2)$ , then no other vertex in the funnel is in  $V(G_1)$ , which is a contradiction. Therefore,  $P$  is funnel-free, and this completes the proof. ■

The following proof is very similar to that of Lemma 7.4, and certain details are not repeated here.

**Lemma 7.5.** In the graph  $G_t(n)$ , there exists a funnel-free Hamiltonian cycle.

*Proof.* As in the proof of Lemma 7.4, let  $G_1$  and  $G_2$  be subgraphs of  $G_t(n)$  induced respectively by the set of all vertices  $(x_0, x_1, \dots, x_t)$  where  $x_0 > 0$  and the set of all vertices  $(y_0, y_1, \dots, y_t)$  where  $y_0 = 0$ . Let  $P_1$  be a funnel-free Hamiltonian path of  $G_1$  between  $(1, n-1, \dots, 0)$  and  $(1, 0, \dots, n-1)$  and let  $P_2$  be a Hamiltonian path of  $G_2$  between  $(0, n, \dots, 0)$  and  $(0, 0, \dots, n)$ . We concatenate these two paths with the two edges  $(1, n-1, \dots, 0) - (0, n, \dots, 0)$  and  $(1, 0, \dots, n-1) - (0, 0, \dots, n)$  to get a Hamiltonian cycle  $C$  of  $G_t(n)$ .

Suppose that  $C$  is not funnel-free. A contradiction can be obtained with an argument similar to that used in the proof of Lemma 7.4, completing this proof. ■

**Lemma 7.6.** For  $G = A_t^k$  and funnel  $S$ , the elements of  $S' = S \cap G$  where  $|S'| \geq 3$  are either all in the same block or are in  $|S'|$  different blocks. In the latter case, the set of the corresponding blocks is the subset of a block-funnel in the block-graph  $G'$ .

*Proof.* Let  $S = F((x'_0, x'_1, \dots, x'_t))$  and let  $S' = S \cap G$ . Assume that  $S_a = (y_0, y_1, \dots, y_t)$  and  $S_b = (x_0, x_1, \dots, x_t)$  are in  $S'$  and are in different blocks. Since  $S_a$  and  $S_b$  are adjacent, the blocks are b-adjacent, which means that

$$\left\lfloor \frac{y_a}{p^{k-1}} \right\rfloor + 1 = \left\lfloor \frac{x_a}{p^{k-1}} \right\rfloor \quad \text{and} \quad \left\lfloor \frac{y_b}{p^{k-1}} \right\rfloor = \left\lfloor \frac{x_b}{p^{k-1}} \right\rfloor + 1 \quad . \quad \text{Therefore,} \quad \sum_{i=0}^t \left\lfloor \frac{x'_i}{p^{k-1}} \right\rfloor = \sum_{i=0}^t \left\lfloor \frac{x_i}{p^{k-1}} \right\rfloor + 1 = p = \sum_{i=0}^t \frac{x'_i}{p^{k-1}}, \text{ so } p^{k-1} | x'_i \text{ for all } 0 \leq i \leq t. \text{ So if } S_{a'} \text{ and}$$

$S_{b'}$  are two elements of  $S'$ , then  $\left\lfloor \frac{x'_{a'}-1}{p^{k-1}} \right\rfloor < \left\lfloor \frac{x'_{a'}}{p^{k-1}} \right\rfloor$ , which means that  $S_{a'}$  and  $S_{b'}$  are in different blocks.

Now observe that by an isomorphism from  $G'$  to  $A_{tp}^1$ , the block that contains  $S_{a'}$  is mapped to an element of  $F\left(\left(\left\lfloor \frac{x'_0}{p^{k-1}} \right\rfloor, \left\lfloor \frac{x'_1}{p^{k-1}} \right\rfloor, \dots, \left\lfloor \frac{x'_t}{p^{k-1}} \right\rfloor\right)\right)$ . ■

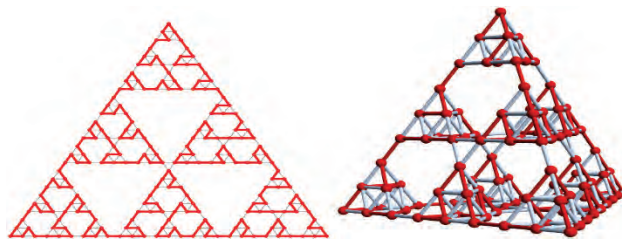
**Lemma 7.7.** In  $A_{tp}^k$  there exist funnel-free Hamiltonian paths between  $c_i$  and  $c_j$  for  $0 \leq i < j \leq t$ .

*Proof.* We proceed with induction. By Lemma 7.4, this lemma holds when  $k = 1$ . Assume that it holds for  $A_{tp}^k$ . Let  $G = A_{tp}^{k+1}$ , and for every block there exists funnel-free Hamiltonian paths between  $c_i$  and  $c_j$  for  $0 \leq i < j \leq t$ . Let  $P'$  be a funnel-free Hamiltonian path of the block-graph  $G'$  between  $B_{p-1,0,\dots,0}$  and  $B_{0,0,\dots,p-1}$ . Let  $P$  be a Hamiltonian path of  $G$  between  $(p^{k+1} - 1, 0, \dots, 0)$  and  $(0, 0, \dots, p^{k+1} - 1)$ , obtained by concatenating funnel-free Hamiltonian paths of the blocks in the order of  $P'$ .

Suppose for contradiction that  $P$  is not funnel-free. Then let  $(u_1, u_2)$  and  $(v_1, v_2)$  be the two pairs of consecutive vertices in  $P$  whose union is the subset of a funnel. The vertices cannot all be in the same block, and by Lemma 7.6,  $P'$  is not funnel-free, which is a contradiction. ■

**Theorem 7.8.** There is a funnel-free Hamiltonian cycle in  $A_{tp}^k$ .

*Proof.* Let  $G = A_{tp}^k$ , and let  $C'$  be a funnel-free Hamiltonian cycle of the block-graph  $G'$ . Let  $C$  be a Hamiltonian path of  $G$  obtained by concatenating funnel-free Hamiltonian paths of the blocks in the order of  $C'$ . Suppose that  $C$  is not funnel-free. A contradiction can be obtained with an argument similar to that used in the proof of Lemma 7.7, completing this proof. ■



**Figure 7.** Funnel-free Hamiltonian cycles of  $A_3^3$  and  $A_{33}^2$ .

## 8. A problem proposed by Teguia and Godbole

In this section, we apply results we obtained to prove the Hamiltonicity of the generalization of the Sierpiński gasket graphs proposed by Teguia and Godbole [17] by presenting a method to construct a Hamiltonian cycle of a generalized Sierpiński gasket graph from a Hamiltonian cycle of  $A_p^k$ . The proof is illustrated by Figure 8.

**Definition 8.1.** If  $x = (x_0, x_1, x_2)$  is a vertex of  $G(n)$ , then  $t(x)$  is the induced subgraph of  $G(n+1)$  with vertices  $\{(x_0, x_1, x_2 + 1), (x_0, x_1 + 1, x_2), (x_0 + 1, x_1, x_2)\}$ .

**Definition 8.2.** If  $G$  is an induced subgraph of  $G(n)$ , then  $T(G)$  denotes the graph  $\cup_{v \in G} t(v)$ .

Now the previous two definitions map every vertex of  $G(n)$  to a triangle in  $G(n+1)$ . The generalized Sierpiński gasket graphs are composed of triangles that are arranged in the same way as the vertices of  $A_p^k$ , so we can use  $A_p^k$  to generate (and thereby give a rigorous definition to) the corresponding gasket graph.

**Definition 8.3.** The generalized Sierpiński gasket graph  $SG_p^k$  denotes  $T(A_p^k)$ .

Now  $SG_2^k$  is isomorphic to the Sierpiński gasket graph  $S_{k+1}$ , and the desired generalization is  $SG_p^k$  where  $p \geq 3$ .

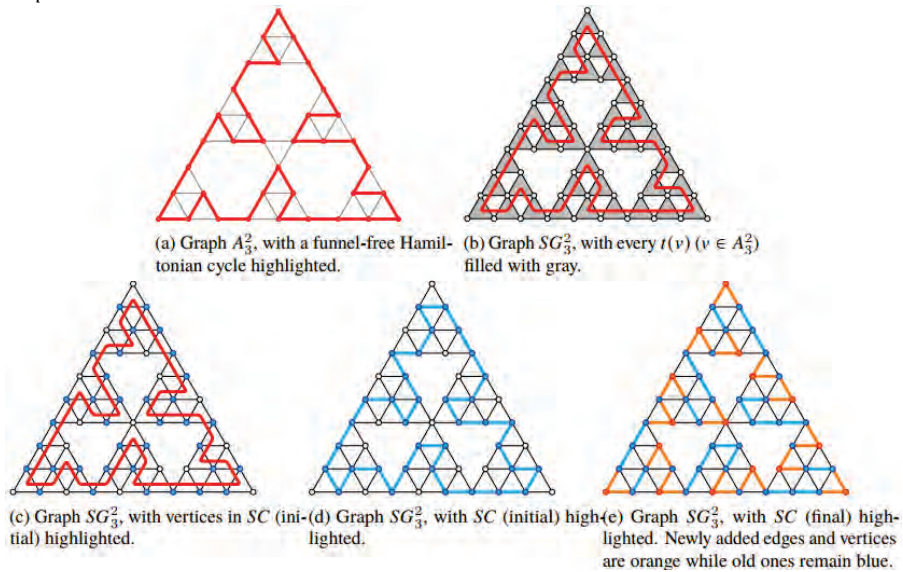
**Lemma 8.1.** If  $u, v$  and  $w$  are different vertices in  $G(n)$  and  $t(u), t(v)$  and  $t(w)$  have a common vertex, then  $u, v, w$  is a funnel.

*Proof.* Let such a common vertex be  $(x_0, x_1, x_2)$ . By definition,  $u, v, w \in \{(x_0, x_1, x_2 - 1), (x_0, x_1 - 1, x_2), (x_0 - 1, x_1, x_2)\}$ . Since  $u, v, w$  are pairwise different,  $\{u, v, w\} = F((x_0, x_1, x_2))$ . ■

**Theorem 8.2.**  $SG_p^k$  is Hamiltonian.

*Proof.* According to Theorem 7.8, there exists a funnel-free Hamiltonian cycle of  $A_p^k$ . Let such a cycle be  $C = (v_1, v_2, \dots, v_l)$ . Let  $v_0$  denote  $v_l$  and let  $v_{l+1}$  denote  $v_1$ . Let  $u_{i,j}$  denote the common vertex of  $t(v_i)$  and  $t(v_j)$ . By Lemma 8.1, for any  $1 \leq i \leq l$ , we have  $u_{i-1,i} \neq u_{i,i+1}$ . So, let  $w_i$  denote the vertex in  $t(v_i)$  that is neither  $u_{i-1,i}$  nor  $u_{i,i+1}$ .

Then we apply the following procedure to find a Hamiltonian cycle of  $SG_p^k$ . Let  $SC$  be the cycle  $(u_{1,2}, \dots, u_{l-1,l}, u_{l,1})$ . We go from  $w_1$  to  $w_l$ , and for each vertex  $w_i$ , if it is not already in  $SC$ , we insert it between  $u_{i-1,i}$  and  $u_{i,i+1}$ . When we finish this process,  $SC$  will have become a Hamiltonian cycle of  $SG_p^k$ . ■



**Figure 8.** Illustration of the process generating a Hamiltonian cycle of  $SG_p^k$ .

## Appendix. Program to find Hamiltonian cycles of $A_{t,p}^k$

```

from numpy import array, concatenate, append, roll, where, argsort, arange
from math import comb
from numpy.linalg import norm
import matplotlib.pyplot as plt
from matplotlib.colors import LinearSegmentedColormap

def simplex_grid_path(t, n):
    """
    Input: t, n

    Output: A numpy array describing a Hamiltonian path of
    the simplex grid graph  $G_t(n)$  with end vertices  $c_0$  and  $c_t$ .

    Example:
    >>> simplex_grid_path(2,2)
    array([[2, 0, 0],
           [1, 0, 1],
           [1, 1, 0],
           [0, 2, 0],
           [0, 1, 1],
           [0, 0, 2]])
    """
    if t==0:
        return array([[n]])

    return array([append([n-i], roll(v, n-i))
                  for i in range(n+1)
                  for v in simplex_grid_path(t-1, i)])

def Hamiltonian_cycle_of_A(t, p, k):
    """
    Input: t, p, k

    Output: A numpy array describing a Hamiltonian cycle of
    the graph  $\{A_t\}_p^k$ .

    Example:
    >>> Hamiltonian_cycle_of_A(2, 2, 2)
    
```

```

array([[2, 1, 0],
       [3, 0, 0],
       [2, 0, 1],
       [1, 0, 2],
       [0, 0, 3],
       [0, 1, 2],
       [0, 2, 1],
       [0, 3, 0],
       [1, 2, 0]])
'''
basic_path = simplex_grid_path(t, p-1)
def get_basic_path(start, end):
    '''
    Returns a Hamiltonian path of G_t(p-1)
    starting at c_{start} and ending in c_{end}.
    '''
    inverse_permutation = [start] + list(set(range(t+1)) - {start, end}) + [end]
    permutation = argsort(inverse_permutation)
    return basic_path[:, permutation]

component1 = simplex_grid_path(t, p-2)
component1[:, [0, 1]] = component1[:, [1, 0]]
component1 += array([1]+[0]*t)
component2 = append(array([[0]]*comb(p+t-2, t-1)),
                    simplex_grid_path(t-1, p-1)[::-1],
                    axis=1)
cycle = concatenate((component1, component2))

for i in range(2, k+1):
    cycle = array([p * b + v
                  for a, b, c in zip(roll(cycle, 1, axis=0),
                                     cycle,
                                     roll(cycle, -1, axis=0))
                  for v in get_basic_path(where((a-b)==1)[0][0],
                                           where((c-b)==1)[0][0])])

return cycle

plt.rcParams["figure.figsize"] = (8,4*(3**0.5))
position_matrix = [None, None, [[0,1,2],[0,3*0.5,0]], [[0,1,1,4],[0,1,4,1],[0,4,1,1]]]
def visualize_cycle(t, p, k):
    '''
    Input: t, p, k

    Output: Shows a figure showing a Hamiltonian cycle of {A_t}_p^k.
    No value is returned.

    When t=3, the interactive figure can be rotated and zoomed.

    It should be noted that 3D figures produced by matplotlib
    are subject to some glitches. For example, it is very often
    that an edge that should be under another edge is above it.
    However, matplotlib does have the virtue of being
    comparatively fast and interactive.
    '''
    if t==2:
        fig, ax = plt.subplots()
    elif t==3:
        fig = plt.figure(1)
        fig.add_subplot(111, projection='3d')
        ax = plt.axes(projection='3d')
    else:
        raise ValueError('Bad dimension %s.' % t)

    cycle = Hamiltonian_cycle_of_A(t, p, k)
    coord = array([sum((v*position_matrix[t]).transpose()).transpose() for v in cycle])
    length = len(cycle)

    for i in range(length):
        for j in range(i):
            if norm(cycle[i]-cycle[j], 1)!=2:
                continue
            ax.plot(*array([coord[i], coord[j]]).transpose(),
                   c='red' if ((i-j-1)%length)*((j-i-1)%length)==0 else '#e0e0e0')
    ax.scatter(*coord.transpose(), c=arange(length), cmap="hsv", zorder=1000)

    plt.axis('off')
    plt.show()

```

## References

- [1] T. Bannick and H. Buhrman, Quantum Pascal's Triangle and Sierpiński's Carpet. Preprint (2017). <https://arxiv.org/abs/1708.07429>
- [2] A. Bogomolny, Patterns in Pascal's Triangle. *Cut-the-knot* <https://www.cut-the-knot.org/arithmetic/combinatorics/PascalTriangleProperties.shtml>
- [3] G.-H. Chen and D.-R. Duh, Topological Properties, Communication, and Computation on WK-recursive Networks. *Networks* **24** (1994), 303–317.
- [4] R. H. Cooper, R. Hunter-Duvar and W. Patterson, A More Efficient PublicKey Cryptosystem Using the Pascal Triangle. *IEEE International Conference on Communications, World Prosperity Through Communications* **3** (1989), 1165-1169.
- [5] K. S. Davis and W. A. Webb, Pascal's Triangle modulo 4. *Fibonacci Quarterly* **29** (1991), 79-83.
- [6] L. E. Dickson, *History of the Theory of Numbers*. Chelsea Publishing Company (1952), 263-278.
- [7] A. M. Hinz, Pascal's Triangle and the Tower of Hanoi. *American Mathematical Monthly* **99** (1992), 538-544.
- [8] N. Holter, A. Lakhtakia, V. K. Varadan, V. V. Varadan and R. Messier, On a New Class of Planar Fractals: the Pascal-Sierpiński Gaskets. *Journal of Physics A: Mathematical and General* **19** (1986), 1753-1759.
- [9] J. G. Huard, B. K. Spearman and K. S. Williams, Pascal's Triangle (mod 8). *European Journal of Combinatorics* **19** (1998), 45-62.
- [10] J. G. Huard, B. K. Spearman and K. S. Williams, Pascal's Triangle (mod 9). *Acta Arithmetica* **78** (1997), 331-349.
- [11] S. Klavžar, Coloring Sierpiński Graphs and Sierpiński Gasket Graphs. *Taiwanese Journal of Mathematics* **12** (2008), 513–522.
- [12] S. Klavžar and U. Milutinović, Graphs („=, :” and a Variant of the Tower of Hanoi Problem. *Czechoslovak Mathematical Journal* **1** (1997), 95–104.
- [13] S. L. Lipscomb, *Fractals and Universal Spaces in Dimension Theory*. Springer (2009).
- [14] C.-S. Liu and C.-L. Kuo, A Multiple-Scale Pascal Polynomial Triangle Solving Elliptic Equations and Inverse Cauchy Problems. *Engineering Analysis with Boundary Elements* **62** (2016), 35-43.
- [15] A. M. Reiter, Determining the Dimension of Fractals Generated by Pascal's Triangle. *Fibonacci Quarterly* **31** (1993), 112-120.

- [16] E. B. Rosenthal, A Pascal Pyramid For Trinomial Coefficients. *The Mathematics Teacher* **54** (1961), 336-338.
- [17] A. M. Teguia and A. P. Godbole, Sierpiński Gasket Graphs and Some of Their Properties. *Australasian Journal of Combinatorics* **35** (2006), 181-192.
- [18] D. B. West, *Introduction to Graph Theory, 2nd ed.* China Machine Press (2020), 390-391.
- [19] D. Woods and M. Kohlenberg, Pascal's  $k$ -Simplex. *The Two-Year College Mathematics Journal* **4** (1973), 38-43.

