SCHOLARLY REVIEW

ISSUE#2

Fall 2021

contact@scholarlyreview.org

About

The Scholarly Review is an open-access, quarterly journal dedicated to publishing academic research in the natural sciences, social sciences and humanities. Our independent editorial board is made of highly accomplished professors and academics in corresponding fields who review student submissions and select pieces for publication.

Editorial Board

TABLE OF CONTENTS

Natural Sciences and Computer Science

Social Sciences, Health and Public Policy

TABLE OF CONTENTS

Arts, Activism and Humanities

Editorial: SR Issue Fall 2021:

Welcome to the second issue of the Scholarly Review, a multi-disciplinary journal for student authors, providing young and emerging scholars with a platform from which to showcase their work, while at the same time providing readers with insights into a wide range of topics. This issue covers subjects ranging from Renaissance art to the Chinese pharmaceutical industry, the MeToo movement, space exploration and more. The editorial board was delighted to receive such high-quality papers, and while papers are organized under the headings *Natural sciences and computer science; Social Sciences, health and public policy; Arts, activism and humanities*, some papers cross disciplines and these divisions are merely a guide. The depth and breadth of topics covered may provide absorbing material for the reader as well as inspiration for other would-be authors.

Donggeon Kim's paper is concerned with the characteristics of exoplanets and their habitability before going on to propose a possible future exoplanet classification scheme. Michael Zheng's paper reveals that when using an object detection system on the ExDark dataset, applying frozen layers transfer learning can effectively compensate for small reductions in the amount of data in a small dataset. Larry Hui's paper uses a two-parameter linear model, performing a chi-square data test to determine the fractal dimensions of crumpled paper.

Daisy Lin's paper assesses COVID-19 vaccine policies in China and the USA, assessing the effectiveness of approaches adopted by these countries for controlling the pandemic. Jinghzi Zhuang's paper is a literature review exploring possible reasons for inconsistent results in studies on the arousal-emotion theory, studying the effects of music on the individual. Valerie Xu's paper analyzes the Green New Deal in the USA and the economic implications of the transition from fossil to renewable energy sources. Walter Liu's paper discusses the pharmaceutical market in China, analyzing disadvantages and potential benefits of policies introduced recently by the Chinese government. Zach Wu's paper looks at broadband infrastructure and policy in the USA, with the aim of finding ways to reduce disparities between those with and without a high-speed connection.

João Victor Ataides Marques's paper describes military movements in the state of Goiás during the Brazilian military dictatorship (1964-1985), demonstrating the negative impact of dictatorship on the public. Megan O'Sullivan's paper examines the MeToo movement and the negative impact on people's lives for sexual harassment survivors, including celebrities. Summer Xu's paper uses case studies to illustrate the complexities of censorship in China's entertainment industry, with a particular focus on LGBTQ film content. Vincent Luo's paper provides in-depth analysis of Raphael's well-known 16th century altarpiece, addressing the harmony of composition, the use of symmetry and asymmetry as well as the choice of colors, use of realism and linear perspective techniques.

Scholarly Review

An Up-To-Date Exploratory Survey of 4434 Confirmed Exoplanets and an Insight into the Sensitivity of Detection Methods

By Donggeon Kim

Author Bio:

Donggeon Kim, born in Seoul, South Korea, is a senior at the Shanghai American School – Puxi Campus. He has an exceptional passion for astrophysics, particularly in exoplanets, and is intending to major in either Applied Mathematics, Astrophysics, or Aerospace Engineering. He first became interested in the universe after visiting an observatory with his family in grade 5. Now, he loves applying his math and physics skills to many different disciplines to solve real-world problems. He is also an avid communicator, striving to share his passion with his community through technology and education. Outside of academics, you can find him playing baseball, table tennis, solving Rubik's cube, and watching *Interstellar*.

Abstract

The recent development of detection methods has produced large metadata of exoplanets, which are stored and regularly updated on the NASA Exoplanet Archive. In this paper, classification and habitability schemes by Kopparapu, R. et al are incorporated to classify 4434 confirmed exoplanets into 5 different categories—Rocky, Super-Earths, Sub-Neptune, Sub-Jovian, and Jovian—and to compute their habitability. In addition, 20 graphs are plotted to analyze the trends of exoplanets' characterization, habitability, and detection methods sensitivity. This paper also upgrades a current NASA Exoplanet Archive catalog to be more comprehensive by appending classification and habitability flag columns. Moreover, a novel exoplanet classification scheme is proposed to utilize Weighted Average for more comprehensive metadata from futuristic telescopes such as the James Webb Space Telescope. Codes, datasets, plots, and the upgraded exoplanet catalog from this research are published on the GitHub page, for readers to repeat this research's analysis with newly updated exoplanet data.

Keywords: Physics and Astronomy; Planetary Science; Exoplanet; NASA Exoplanet Archive; Exoplanet Classification; Habitable Zone; Goldilocks Zone; Detection Method Sensitivity

Introduction

Our search for worlds beyond the Earth - the exoplanets, or the planets outside the Solar System - was sparked over 2000 years ago, when the Greek philosopher Epicurus (ca. 300 B.C.E.) asserted "There are infinite worlds both like and unlike this world of ours" [1]. However, it was only in 1992 that the first exoplanet was officially detected and confirmed [2]. Thus, the field of exoplanet research is in a relatively new era, but a rapid one: over the past two decades, with a help of advanced detection techniques, more than 4000 exoplanets have been confirmed along with more than 7000 "candidate" exoplanets [3]. As a side note, exoplanets are considered confirmed only once they are verified through additional observation using two other telescopes [4].

1. Classification of Exoplanets

The exponential detection of exoplanets signaled a growing need to classify exoplanets into certain categories to understand their diversity[5]. Exoplanets are typically classified into certain categories based on their characteristics, such as radius and mass, compared to our own known solar planets. For example, NASA categorized exoplanets into Terrestrial, Super Earth, Gas Giant, and Neptune-like [3]. While there is no one officially accepted classification scheme because of the complexity of exoplanets that cannot be described by just one mathematical model, this paper uses the classification scheme proposed by Dr. Ravi Kumar Kopparapu et al. in 2018 in their paper "Exoplanet Classification and Yield Estimates for Direct Imaging Missions". This classification scheme prioritizes the size of exoplanets and the starlight flux on the planet as main factors on planet composition between 5 categories: Rocky, Super Earths, Sub-Neptune, sub-Jovian, and Jovian [5].

2. Habitability of Exoplanets

Another way to classify exoplanets is by determining the habitability of exoplanets: whether the exoplanet is within the habitable zone boundary. The habitable zone, also known as the Goldilocks zone, is the circular boundary of orbits around a star within which a planetary surface is not too cold and not too hot to support essential substances for life such as CO_2 and H_2O [6]. Again, there is no one officially accepted habitable zone scheme, but this paper applies the habitable zone scheme proposed by Dr. Ravi Kumar Kopparapu et al. in 2013 in their paper "Habitable Zones Around Main-Sequence Stars: New Estimates", which takes into account the stellar luminosity computed by stellar effective temperature and stellar radius [6].

3. Detection Methods of Exoplanets

Since exoplanets are small, and host stars are so bright that they outshine their planets, it is generally difficult to directly detect exoplanets. In light of this, there have been many approaches to indirectly detect exoplanets by searching for characteristics of host stars, which are easier to detect. As of now, there are largely 5 detection methods: transit method, radial velocity, microlensing, direct imaging, and pulsar timing. The transit method detects the tiny dips in light when the exoplanet crosses the host star and blocks the starlight in the direction to the Earth [1]. The radial velocity method detects the Doppler's effect of a host star's changes in radial velocity caused by the gravitational pulls between the star and the planet. Keep in mind that the star orbits around the center of its planetary system, not just staying in its center. If the host star moves toward the direction of the Earth, the wavelengths of starlight are squeezed, emitting blue-shifted light signal, and if it moves away, the wavelengths of starlight are stretched, emitting red-shifted light signal [1]. The microlensing method is used when a foreground star happens to pass a more distant background star. As the foreground star passes the background star, the background star's brightness will increase due to the Gravitational Lensing effect. If the foreground star happens to host a planet, the planet will also act as a Gravitational lens, resulting in a unique peak in the background star's brightness [1]. Direct imaging uses infrared wavelengths to directly observe planets [1]. Lastly, pulsar timing detects exoplanets around pulsars, which emit an intense electromagnetic radiation on a regular rate. The slight regular variations in the timing of the pulses indicate that pulsar orbits around the center of mass of a system with one or more planets, suggesting the existence of the exoplanets [1].

4. Outline of the Paper

This paper focuses on analyzing the demographics of currently confirmed 4434 exoplanets based on classification and habitability schemes, as well as an insight of the sensitivity of each detection method. Along with this analysis, I made two contributions to the exoplanet community. While the exoplanets catalog provides extensive information about each exoplanet and its host star, the catalog does not include the habitability and the classification category of each exoplanet. Employing a highly respected habitable zone scheme (cited 1102 times) and classification scheme proposed by the same author, I updated the current exoplanet catalog with the habitability flag column and classification column. The updated catalog will provide more comprehensive information and trends of currently confirmed exoplanets.

This paper is divided into 6 parts: *introduction, materials and method, plots, discussion, caveat and future research, and conclusion*. I also shared my code, plots, dataset, and the updated exoplanet catalog on my GitHub page (https://github.com/SteveHawKim03/exoplanet-analysis) so that readers can repeat my analysis whenever new confirmed exoplanets are added to the NASA Exoplanet Archive catalog.

Materials and Methods

1. Exoplanet Catalog

This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program [7]. There were 4434 confirmed exoplanets as of July 12th 2021.

While there are more than 100 variables to each exoplanet, my research focuses on 10 variables. These 10 variables are explained in more detail in the table below [8]:

I. Variables and Descriptions

Variable Name	Unit	Description
Detection Method	N/A	Method by which the planet was first identified.
Orbital Period	Days	Time the planet takes to make a complete orbit around the host star or system.
Orbit Semi-Major Axis	AU	The longest radius of an elliptic orbit of the planet. Used to represent the separation between the host star and the exoplanet.
Planet Radius	Earth Radius	Length of a line segment from the center of the planet to its surface, measured in units of radius of the Earth.
Planet Mass	Earth Mass	Best planet mass measurement/approximation in units of masses of Earth
Insolation Flux	Earth Flux	Flux of solar radiation per unit of horizontal area for a planet. Another way to give the equilibrium temperature, which is the temperature of the planet as modeled by a black body heated only by its host star.
Stellar Effective Temperature	Kelvins	Temperature of the star as modeled by a black body emitting the same total amount of electromagnetic radiation.
Stellar Radius	Solar Radius	Length of a line segment from the center of the star to its surface, measured in units of radius of the Sun.
RA	Degree	Right Ascension - east and west of the celestial equator - of the planetary system.
Dec	Degree	Declination - north and south of the celestial equator - of the planetary system.

2. Python Packages

For my analysis on these data, I used Python along with four packages: Pandas, NumPy, Matplotlib, and Astropy [9][10]. I used Pandas to read and write csv files of the exoplanets catalog, NumPy to generate arrays of data and compute mathematical operations, Matplotlib to plot graphs, and Astropy to bring astrophysical constants and analyze the plots.

3. Habitability Scheme

As mentioned in the introduction, I used the habitable zone scheme proposed by Dr. Ravi Kumar Kopparapu et al. in their paper "Habitable Zones Around Main-Sequence Stars: New Estimates" [6]. The scheme first calculates the habitable zone stellar fluxes (S_{eff}), measured in Kelvins [1], reaching the top of the atmosphere of an Earth-like planet in relation to the stellar effective temperature (T_{eff}), measured in Kelvins:

where T. = T_{eff} − 5780 K and the coefficients , *a, b, c*, and *d* are as follows:

II. Habitable Zone Constants

Constant	Recent Venus	Runaway Greenhouse	Moist Greenhouse	Maximum Greenhouse	Early Mars
	1.7753	1.0512	1.0140	0.3438	0.3179
a	1.4316×10^{-4}	1.3242×10^{-4}	8.1774×10^{-5}	5.8942×10^{-5}	5.4513×10^{-5}
b	2.9875×10^{-9}	1.5418×10^{-8}	1.7063×10^{-9}	1.6558×10^{-9}	1.5313×10^{-9}
c	-7.5702×10^{-12}	-7.9895×10^{-12}	-4.3241×10^{-12}	-3.0045×10^{-12}	-2.7786×10^{-12}
d	-1.1635×10^{-15}	-1.8328×10^{-15}	-6.6462×10^{-16}	-5.2983×10^{-16}	-4.8997×10^{-16}

While this scheme proposes two types of definition, the narrower 'conservative habitable zone' and wider 'optimistic habitable zone', I chose to use the wider 'optimistic habitable zone' definition because the wider definition of the habitable zone is more comprehensive as it entails the potential factors of water and CO2 clouds on the planet [11]. The wider optimistic habitable zone definition is bounded by the 'Recent Venus' and 'Early Mars' limits whereas the narrower conservative habitable zone definition is bounded by the 'Moist Greenhouse' and 'Maximum Greenhouse' limits. Thus, we will only need data in the 'Recent Venus' and 'Early Mars' columns.

Once we calculate the habitable zone stellar fluxes (S_{eff}), we can calculate the corresponding HZ distance limits (d) by using the relation [2] where $L/L_{\odot}$ is the luminosity of the star compared to the Sun, which can be calculated by [3] where R is the star's radius, $R_{\odot}$ is the Sun's radius, equal to 695700 km, and $T_{eff_{\odot}}$ is the temperature of the Sun, equal to 5778 K [12].

Using the HZ distance limits (d) obtained with coefficients that correspond to 'Recent Venus' and 'Early Mars', we can calculate the boundary of optimistic habitable zones. If the exoplanet's separation (distance) from the star, defined by the orbital semi-major axis, is within this habitable-zone boundary, then that exoplanet is classified as habitable with a Boolean value of True. If not, meaning that the exoplanet's semi-major axis is either greater or smaller than the maximum limit or the minimum limit of the boundary, respectively, the exoplanet is classified as not habitable with a Boolean value of False.

4. Exoplanet Classification Scheme

The exoplanet classification scheme proposed by Dr. Ravi Kumar Kopparapu et al. in their paper "Exoplanet Classification and Yield Estimates for Direct Imaging Missions" follows a more simple relation, based on chemical species' condensation sequences in planetary atmospheres. Their classification scheme can be summarized by the table below [5]:

III. Exoplanet Classification Scheme

Planet Type (Stellar Flux Range) [Earth Flux]	Planet Radius [Earth Radius]
Hot rocky (182 - 1.0)	0.5-1.0
Warm rocky (1.0 - 0.28)	0.5-1.0
Cold rocky (0.28 - 0.0035)	0.5-1.0
Hot super-Earths (187 - 1.12)	1.0-1.75
Warm super-Earths (1.12 - 0.30)	1.0-1.75
Cold super-Earths (0.30 - 0.0030)	1.0-1.75
Hot sub-Neptune (188 - 1.15)	1.75-3.5
Warm sub-Neptune (1.15 - 0.32)	1.75-3.5
Cold sub-Neptune (0.32 - 0.0030)	1.75-3.5
Hot sub-Jovian (220 - 1.65)	3.5-6.0
Warm sub-Jovian (1.65 - 0.45)	3.5-6.0
Cold sub-Jovian (0.45 - 0.0030)	3.5-6.0
Hot Jovian (220 - 1.65)	6.0-14.3
Warm Jovian (1.65 - 0.40)	6.0-14.3
Cold Jovian (0.40 - 0.0025)	6.0-14.3

Results

Using the Python packages, variables, and schemes mentioned in the previous section, I analyzed the data and produced two types of graphs - bar graph and scatter plot. Bar graphs show the frequency of different demographical categories, and scatter plots show the trends that the data of 4434 confirmed exoplanets follow. I categorized each plot with a distinct alphanumerical index, as shown below. The plots are summarized in the appendix.

A. Bar Graph

1. Classification Frequency (*A-1.1*)
   - Rocky Classification Frequency (*A-2.2*)
   - Super Earths Classification Frequency (*A-.3*)
   - Sub-Neptune Classification Frequency (*A-3.4*)
   - Sub-Jovian Classification Frequency (*A-3.5*)
   - Jovian Classification Frequency (*A-3.6*)

2. Habitability Frequency (*A-2.1*)
   - Habitable Classification Frequency (*A-2.2*)
   - Habitable Classification Percentage Frequency (*A-2.3*)
   - Detection Method Frequency (*A-3*)

B. Scatter Plot

1. Exoplanet Characterization Analysis
   - *Orbit Semi-Major Axis vs Insolation Radius vs Stellar Effective Temperature (B-1.1)*
   - *(Orbit Semi-Major Axis)^3 vs (Orbital Period)^2 (B-1.2)*
   - *Mass Radius Ratio vs Orbit Semi-Major Axis vs Classification (B-1.3)*
   - *Mass Radius Ratio vs Flux vs Classification (B-1.4)*
   - *Planet Mass vs Planet Radius vs Classification (B-1.5)*

2. Habitability Analysis
   - *Planet Mass vs Planet Radius vs Habitability (B-2.1)*
   - *Mass Radius Ratio vs Flux vs Habitability (B-2.2)*

3. Detection Method Analysis
   - *Orbit Period vs Planet Radius vs Detection Method (B-3.1)*
   - *Orbit Period vs Planet Radius vs Detection Method (B-3.2)*
   - *Skymap of Confirmed Exoplanets (RA vs Dec vs Detection Method) (B-3.3)*

Materials and Methods

With the following plots, I surveyed currently confirmed exoplanets in three different perspectives: characterization and classification, detection method, habitability of confirmed exoplanets.

There are four variables that characterize exoplanets (orbital period, orbit semi-major axis, planet radius, and planet mass), three variables that characterize the host star (insolation flux, stellar effective temperature, and stellar radius), and two variables that indicate the location of the exoplanet (right ascension and declination).

1. Exoplanet Characterization Analysis

First, I analyzed the characterization and classification of currently confirmed exoplanets with bar graphs *A-1.1 to A-1.6* and scatter plots *B-1.1 to B-1.5*. The bar graph A-1.1 shows the number of categories that each exoplanet is assigned to based on the classification scheme [5]. According to this graph, Sub-Neptunes are the most confirmed exoplanet type with 1554, exoplanets followed by Jovians with 1202, Super Earths with 944, Sub-Jovians with 328, and Rocky with 164. This is quite similar to NASA's manual classification, which has classified exoplanets into 1497 Neptune-like, 1403 Gas Giant, 1364 Super Earths, 165 terrestrial, and 5 unknown [3]. Unlike typical exoplanet classifications, [5] this classification scheme also proposed classification by flux to further classify into 'hot', 'warm', and 'cold' exoplanets of each 5 category. The trend that more bigger exoplanets are frequent in the dataset shows the sensitivity of our current telescopes on bigger exoplanets. Since not all exoplanets have flux data, there are exoplanets that cannot be further classified, marked as 'Flux N/A' in the graphs. A huge number of 'Flux N/A' shows the limitations of our current telescopes. *Figures A-1.2 to A-1.6* all conform to the agreement that hot exoplanets dominate all of rocky, super earth, sub-neptune, sub-jovian, and jovian exoplanets by far. Since hot exoplanets are likely to reside closer to the host star, this trend implies that our current telescopes are more keen and sensitive to detecting exoplanets orbiting closer to the host star.

The scatter plots can show a more unique trend for the characterization of exoplanets. *Figure B-1.1* shows the correlation between flux, orbital semi-major axis, and stellar effective temperature. This figure is plotted in the idea that stellar flux on the planet is dependent on the temperature of the host star and the separation between host star and the planet. Specifically, the insolation flux should be greater with greater stellar effective temperature and shorter semi-major axis, as the insolation flux is computed by the inverse square law [13] [4] where a is the semi-major axis and the luminosity $L/L_{\odot}$ is computed by equation (3), which is proportionally related to the stellar effective temperature by a power of 4. *Figure B-1.1* plot clearly confirms this equation as exoplanets at the top left (higher stellar effective temperature and shorter orbit semi-major axis) are marked by higher insolation flux (color yellow).

Figure B-1.2 shows the correlation between (Orbit Semi-Major Axis)^3 and (Orbital Period)^2 to test Kepler's third law on the actual confirmed exoplanet dataset. The Kepler's third law states that (Orbit Semi-Major Axis)^3 is directly proportional to (Orbital Period)^2 by the equation [14] where G is Newton's Gravitational Constant, P is the orbital period, M is the mass of the star, m is the mass of the planet, and a is the orbital radius (orbit semi-major axis). *Figure B-1.2* confirms this law by showing a linear trendline with a R2 value of 0.9967, which is very close to 1. Based on this observation, we can conclude that the trend from the orbit semi-major axis can also tell the trend from the orbital period, and vice versa. One outlier from this plot at around 17 (Orbit Semi-Major Axis)^3 and around 108 (Orbital Period)^2 is CFBDSIR J145829+101343 b and may be worth a future investigation.

Figure B-1.3 test the classification of exoplanets on two of the most important factors of exoplanet characteristics - planet mass and planet radius [15]. As the three discrete, yet connected, straight lines suggest, there is indeed a clear, linear, proportional correlation between planet mass and planet radius [16]: the greater the planet mass is, the longer the planet radius is. Rocky planets (blue) correlate with lower planet mass and planet radius, followed in order of super earth (red), sub neptune (yellow), sub jovian (green), and jovian (purple). Also, note that at around planet mass of 10^2, the positive straight line stops, and the negative line starts. This plot also convinces us that there could be more a refined classification scheme based on exoplanet mass, instead of that on radius that [5] uses, or the ratio between mass and radius [17] [18].

Based on the idea about planet mass radius ratio, I plotted *figures B-1.4* and *B-1.5* to test classification in relation to insolation flux and orbit semi-major axis. The mass radius ratio at the x-axis clearly proves that exoplanets can be classified into a certain range of mass radius ratio. Rocky exoplanets, clustered on the leftist side, can be categorized with lower mass radius ratio, whereas Jovian exoplanets, clustered on the far right side, can be categorized with higher mass radius ratio. Super Earth, sub-Neptune, and sub-Jovian exoplanets can also be classified with mass radius ratio - from lower to higher mass radius ratio. While there is no large trend with insolation flux, the exoplanets classified with larger mass radius ratio tend to have greater separation (orbit semi-major axis) from the host star.

2. Exoplanet Habitability Analysis

Next, I analyzed the habitability of currently confirmed exoplanets with bar graphs *A-2.1*, *A-2.2, A-2.3* and scatter plots *B-2.1* and *B-2.2*. Based on the habitable zone scheme [6], I calculated the number of exoplanets within the habitable zone and made a bar graph *A-2.1*. The bar graph *A-2* shows that only around 5 percent of currently confirmed exoplanets (231 out of 4434) are habitable according to the scheme proposed in "Habitable Zones Around Main-Sequence Stars: New Estimates". Then, I classified these habitable exoplanets into rocky, super earth, sub neptune, sub jovian, and jovian exoplanets based on the classification scheme [5] and generated bar graphs *A-2.2* and *A-2.3*. These two graphs show that jovian is the most common habitable exoplanet classification and that rocky is the least common habitable exoplanet classification not only by number but also by percentage. I plotted *B-2.1* and *B-2.2* to show trends within habitable exoplanets. It turns out that the habitable exoplanets almost perfectly follow the mass radius ratio lines, identified in *figure B-1.3* and have relatively lower insolation flux, from 0.1 to 10 Earth Flux. This range is closer to the Earth's own flux, 1 Earth Flux, indicating that the right amount of flux is a significant factor of habitability of the exoplanets.

3. Exoplanet Detection Method Analysis

Finally, I analyzed the detection method with bar graph *A-3* and *B-1 to B-3* plots. First, bar graph *A-3* shows that transit method detected the most exoplanets, thanks to transit missions like Kepler Space Telescope and Transiting Exoplanet Survey Satellite (TESS), followed by radial velocity, microlensing, imaging, and pulsar timing.

On the other hand, plots *B-3.1* to *B-3.3* show the sensitivity of the detection methods. Plot *B-3.1* and *B-3.2* show a large cluster of transit method plots (orange) concentrated at a relatively small planet mass (1 ~ 30 Earth Mass), planet radius (1~5 Earth Radius) and orbital period (1 ~ 10^2 Earth Days). This conforms with the sensitivity of transit method; since transit method detects the change in starlight blocked by exoplanets, and exoplanets orbiting closer to the host

star are more likely to block the starlight by greater area, it is more likely to detect exoplanets with smaller orbit semi-major axis. Since orbit semi-major axis is related to orbital period by Kepler's third law, as the plot *B-1.2* shows, transit method is more sensitive to exoplanets with relatively shorter orbital periods [1]. However, this intuition fails for the cluster of shorter planet-radius exoplanets: the transit method should be more sensitive to exoplanets with greater radius because bigger exoplanets will block more of the starlight. This suggests that more smaller exoplanets reside closer to the host star. The trend with planet mass directly follows the trend with planet radius by our observation of mass radius ratio from *figure B-1.3*. Another notable trend is the cluster of radial velocity plots (blue) concentrated at a relatively longer planet radius (10 to 15 Earth Radius) and heavier planet mass (102 to 104 Earth Mass). This shows the sensitivity of radial velocity exoplanets on heavier exoplanets, as radial velocity is detecting the gravitational pull between host star and the planet and the gravitational pull between host star and planet proportionally depends on their mass [1]. The trend with planet radius directly follows the trend with planet radius by our observation of mass radius ratio from *figure B-1.3*. Lastly, we can note that the direct imaging plots (green) are concentrated at higher planet radius (12 to 15 Earth Radius) and longer orbital period (10^4 to 10^5 Earth Days). This is because our current direct imaging method is not sensitive enough: it is limited to exoplanets around nearby stars with very large radius and longer separation from stars [1]. Trends for other detection methods are not that clear due to lack of exoplanets confirmed by those methods.

Plot *B-3.3*, which is a skymap of all the confirmed exoplanets, shows a large concentration of yellow-colored exoplanets around 300 degree RA and 50 degree DEC, which coincides with the Kepler Space transit Telescope field of view [19], as well as a large concentration of red-colored exoplanets around 275 degree RA and -30 degree DEC. On the other hand, radial velocity and imaging seem to detect exoplanets regardless of their location.

Caveat and Future Research

Since the study of exoplanets is a relatively new field, there exist some caveats and limitations in this paper. The three most notable caveats are the limitation of data, disregard for the uncertainties, and the complex nature of exoplanets and their habitability. Along with the identification of these caveats, I also propose some future works that can remedy these caveats.

1. Limitation of Data

Since the search for exoplanets has been underway for only around two decades, the current datasets are greatly limited in three ways. Firstly, most of the 4434 confirmed exoplanets are in a relatively small, concentrated region of the Milky Way galaxy because that is as far as current telescopes have been able to probe. These 4434 exoplanets represent less than a 0.000004434 percent of the planets within our Milky Way galaxy as it has been shown that there are at least 100 billion exoplanets in the Milky Way galaxy [20]. Therefore, the survey of demographics in this paper only applies to these confirmed 4434 exoplanets, so trends and analysis found in the plots may be completely irrelevant as more exoplanets are detected and confirmed. Secondly, some of the data were missing for some exoplanets. For example, more than 1600 exoplanets missed their insolation flux data, so only around 2800 exoplanets were further classified into "Hot", "Warm", and "Cold". In order to obtain more complete demographics of exoplanets, future missions can be undertaken to find missing data of currently confirmed exoplanets by detecting them using different detection methods or by detecting them with more powerful, overarching telescopes such as the James Webb Telescope, which will be launched in November this year [21]. Moreover, future research can be done by manipulating more variables than the 10 variables I used, as well as with data of candidate exoplanets, which include almost twice as many exoplanets (7,472) as the confirmed exoplanets [3]. Lastly, some of the data may be underestimated because most indirect detection methods, especially radial velocity, are heavily dependent on the orientation of the planetary system. These orientation-dependent detection methods only show the component of the velocity in the observer's direction, leading to the underestimation of data. Indeed, the mass from the NASA Exoplanet Archive catalog show

the minimum value, and NASA acknowledges this caveat [22] [23]. Future work can be done to improve the high-contrast direct imaging technology to find more accurate exoplanet data because direct imaging is independent of the orientation of the planetary systems [24].

2. Disregard for Uncertainties

Another caveat of my research is that I ignored the uncertainties when plotting graphs in order to simplify the plots. Since the main goal of this paper is to survey the general demographical trends of currently confirmed planets, uncertainties are negligible, especially because uncertainties are really small compared to the actual values. However, future surveys may be done to include these uncertainties to provide a more complete representation of the data.

3. Complex Nature of Exoplanet and Habitability

The last caveat to note is the complexity of exoplanets and habitability. There is no one official exoplanet classification scheme and habitability scheme. Instead, each exoplanet has to be individually examined to find out the information about its classification and habitability. Although this paper follows two of the most respected schemes, they may not represent exoplanets' actual classification and habitability. In particular, the definition of current habitable zones is made in respect to the condition of Earth. However, different creatures may live in conditions different from Earth such as the dependency on CO^2 and H_2O [25]. Moreover, different planetary systems, such as pulsar [26] and binary [27], may require a completely different habitable zone scheme due to their completely different environments.

Therefore, while this paper employs habitable zone schemes cited by more than 1000 other papers, and exoplanets deemed habitable by proposed habitability schemes may be worth further probes, it is by no means the perfect formula to determine if a planet hosts life or not. Future research can be done to apply different classification and habitable zone schemes based on the data and code presented in my paper.

Conclusion

The NASA Exoplanet Archive is an effective open catalog that provides data of both confirmed and candidate exoplanets. However, there are some limitations in this catalog such as lacking information about rough classification and habitability of the exoplanet. Throughout this research, I have analyzed 10 variables of 4434 confirmed exoplanets from the NASA Exoplanet Archive by generating 20 different plots and updated the catalog by including classification and habitability flags. I analyzed the data in three different ways, by examining the characterization and classification of exoplanets, habitability of exoplanets, and sensitivity of exoplanet detection methods.

I employed four different open Python packages – namely, Pandas, NumPy, Matplotlib, and Astropy – in reading, writing, analyzing data and generating plots. The analysis methodology is implementation-friendly, as all the codes, plots, updated catalogs, and dataset are publicly uploaded on my GitHub page for readers to repeat my analysis whenever new confirmed exoplanets are added on the NASA Exoplanet Archive.

The main take-aways from the analysis of the dataset include the confirmation of flux-stellar temperature-separation relation, Kepler's third law, mass radius ratio and correlation, habitability on 'hot' and Jovian classified exoplanets, and sensitivity of transit, radial velocity, and direct imaging detection methods.

As more exoplanets are detected with more extensive missions, such as the James Webb Space Telescope and the Roman Telescope [28], many more exoplanets with more diverse environments will be detected, getting humans closer to the goal of finding other planetary life. Therefore, there will be more need for sophisticated, comprehensive habitability and classification schemes. I end this paper by proposing a possible exoplanet classification scheme that uses a weighted average method for future purposes, when there are more advanced, comprehensive telescopes to obtain more complete exoplanet metadata. Although some scientists argue that classification should be based on easily detectable characteristics of exoplanets and based on the fewest possible criteria [29], I

noticed that there are many different important factors and correlations in categorizing exoplanets, such as planet mass, planet radius, insolation flux, and semi-major axis. Some combination of weighted average of these variables may lead to a more sophisticated classification scheme because the world of exoplanets is so complex that they cannot be categorized just by a single variable; variables that are deemed more important in the classification of exoplanets may take greater weight than do other variables. Although current detection methods are quite limited to what types of variables they can accurately detect, such as the limitations of accurate planet mass data for transit method and the limitation of accurate planet radius data for radial velocity method [14], more advanced future telescopes will allow us to find more comprehensive data of exoplanets and to obtain more complete knowledge on the world of exoplanets.

Acknowledgments

I thank Dr. Dan Sirbu, a Research Scientist and Engineer at BAERI, NASA Ames Research Center, for being my mentor on this project. The completion of this project would not have been possible without his guidance, review, and feedback.

I also acknowledge that this research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program.

References

1. Seager, S. (Ed.). (2011). Exoplanets. Tucson, AZ: The University of Arizona Press.

2. National Aeronautics and Space Administration. (2019, June 13). Historic Timeline. Exoplanet Exploration: Planets Beyond Our Solar System. Retrieved July 11, 2021, from https://exoplanets.nasa.gov/alien-worlds/historic-timeline/#first-exoplanets-discovered

3. National Aeronautics and Space Administration. (2021, July 17). *Discoveries Dashboard*. Exoplanet Exploration: Planets Beyond Our Solar System. https://exoplanets.nasa.gov/discovery/discoveries-dashboard/

4. National Aeronautics and Space Administration. (2021, April 28). *In Depth | Confirmed vs. Candidate*. Exoplanet Exploration: Planets Beyond Our Solar System. https://exoplanets.nasa.gov/what-is-an-exoplanet/in-depth/

5. Kopparapu, R. et al. (2018). Exoplanet Classification and Yield Estimates for Direct Imaging Missions. The Astrophysical Journal, 856(2). 10.3847/1538-4357/aab205

6. Kopparapu, R. et al. (2013). Habitable Zones Around Main-Sequence Stars: New Estimates. The Astrophysical Journal, 765(2). 10.1088/0004-637X/765/2/131

7. NASA Exoplanet Archive. (2021). Planetary Systems Composite Data [Data set]. Retrieved from https://exoplanetarchive.ipac.caltech.edu/cgi-bin/TblView/nph-tblView?app=ExoTbls&config=PSCompPars. 10.26133/NEA13

8. NASA Exoplanet Archive. (2020, August 4). Extended Planet Data Table Data Column Definitions. NASA Exoplanet Archive. Retrieved July 11, 2021, from https://exoplanetarchive.ipac.caltech.edu/docs/API_exomultpars_columns.html

9. Robitaille, T. P., Tollerud, E. J., Greenfield, P., Droettboom, M., Bray, E., Aldcroft, T., Davis, M., Ginsburg, A., Price-Whelan, A. M., Kerzendorf, W. E., Conley, A., Crighton, N., Barbary, K., Muna, D., Ferguson, H., Grollier, F., Parikh, M. M., Nair, P. H., Günther, H. M., . . . Streicher, O. (2013). Astropy: A community Python package for astronomy. Astronomy & Astrophysics, 558(33). https://doi.org/10.1051/0004-6361/201322068

10. Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., Lim, P. L., Crawford, S. M., Conseil, S., Shupe, D. L., Craig, M. W., Dencheva, N., Ginsburg, A., VanderPlas, J. T., Bradley, L. D., Pérez-Suárez, D., de Val-Borro, M., Aldcroft, T. L., Cruz, K. L., Robitaille, T. P., Tollerud, E. J., Ardelean, C., . . . Zabalza, V. (2018). The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package.

The Astronomical Journal, 156(3). https://doi.org/10.3847/1538-3881/aabc4f

11. Planetary Habitability Laboratory. (2013, March 24). Summary of the Limits of the New Habitable Zone. Planetary Habitability Laboratory @ UPR Arecibo. https://phl.upr.edu/library/notes/summarylimitsofthenewhabitablezone

12. Szyk, B. (2018, January 31). Luminosity Calculator. Omni Calculator. https://www.omnicalculator.com/physics/luminosity

13. NASA Exoplanet Archive. (2016, November 2). *How Predicted Observables for Exoplanets are Calculated*. https://exoplanetarchive.ipac.caltech.edu/docs/poet_calculations.html#insol_flux

14. Simon Fraser University. (n.d.). *Calculating exoplanet properties - President's Dream Colloquium - Simon Fraser University*. Retrieved July 19, 2021, from https://www.sfu.ca/colloquium/PDC_Top/astrobiology/discovering-exoplanets/calculating-exoplanet-properties.html

15. Ulmer-Moll, S., Santos, N. C., Figueira, P., Brinchmann, J., & Faria, J. P. (2019). Beyond the exoplanet mass-radius relation. *Astronomy & Astrophysics*, 630, A135. https://doi.org/10.1051/0004-6361/201936049

16. Sotin, C., Grasset, O., & Mocquet, A. (2007). Mass–radius curve for extrasolar Earth-like planets and ocean planets. *Icarus, 191*(1), 337–351. https://doi.org/10.1016/j.icarus.2007.04.006

17. Seager, S., Kuchner, M., Hier□Majumder, C. A., & Militzer, B. (2007). Mass□Radius Relationships for Solid Exoplanets. *The Astrophysical Journal*, 669(2), 1279–1297. https://doi.org/10.1086/521346

18. Swift, D. C., Eggert, J. H., Hicks, D. G., Hamel, S., Caspersen, K., Schwegler, E., Collins, G. W., Nettelmann, N., & Ackland, G. J. (2011). MASS-RADIUS RELATIONSHIPS FOR EXOPLANETS. The Astrophysical Journal, 744(1), 59. https://doi.org/10.1088/0004-637x/744/1/59

19. Batalha, N. M., Borucki, W. J., Koch, D. G., Bryson, S. T., Haas, M., Brown, T. M., Caldwell, D. A., Hall, J. R., Gilliland, R. L., Latham, D. W., Meibom, S., & Monet, D. G. (2010). SELECTION, PRIORITIZATION, AND CHARACTERISTICS OF KEPLER TARGET STARS. *The Astrophysical Journal*, 713(2), L109–L114. https://doi.org/10.1088/2041-8205/713/2/l109

20. National Aeronautics and Space Administration. (2013, June 7). Billions and Billions of Planets. NASA. https://www.nasa.gov/mission_pages/kepler/news/kepler20130103.html.

21. Kalirai, J. (2018). Scientific discovery with the James Webb Space Telescope. *Contemporary Physics*, 59(3), 251–290. https://doi.org/10.1080/00107514.2018.1467648

22. Otegi, J. F., Bouchy, F., & Helled, R. (2020). Revisited mass-radius relations for exoplanets below 120 M⊕. *Astronomy & Astrophysics*, 634. https://doi.org/10.1051/0004-6361/201936482

23. Changeat, Q., Keyte, L., Waldmann, I. P., & Tinetti, G. (2020). Impact of Planetary Mass Uncertainties on Exoplanet Atmospheric Retrievals. *The Astrophysical Journal*, 896(2). https://doi.org/10.3847/1538-4357/ab8f8b

24. Jovanovic, N., Guyon, O., N'Diaye, M., Galicher, R., Sirbu, D., Kenworthy, M. A., Ygouf, M., Baudoz, P., Kühn, J., Huby, E., Wilby, M. J., Por, E. H., Haffert, S. Y., Keller, C. U., Snik, F., Miller, K. L., Wallace, J. K., Beaulieu, M., Cady, E., . . . Riggs, A. J. E. (2018). Review of high-contrast imaging systems for current and future ground-based and space-based telescopes: Part II. Common path wavefront sensing/control and coherent differential imaging. Adaptive Optics Systems VI, 107031. https://doi.org/10.1117/12.2314260

25. Ramirez, R. (2018). A More Comprehensive Habitable Zone for Finding Life on Other Planets. Geosciences, 8(8), 280. https://doi.org/10.3390/geosciences8080280

26. Patruno, A., & Kama, M. (2017). Neutron star planets: Atmospheric processes and irradiation. *Astronomy & Astrophysics*, 608. https://doi.org/10.1051/0004-6361/201731102

27. Whitmire, D. P., Matese, J. J., Criswell, L., & Mikkola, S. (1998). Habitable Planet Formation in Binary Star Systems. Icarus, 132(1), 196–203. *https://doi.org/10.1006/icar.1998.5900*

28. National Aeronautics and Space Administration. (2020, November 10). Mission. Exoplanet Exploration: Planets Beyond Our Solar System. https://exoplanets.nasa.gov/discovery/missions

29. Stern, S. A., & Levison, H. F. (2002). Regarding the Criteria for Planethood and Proposed Planetary Classification Schemes. *Highlights of Astronomy*, 12, 205–213. https://doi.org/10.1017/s1539299600013289

The Impact of Decreasing Dataset Sizes on Frozen Layer Transfer Learning

By Michael Zheng

Author Bio:

Pengen Zheng, who goes by Michael, is currently a senior at William P. Clements High School located in Sugar Land, Texas. He finds the field of computer vision to be very fascinating and hopes to pursue future studies in machine learning-based computer vision.

Abstract

Transfer learning is a machine learning training method where a model trained for one task gets trained a second time for another task that is usually related yet different from the original task. Applying transfer learning during machine learning training helps to decrease the training time and computational resources needed, as the model has already been pretrained when it was trained for the first task. Transfer learning also helps to compensate for problems like underfitting that might occur when a machine learning model is being trained on a small dataset. This study aims to identify correlations between the effectiveness of transfer learning, specifically frozen layer transfer learning, and the amount of data provided. We accomplish this by applying transfer learning to a machine learning model and feeding it gradually decreasing amounts of data from a data set and observing the model accuracy. We observe that as the amount of provided data decreases, frozen layer transfer learning becomes increasingly less effective.

Keywords: Machine learning, computer vision, transfer learning, frozen layer, YOLOv3, ExDark dataset, computer science, mAP@0.5

Introduction

In recent years, machine learning applications have been increasingly prominent in many different fields. However, with the application of machine learning comes the need to train a machine learning model or multiple models towards a specific goal. During this process, a few problems might arise. For one, depending on the size and complexity of a machine learning model, training such models from the ground up might prove to use a lot of computational resources and time. In some instances, the amount of data in a dataset available for a machine learning model to use for training might be relatively small. If those small datasets are used during ground-up training, it might lead to underfitting, in which a machine learning model is unable to identify dominating trends in a dataset.

A common solution to the aforementioned problems is to apply transfer learning when training a model. Transfer learning is when a model that was originally trained for one task is applied and trained for another different but related task. As the tasks share some similarities, the changes that need to be made to the model when transitioning from the first task to the second task would generally use less computational resources and time compared to training a model for the second task from the ground up. Furthermore, if the amount of training data available for the second task is small, transfer learning helps to compensate for the downfalls in that much of the network would have been set up when the model was trained for the first task, assuming there were large amounts of data available during the process of training the model for the first task. This way, instead of worrying about underfitting, the small amount of data for the second task would act as fine tuners to the model in the transfer learning process. An example of transfer learning might be that a model that was originally trained to identify cars is being re-purposed to identifying trucks, which is a different but related task as there are similar characteristics between a car and a truck.

In our study, we look at a specific transfer learning technique called frozen layers. Frozen layers in transfer learning refer to when a model that has been trained for a task is being trained for a second task, the weights of selected layers won't be changed during the second training process. When layers are frozen, fewer weights need to be modified during backpropagation, which helps to improve the training speed and decrease the computations needed and allows data to make more impact on the unfrozen layers compared to during normal transfer learning. However, given that frozen layers are not updated during the process, this technique may bring decreased accuracy.
In this paper, we will look at how effective frozen layers transfer learning is at compensating for small datasets. We achieve this by training the YOLOv3 Object Detection model (Redmon et al., 2018) using different subsets of the Exclusively Dark (ExDark) Image Dataset (Loh & Chan, 2019). We then measure the final Mean Average Precision of the model after each training as our metrics of evaluation.

Model Training

Algorithm, Dataset, and Training Environment

In this study, we are using YOLOv3, a Real-Time Object Detection system, as our machine learning model. YOLOv3 consists of a Darknet-53 backbone responsible for feature extraction and a YOLOv3 head section responsible for object detection. For our purposes, we are using a PyTorch implementation of YOLOv3 created by Ultralytics (Jocher et al., 2021). In terms of the image dataset, we are using the Exclusively Dark (ExDark) dataset. ExDark is a dataset consisting of 7,363 images in low light environments with 12 different object classes. The ExDark dataset provides enough variability in object classes while maintaining a relatively small number of data. Our models were trained using Google Colab.

Training

We want to accurately observe the effect the size of a dataset would have on task effectiveness. To do so, we first separated the ExDark dataset so that 80 percent of all images are used for training purposes while the other 20 percent are used for validation purposes. The images were separated in a way that

would yield approximately equal distribution of images of different object classes in both the training and validation dataset. After the training images have been separated from the validation images, we then randomly segment the training set into 10 different subsets with equal numbers of images.

To prepare for frozen layer transfer learning, we first train our YOLOv3 model from scratch with the COCO 2017 (Lin et al., 2015) dataset for 300 epochs. COCO 2017 is a dataset with 91 object classes and about 2,118,000 images. Given the large number of images in COCO 2017 and the fact that all object classes in ExDark are a subset of COCO 2017's object classes, this makes COCO 2017 a good candidate as a dataset for the initial training.

After we train our YOLOv3 model from scratch with COCO 2017, we can apply frozen layer transfer learning by first freezing the Darknet-53 backbone, or layers 0 through 10 in the Ultralytics YOLOv3 implementation. This way, we can use transfer learning to further train the parts of YOLOv3 responsible for object detection while leaving the feature extraction as is.

Table 1
Training Configurations and Corresponding Number of Subsets and Epochs

Training configuration	Number of Subsets	Number of Epochs
1	1	1100
2	2	1000
3	3	900
4	4	800
5	5	700
6	6	600
7	7	500
8	8	400
9	9	300
10	10	200

With the Darknet-53 backbone frozen, we train our YOLOv3 model for 10 different training configurations. We first feed our YOLOv3 model 1 subset of the 10 previously segmented ExDark training data and train it for 1100 epochs at 16 images per batch. Then, we reset YOLOv3 to the state immediately after the backbone is frozen and feed it 2 out of the 10 ExDark training subsets and train it for 1000 epochs. These steps are then repeated 8 more times, each time adding one more subset and decreasing the epoch count by 100, until all 10 subsets are being fed into the model, as seen in Table 1. The number of images per batch should remain at 16 images. It should be noted that every new training configuration should have the exact same subsets as before with the exception of the newly added subset. This helps to ensure that for every training configuration except the first configuration, the training data is a strict superset of the previous configuration.

In real-world applications, a machine learning model shouldn't be trained for an extensive amount of epochs as it may lead to the model overfitting. However, for our purposes, we are training each training configuration for more epochs than necessary to extensively observe potential patterns and trends.

Results and Discussion

For our measurement, we will be using specifically mAP@0.5, or the Mean Average Precision for when the IoU threshold is 0.5 and above. In Table 2 we can see the final mAP@0.5 for each training configuration. As expected, as we increase the amount of data provided to the model, the resulting mAP@0.5 increases. Taking a look at Table 3, we find that with increasing amounts of data being provided during training, the increase in final mAP@0.5 shows a generally decreasing trend. This implies that as the amount of data provided decreases, the effectiveness of frozen layer transfer learning becomes more negatively impacted.

Table 2
Amount of Data Provided and the Resulting mAP@0.5

Percentage of Data Provided	Resulting mAP@0.5
0.1	0.1380
0.2	0.1912
0.3	0.2172
0.4	0.2512
0.5	0.2738
0.6	0.2901
0.7	0.3027
0.8	0.3301
0.9	0.3403
1.0	0.3479

Table 3
Final mAP@0.5 Difference Between Each Consecutive Training Configuration

Domain of Percentage of Data Provided	Δ Resulting mAP@0.5
0.1	0.1380
0.2	0.1912
0.3	0.2172
0.4	0.2512
0.5	0.2738
0.6	0.2901
0.7	0.3027
0.8	0.3301
0.9	0.3403
1.0	0.3479

In Figure 1, we graph the progression of mAP@0.5 of each training configuration. Specifically, after each epoch, we calculate the mAP@0.5 by testing each training configuration against the evaluation dataset. We can observe that as the amount of data provided during training decreases, it takes longer for the corresponding configuration's mAP@0.5 to stop fluctuating. The time it takes for a model to stop fluctuating is a reflection of how quickly a model can reach a stage of robustness. From Figure 1, we notice that the less data that's accessible to a model during training, the longer it takes for that model to become robust.

Figure 1
The mAP@0.5 of Each Training configuration During Training.

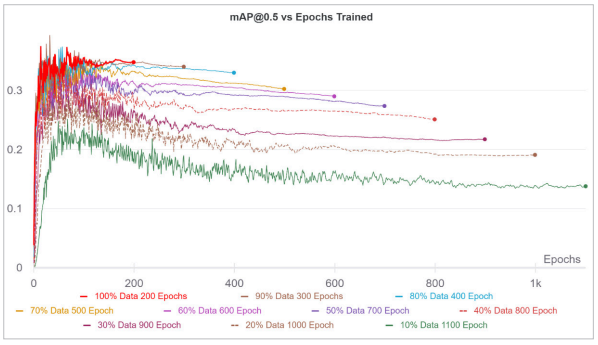


Figure 2
Figure 1 Smoothed Using Running Average Smoothing

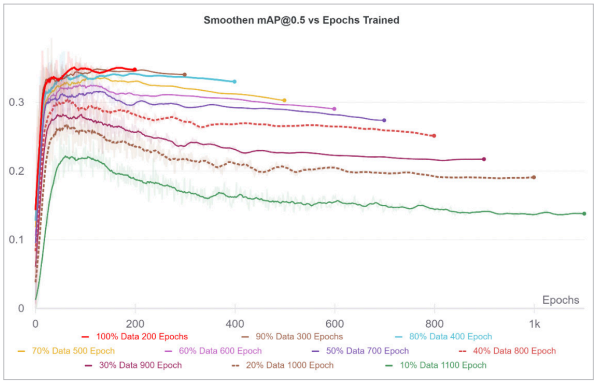


Table 4
Difference Between Highest Final mAP@0.5 after 200 epochs in Figure 2

Percentage of Data Provided	Δ mAP@0.5 Drop Off
0.1	0.0337
0.2	0.0305
0.3	0.0256
0.4	0.0202
0.5	0.0179
0.6	0.0149
0.7	0.0053
0.8	0.0003
0.9	0.0030
1.0	0.0031

As shown in Table 4, as more data is provided to the model during training, the model generally experiences less mAP@0.5 drop-off in the first 200 epochs. We can thus deduce that as less data is provided to a model, the model could overfit easier.

Conclusion

Using the YOLOv3 Object Detection system and the ExDark dataset, we have found that by applying frozen layers transfer learning to machine learning model training, it helps to effectively compensate for minor decreases to the amount of data in a small dataset. However, as more data is removed from a small dataset, frozen layers transfer learning becomes decreasingly effective. Specifically, with decreasing data provided to a model during training, we can determine that frozen layer transfer learning becomes more and more negatively impacted as the model experiences a larger drop-off in mAP@0.5, takes a longer time to reach a stage of robustness, and becomes easier to overfit. In a real-world application, based on time and computational constraints, one could, if needed, reference our results to find a fitting decrease to the transfer learning training dataset that could preserve enough transfer learning effectiveness to produce a final trained model that could meet their expectations.

It should be noted that observations made in this study are based on one specific machine learning model and one specific dataset. In future studies of this project, we plan on implementing the experiment on different combinations of models and datasets outside of YOLOv3 and ExDark to reinforce our conclusion.

References

Jocher, G., Kwon, Y., Guigarfr, perry0418, Veitch-Michaelis, J., Ttayu, Suess, D., Baltacı, F., Bianconi, G., IlyaOvodov, Marc, e96031413, Lee, C., Kendall, D., Falak, Reveriano, F., FuLin, GoogleWiki, Nataprawira, J., … Shead, T. M. (2021, April 12). *Ultralytics/Yolov3: V9.5.0 - YOLOv5 v5.0 RELEASE compatibility update FOR YOLOV3*. Zenodo. Retrieved from https://doi.org/10.5281/zenodo.4681234.

Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., & Dollár, P. (2015, February 21). *Microsoft coco: Common objects in context*. arXiv.org. Retrieved from https://arxiv.org/abs/1405.0312.

Loh, Y. P., & Chan, C. S. (2019). Getting to know low-light images with the exclusively dark dataset. *Computer Vision and Image Understanding*, 178, 30–42. https://doi.org/10.1016/j.cviu.2018.10.010

Redmon, J., & Farhadi, A. (2018, April 8). *Yolov3: An incremental improvement*. arXiv.org. Retrieved from https://arxiv.org/abs/1804.02767.

Determining the Fractal Dimension of Crumpled Paper using Chi-Squared Data Analysis and Model Testing

By Larry Hui

Author Bio:

Larry Hui is a current high school student at Crescent School located in Canada who would like to pursue Mechanical Engineering with a concentration in Aerospace and Aeronautical Engineering. He is a beginner researcher who, under the guidance of Dr. Carey Witkov, has just recently completed one paper regarding the fractal dimensions of crumpled paper using chi-squared model testing. Outside of his academic interests, he is a holder of a private pilot's licence and working on both his night and instrument ratings. He likes to read short fiction and classics and spend time playing classical music on the trumpet and the piano.

Abstract

Most real objects in the natural world are fractals (i.e., non-smooth). Fractal dimensions are non-integer and therefore, by example, may lie between that of a plane (two-dimensional) and a solid object (three-dimensional). Determining the dimension of fractal objects can help describe objects for which suitable geometric representations may be unavailable. The purpose of this study is to determine the fractal dimension of crumpled paper using chi-squared data analysis and model testing. This offers several improvements over linear regression including better parameter estimation, better estimation of parameter uncertainties and model testing.

Crumpled paper does not fully occupy the space it is embedded in, so its fractal dimension is estimated to be between that of a plane and a solid sphere (i.e., between 2 and 3). Data were obtained from photographs of crumpled paper of various masses as the independent variable and the diameter of each crumpled ball was repeatedly measured for the dependent variable. Chi-squared data analysis of a two-parameter linear model gave a best-fit slope estimate of the fractal dimension as 2.17, and chi-squared model testing showed that the model should not be rejected.

Keywords: Fractal, Geometry, Dimension, Chi-Squared, Regression, Estimation, Paper, Crumpled

Nomenclature

χ^2	Chi-squared test statistic
m	Relative mass of crumpled paper ball
ρ	Density of the crumpled paper ball
R	Radius of the crumpled paper ball
n	Fractal dimension of crumpled paper ball
A_{Best}	Best fit slope parameter
B_{Best}	Best fit y-intercept parameter
$P(x)$	Probability density function of the Gaussian
$P(\chi^2)$	Probability density function of the chi-squared for continuous variables
SE	Standard error of the mean
N	Number of data points on the plane
σ	Standard deviation
μ	Population mean
$\bar{y}$	Mean radii
$\bar{x}$	Mean relative mass
$SE_{\bar{y}}$	Standard error of the mean radii

Section I: Introduction to the Study

The problem in adequately describing shapes, forms of curves and surfaces has been a contentious and thought-provoking matter for hundreds of years. Euclidean geometry fails to adequately represent many real-world objects, for example, the coastline of a country is neither straight nor elliptical. Imagine attempting to describe rough or uneven terrain in terms of classical geometry—it cannot easily be done; if at all. However, in 1979, the French polymath Benoit Mandelbrot proposed the interesting idea of using a single number as a measure of roughness, a statistical complexity showing how complicated a self-similar object is, which he called the "fractal dimension" of an object. This idea was used to alleviate the vagueness caused by the absence of any suitable geometric representations (Mandelbrot, 1980). A fractal dimension characterizes curves and surfaces in terms of their complexity by treating dimension as a continuum and not a discrete value.

Normally, dimensions belong to the positive set of integers, Z+; for example, the 1st dimension is a line, the 2nd dimension is a plane, and the 3rd dimension is a real-world object (Devaney, 1995). However, fractal dimensions are non-integer and may lie between the aforementioned examples. Fractal dimension, therefore, is a measure of space-filling and can be related to other physical properties, e.g., rigidity, which is relevant to crumpled paper.

Although knowledge of fractal properties has been around since the beginning of the 20th century, Mandelbrot was the first to recognize their applications outside of mathematics and related them to economics, chaos theory, statistical physics and more (Mirowski, 1990). This paper aims to obtain the fractal dimension of crumpled paper using a chi-squared algorithm written in Python. The main result of this paper is that chi-squared analysis improves on ordinary linear regression for determining fractal dimensions by including measurement uncertainties and model testing in the algorithm.

A. Inquiry Statement, Research Approach, and Hypothesis

Using chi-squared analysis, we can determine if the derived family of models should be rejected or not. If the model is not rejected, how do we determine the fractal dimension of crumpled paper? Based on the fact that the crumpled paper does not fully occupy the space it is embedded in, a general hypothesis we could assume is that the fractal dimension of crumpled paper should be between 2 (a plane) and 3 (a solid object). Indeed, the fractal dimension that was found (2.2) is between 2 and 3.

The research approach used for the chi-squared analysis is as follows: (i) derive the model, perform the experiment, (ii) determine the best-fit parameters, (iii) test if the best-fit is a good fit, (iv) find the uncertainties on the parameters, (v) test if the model should be rejected, and lastly, (vi) determine

if there should be a revised and improved model. The following chi-squared analysis in Eq. 1 was used to perform the model testing using chi-squared for continuous variables as shown below:

$$\chi^2 = \sum_{i=1}^{n} \frac{(y_i - y_m)^2}{\sigma_i^2} \quad (1)$$

While this resembles a weighted least squares, where the contribution of each data point is inversely weighted by its uncertainty, weighted least squares, like ordinary least squares, do parameter estimation but not model testing. Model testing usually involves the use of ad hoc methods (e.g., correlation coefficients, F-tests, confidence intervals, etc.) whereas chi-squared analysis combines parameter estimation and model testing in one methodology (Witkov & Zengel, 2019). The chi-squared test was not done by hand but instead done in a program written in Python. An explanation of what the program does and how it computes the chi-square value will be explained later in this paper. So we ask the question: why use chi-squared analysis rather than a linear regression?

Using chi-squared to test two quantitative variables is not seen as often as two categorical variables since the chi-squared test usually involves testing models which include the number of Observed (O) and Expected (E) occurrences of certain event(s) as evident in Eq. 2.

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \quad (2)$$

This is not, however, the form of chi-squared used for testing continuous models. Chi-squared is used to assess our model because it can be shown that it provides optimal estimates of model parameters for normally distributed errors; in this case, these are—slope *A* and *y*-intercept *B*. In using chi-squared, we should be asking the question, given the model, how probable is the data? This leads to the definition of chi-squared, which, when minimized, provides a maximum likelihood estimate of the best slope and y-intercept. As shown in **Appendix A**, the best-fit parameters are those that minimize chi-squared. This is because the minimum value of chi-squared corresponds to the maximum likelihood that, given the model, the data would be generated.

The most important feature of chi-squared curve-fitting versus Gomes' linear regression method (Gomes, 1987) is that chi-squared, by its definition, includes measurement uncertainties (note that the square of the standard error appears in the denominator of chi-squared). For example, a good fit to the data could be a bad fit to the model. A good model fit requires that the minimum value of 2 should be of the order of *N*, the number of data points. This ensures that the weighted sum of squares between the data and model is within one standard error (Witkov & Zengel, 2019). By including uncertainties, chi-squared is able to combine parameter estimation and model testing in one consistent methodology.

B. System Model and Derivation

For solid objects, we assume a power-law relationship between mass and radius of the crumpled paper, but allow for fractional exponents as the crumpled paper is a fractal object. In the following power-law relation, m is the mass of the paper, R is the radius, ρ is the density, and *n* is the fractal dimension:

$$m = \rho R^n \quad (3)$$

To solve for the fractal dimension of the crumpled paper, we need to solve Eq. 3 for n. Taking logarithms on both sides of the equation, we obtain:

$$\log log\, m = \log log\, \rho + n \log log\, R \quad (4)$$

Rearranging for log R, note that the equation is a linear function in the form y = m(A)x + B.

$$\log log\, R = \frac{1}{n} \log log\, m - \frac{1}{n} \log log\, \rho \quad (5)$$

Eq. 5 is now in the form of the following two-parameter linear model

$$y = A_{Best} x + B_{Best}, \text{ where } n = \frac{1}{A_{Best}} \quad (6)$$

While the mass of the crumpled paper is a function of its radius, it is convenient to express the model in Eq. 5 using radius (log *log R*) as the dependent variable (*y*) and mass (log *log M*) as the independent variable (*x*).

Section II: Experimentation and Data Analysis

To perform this experiment, we used the crumpled paper image in Fig. 1 taken from Yale University's Fractals website, which involves two square sheets each of 8 by 11-inch paper with relative masses of 1, ½, ¼, and ⅛, respectively (Yale, n.d). The raw data was extracted using the online geometry software GeoGebra on a photo of four crumpled balls of different masses. Using GeoGebra, we estimated the diameters of the crumpled paper by taking three measurements: the longest diameter, the shortest diameter, and one in between to reduce sampling variability.

Figure 1 Paper balls of relative mass used for experiment.

Table 1: Raw Data

Crumpled Ball	Relative Mass	Radii (cm)
1	1.0	2.8, 3.2, 3.4
2	0.5	2.2, 2.5, 2.6
3	0.25	1.4, 1.7, 2.0
4	0.125	0.9, 1.3, 1.3

Table 2: *Radii,* $\bar{x}$, $\bar{y}$, *and* $SE_{\bar{y}}$

Crumpled Ball	$\bar{x}$	Radii (cm)	$\bar{y}$	$SE_{\bar{y}}$
1	1.0	2.8, 3.2, 3.4	3.1	0.06
2	0.5	2.2, 2.5, 2.6	2.4	0.05
3	0.25	1.4, 1.7, 2.0	1.7	0.10
4	0.125	0.9, 1.3, 1.3	1.2	0.11

In Table 1 and Table 2, N is the number of data points which is equal to the number of independent variables: 4. $\bar{y}$ is the mean radii, $\bar{x}$ is the mean mass, and $SE_{\bar{y}}$ is the standard error of the mean radii $\bar{y}$.

A. Probabilistic Basis for Chi-Square Model Testing

Looking at the probability density function (*pdf*) of the normal distribution, we get the following equation *P(x)*, σ represents the standard deviation, and μ represents the mean:

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (7)$$

The normal distribution can thus be represented as a model for the 2-parameter linear equation:

$$P(x) \propto e^{-\sum_{i=1}^{n} \frac{(y_i - y_{model_i})^2}{2\sigma^2}} \quad (8)$$

Eq. 8 is based on the difference of each data point yi and the corresponding model value y_{modeli} where the model is the two-parameter linear model *y=Ax+B*. After transforming the normal distribution into a model, we now have a probabilistic basis for chi-squared model testing.

The Central Limit Theorem allows us to construct a normal distribution centered around each data point. We can then ask, given the model, what is the probability of generating all the data? Starting with the probability of generating one point of data and using the probability rule for a sequence of independent events, $PA \bigcap B = PA \times P(B)$, we obtain for the probability of generating all the data:

$$P(all\ data) = e^{-\frac{(y_1 - y_{model_1})^2}{2\sigma_1^2}} \times e^{-\frac{(y_1 - y_{model_2})^2}{2\sigma_2^2}} \ldots \times e^{-\frac{(y_1 - y_{model_n})^2}{2\sigma_n^2}} = e^{-\sum_{i=1}^{n} \frac{(y_i - y_{model_i})^2}{2\sigma_i^2}} \quad (9)$$

Defining chi-squared $\chi^2 = \sum_{i=1}^{n} \frac{(y_i - y_m)^2}{\sigma_i^2}$ to simplify Eq. 9, we get the following result:

$$P(Generating\ Data \mid The\ Model) \propto e^{-\frac{\chi^2}{2}} \quad (10)$$

This means that the probability of generating the data will be greatest, given the model, when 2is least. The normalized probability density function for chi-squared (**Appendix C**) that will be used later to find probability contours is given by:

$$P(\chi^2) = \frac{1}{2} e^{-\chi^2/2} \quad (11)$$

B. Description of the Chi-Squared Script

The Python script does chi-squared curve fitting and tests the two-parameter linear model *y=Ax+B*. Three arrays must be populated to use the script. These are array *x*, which is the array of $\bar{x}$ (mean masses), array *y*, which is the array of $\bar{y}$ (mean radii), and *yerr*, which is the array of the standard errors of the mean radii (error bars) where σ represents the standard deviation and *n* represents the number of samples shown in Eq. 12.

$$SE = \frac{\sigma}{\sqrt{n}} \quad (12)$$

N represents the total number of data points (i.e., the number of values for the independent variable). *sigma_r* is the standard error of the radii; chi-squared is the sum of the squared differences between data and model, weighted inversely by the square of measurement standard errors.

The process for computing chi-squared in this script takes advantage of the fact that min χ^2, A_{Best}, and B_{Best} can be obtained in closed-form for linear models. The reason is that chi-squared for two parameters is a paraboloid whose minimum can be obtained by an equation rather than a search.

The python script outputs two plots. The first plot is a log-log data plot whose slope is the fractal dimension. Error bars represent standard errors as defined in Eq. 12. The second plot is a contour plot in the A-B parameter plane and displays two contours.

The inner contour represents the range of parameter values that, given the model with parameter values within this contour, there is a 68% likelihood that the data would have been generated. The outer contour corresponds to a 95% likelihood. The likelihood of the two parameters in the experiment follow a 2probability density function for two measurements as stated in Eq. 11. Thus, to calculate the contours we need to find χ^2_{68} and χ^2_{95}. To obtain the value χ^2 that encloses 95% of the likelihood, we integrate the probability density function in Eq. 11:

$$0.95 = \int_0^{\chi^2_{95}} P(\chi^2)d(\chi^2) = \int_0^{\chi^2_{95}} \frac{1}{2}e^{-\chi^2/2}d(\chi^2) \quad (13)$$

So, we can solve $\chi^2_{68} = 2.3\sigma$, and $\chi^2_{95} = 6.2\sigma$. Now, we can build the contour with $\chi^2_{min} + 2.3\sigma$ and $\chi^2_{95} + 6.2\sigma$ which is highlighted in the code. This is especially helpful if we have a parameter reference value. If that reference value is outside the yellow line (2σ), then we should reject the model.

Section III: Results, Discussion, and Conclusions

From the outset, the main goal of this paper was to investigate the fractal dimension of crumpled paper using chi-squared instead of ordinary least-squares. Although least-squares generates feasible results, chi-squared provides a more accurate insight into the fractal dimension including better parameter estimation and model testing criteria to reject the model. Fig. 4 below shows the results generated by the chi-squared script that does curve fitting to the 2-parameter linear model.

The results generated include A_{Best}, the fractal dimension (the reciprocal of A_{Best}), B_{Best}, min χ^2, N value, and the good fit min χ^2 range. **Fig. 4** shows that a fractal dimension of 2.17 is obtained.

Figure 2. log(m) and log(R)

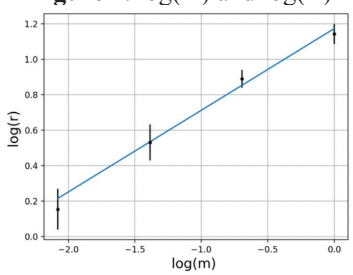


Figure 3. min χ^2 contour

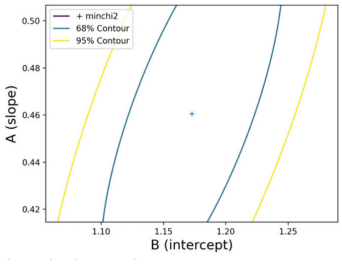


Figure 4. Chi-squared analysis results

```
CHI-SQUARED ANALYSIS RESULTS
A_best = 0.46
fractal dimension = 2.17
B_best = 1.17
minchi2 = 1.11
N = 4
SigmaR = [0.17638342 0.12018504 0.17320508 0.13333333]
Y Error = [0.05629258 0.04939111 0.10188534 0.11428571]
minchi2 good fit range = [ 1.17 , 6.83 ]
chi2 for 68% Contour = 3.41
chi2 for 95% Contour = 7.31
```

This makes sense as fractal dimensions cannot be whole integers and the crumpled paper has dimensionality between a plane $n = 2$, and a sphere $n = 3$. This is because the paper used in **Fig. 1** is likely more rigid than standard sheets of printing paper. The min 2 is 1.47 and the min 2 within a good fit range is [1.17, 6.83], meaning that it falls within the range. Referring to **Fig. 2**, the error bars or variance for the radii for the smaller ball *r4* is much larger than the rest of them. This is most likely due to measurement error in GeoGebra as zooming in on bigger images is easy and zooming in on smaller images creates inaccuracies. However, because our fractal dimension is 2.17 and the min χ^2 value of 1.47 falls within the good fit range from 1.17 to 6.83, so we should not reject this model. The good fit range was determined using the following:

$$-\ GFR = N - \sqrt{2N} \text{ and } +\ GFR = N + \sqrt{2N} \text{ since } \chi^2_{min} \approx N \quad (14)$$

Looking at **Fig. 3**. to find out what min χ^2 is, we simply find the intersection of A_{Best} and B_{Best} as both are quadratic functions in a one-dimensional plot, with χ^2 the y-axis and A(slope) or B(intercept) as the x-axis, shown in **Fig. 2** and **3**. Knowing this, χ^2 will always appear at the vertex of the parabola where $x=A_{Best}$ *or* B_{Best}

There is a slant to the shape of the paraboloid since we expect there to be a high correlation coefficient value. If there was no slant, the change in the *y*-intercept would not affect the slope which is not the case. Furthermore, looking at all the points mentioned above, the model should not be rejected.

A. Conclusion, Applications, Limitations

In conclusion, chi-squared analysis was used for optimal parameter estimation of the fractal dimension and model testing. In doing so, a reasonable value for the fractal dimension, 2.17, was obtained. Looking back at the model, we could also find B_{Best} represented by ρ as density. However, we are only interested in A_{Best} as the reciprocal of A_{Best} being the fractal dimension. Using the chi-squared code has helped simplify the process a lot rather than doing the chi-squared analysis by hand as well. The code has some drawbacks including flexibility of the code only on linear models. A potential improvement would be to make a more universal chi-squared test that could apply to multiple situations. Error in testing was reduced by taking three measurements all of different lengths. However, using GeoGebra did have some downsides as it was not as exact as we were hoping for. By actually taking the mass of the balls and using a caliper, for example, we could get more precise results. Another way of increasing the precision of obtaining an accurate fractal dimension is to increase the number of sample measurements taken. Looking at the bigger picture, however, fractals and especially chi-squared model testing can provide a unique solution to many problems and should be considered whenever practical, especially in new fields of research. Furthering this study, a potentially interesting approach to solving this problem could be increasing the number of parameters or using a non-linear model to express the fractal dimension.

Acknowledgements

The Python code was written by Dr. Carey Witkov (former Preceptor) and Keith Zengel, Preceptor. Harvard University, Physics Department, Cambridge, MA. I thank and acknowledge assistance from Eric Arsenault and Dr. Carey Witkov for making comments to improve the readability of the manuscript and for providing their expertise in the field.

References

Carey Witkov, K. Z. (2019). *Chi-Squared Data Analysis and Model Testing for Beginners*. Oxford University Press.

Carey Witkov, K. Z. (2020). *Chi-Squared script in Python public software release*. Retrieved from Github: https://github.com/witkov/chi-squared

Crumpled Paper Lab. (n.d.). Retrieved from Yale Mathematics: https://users.math.yale.edu/public_html/People/frame/Fractals/Labs/CrumpledPaperLab

Devaney, R. L. (1995, April 2). *Fractal Dimension*. Retrieved from Boston University Mathematics and Statistics: https://math.bu.edu/DYSYS/chaos-game/node6.html

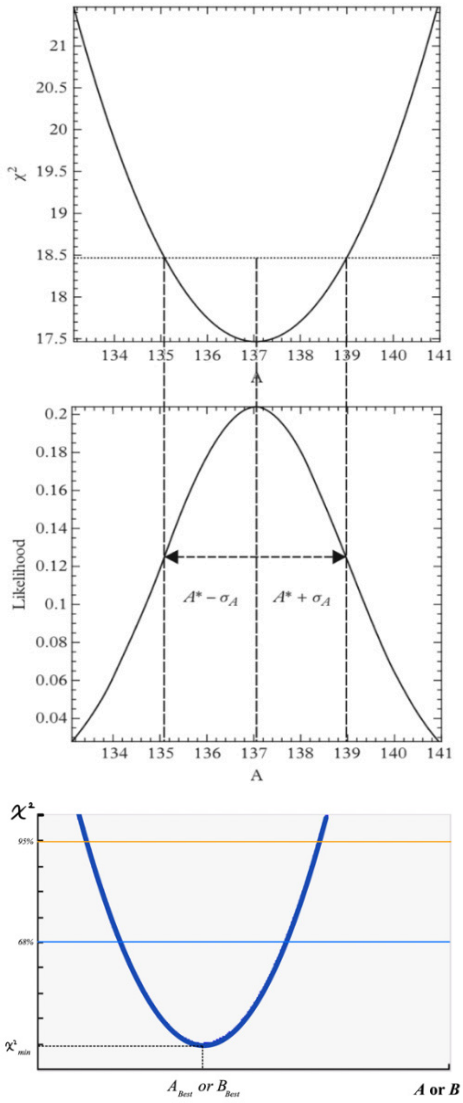
Gomes, M. (1987). Fractal Geometry in Crumpled Paper Balls. *American Journal of Physics*.

Gomes, M. (1987). Paper Crushes Fractally. *Journal of Physics A: Mathematical and General*.

Mandelbrot, B. B. (1981). Fractals and the Geometry of Nature. *Yearbook of Science*.

Mirowski, P. (1990). From Mandelbrot to Chaos in Economic Theory. *Southern Economic Journal*.

APPENDIX A: Gaussian vs. Chi-Squared Distribution (Witkov & Zengel, 2019)



APPENDIX B: Chi-Square Curve Fitting for 2-Parameters

```
# 2dchi2.py version 5/12/20
# This script does chi-square curve fitting to the
2-parameter linear model y = Ax + B
# SOFTWARE DEPENDENCIES: Numpy, Matplotlib
# ----------------------------------------------------------------
# Copyright (C) 2020 Carey Witkov and Keith Zengel
 # This program is free software: you can redistribute it
and/or modify it under
# the terms of the GNU General Public License as
published by the Free Software
# Foundation; either version 3 of the License, or (at
your option) any later version.

# This program is distributed in the hope that it will be
useful, but WITHOUT ANY WARRANTY;
# without even the implied warranty of
MERCHANTABILITY or FITNESS FOR A
PARTICULAR PURPOSE.
# See the GNU General Public License for more
details.

# You should have received a copy of the GNU
General Public License along with this program.
# If not, see https://www.gnu.org/licenses/.
#----------------------------------------------------------------

import matplotlib.pyplot as plt
from numpy import *
import numpy as np

# relative mass of crumpled paper ball
m = np.array([1, 0.5, 0.25, 0.125])

# Radii of crumpled paper balls
r1 = np.array([3.4, 3.2, 2.8])
r2 = np.array([2.5, 2.2, 2.6])
r3 = np.array([2, 1.7, 1.4])
r4 = np.array([1.3, 1.3, 0.9])

r = np.array([mean(r1), mean(r2), mean(r3),
mean(r4)])

# number of data points
N = len(r)
# number of samples for each data point
n = len(r1)
sigma_r = np.array([np.std(r1, ddof=1), np.std(r2,
ddof=1), np.std(r3, ddof=1), np.std(r4, ddof=1)]/
np.sqrt(n));

# x,y data arrays and y-error array
x = log(m);
y = log(r);
yerr = sigma_r/r;
```

```
# calculate sums needed to obtain chi-square
Syy=sum(y**2/yerr**2)
Sxx=sum(x**2/yerr**2)
S0=sum(1/yerr**2)
Sxy=sum((y*x)/yerr**2)
Sy=sum(y/yerr**2)
Sx=sum(x/yerr**2)

A_best = (S0*Sxy - Sx*Sy)/(S0*Sxx - (Sx*Sx))
B_best = (Sy*Sxx - Sx*Sxy)/(S0*Sxx - (Sx*Sx))
minchi2 = Syy+(S0*(Sxy*Sxy) - 2*Sx*Sy*Sxy +
Sxx*(Sy*Sy))/((Sx*Sx) - (S0*Sxx))
sigmaA = 1/sqrt(Sxx)
sigmaB = 1/sqrt(S0)

# create parameter grid for the minchi2
a = np.linspace(Abest-
0.2*Abest,Abest+0.2*Abest,500)
b = np.linspace(Bbest-
0.2*Bbest,Bbest+0.2*Bbest,500)
A,B = np.meshgrid(a,b)

# calculate chi-square over parameter grid (scan and
find using a matrix type calculation)
chi2 =(Syy) + (A**2)*(Sxx) + (B**2)*(S0) -
2*A*Sxy - 2*B*Sy + 2*A*B*Sx

# plot data with error bars
plt.figure()
plt.grid(True)
plt.errorbar(x,y,yerr,linestyle='None',fmt='.k')
plt.xlabel('log(m) ',fontsize=16)
plt.ylabel('log(r) ',fontsize=16)

# plt.title('y vs x data and line-of-best-fit',
fontsize=16)
plt.plot(x,Abest*x+Bbest)

# plot chi-square in A-b parameter plane with 68% and
95% contours
plt.figure()
levels=[minchi2,minchi2+2.3,minchi2+6.2]
Z=plt.contour(B,A,chi2,levels)
plt.ylim(Abest-0.1*Abest,Abest+0.1*Abest)
plt.xlim(Bbest-0.1*Bbest,Bbest+0.1*Bbest)
plt.plot(Bbest,Abest,'+')
plt.ylabel('A (slope)',fontsize=16)
plt.xlabel('B (intercept)',fontsize=16)

# add legend
artists, labels = Z.legend_elements()
plt.legend(handles=artists, labels=["+ minchi2","68%
Contour","95% Contour"], loc='upper left')
plt.show()

# display results
print("CHI-SQUARED ANALYSIS RESULTS")
Abest_float = "{}

# display results
print("CHI-SQUARED ANALYSIS RESULTS")
Abest_float = "{:.2f}".format(Abest)
FractalDimension = "{:.2f}".format(1/Abest)
print("A_best =", Abest_float)
print("fractal dimension =", FractalDimension)
Bbest_float = "{:.2f}".format(Bbest)
print("B_best =", Bbest_float)
minchi2_float = "{:.2f}".format(minchi2)
print("minchi2 =", minchi2_float)
print("N =", N)
print("SigmaR =", sigma_r)
print("Y Error =", yerr)
N_min_float = "{:.2f}".format(N-math.sqrt(2*N))
N_max_float = "{:.2f}".format(N+math.sqrt(2*N))
print("minchi2 good fit range = [", N_min_float, ",",
N_max_float, "]")
contour68 = "{:.2f}".format(levels[1])
print("chi2 for 68% Contour =", contour68)
contour95 = "{:.2f}".format(levels[2])
print("chi2 for 95% Contour =", contour95)
```

APPENDIX C: Derivation of Probability Density Function

The derivation of the probability density function (*pdf*) for two measurements x and y follows. We need to take the combined multiple probability density function for both *x* and *y*

$$C(x,y) = C(x)C(y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2} - \frac{(x-\mu_y)^2}{2\sigma_y^2}}$$

Since we are only interested in contours of equal probability in the function CC(*x*,*y*) which are defined as $\chi^2_x + \chi^2_y$ which is equal to some constant τ:

$$\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(x-\mu_y)^2}{\sigma_y^2} = \chi_x^2 + \chi_y^2 = \tau$$

So we are interested in how the probability changes when jumping from one contour of equal probability to another. We write this probability in terms of

$$\chi = \sqrt{\chi_x^2 + \chi_y^2}$$

Using polar coordinates, we change *x* and *y* to χ_x and χ_y respectively:

$$\iint G(x)G(y) = \iint G(x)G(y)\frac{dx}{d\chi_x}d\chi_x\frac{dy}{d\chi_y}d\chi_y = \iint \frac{1}{2\pi} e^{-\frac{\chi_x^2}{2}+\frac{\chi_y^2}{2}} d\chi_x d\chi_y$$

We introduce an angular coordinate,ϕ, to convert cartesian coordinates to polar coordinates so we can rewrite the double integral as:

$$\iint \frac{1}{2\pi} e^{-\frac{\chi_x^2}{2}} \chi d(\chi) d(\phi) = \int e^{-\chi^2/2} \chi d(\chi)$$

Changing the variables from $\chi \rightarrow \chi^2$ and $d\chi = d\chi^2/(2\chi)$

$$P(\chi^2) = \frac{1}{2} e^{-\chi^2/2}$$

A Comparative Analysis of Public Policy Regarding Vaccination in America and China

By Daisy Lin

Author Bio:

Daisy Lin is a senior studying at Stuyvesant High School in New York City. She first heard of the coronavirus outbreak through relatives living in China during late 2019 and has since been interested in the different ways that the pandemic has been handled. In the United States alone there is strong rhetoric either supporting or opposing preventative measures like vaccines, lockdowns, quarantine, and more. After researching the vaccination policy for the United States of America and the People's Republic of China, she realized the stark difference between the two. In the future, she hopes to continue studying various illnesses along with how they are controlled and treated in different regions.

Abstract

COVID-19, originating from Wuhan, China, quickly spread around the world over the course of a few months during late 2019 and early 2020 and has become a worldwide pandemic. Since then, nations have been in search of numerous viable treatments for nearly two years, including vaccinations. This comparative analysis of second-hand documents published by the United States of America and the People's Republic of China reveals the differences between the two nations in regard to their public policy regarding coronavirus vaccination. China has taken a more strict, direct approach, while the United States has relied on more lenient methods. These differences are likely due to the nations' distinct social and political environments. Given the stark contrasts in these procedures, it is crucial to further explore the reasoning behind the variation and efficacy of each technique to gain relevant insight on the way future pandemics should be handled to minimize their spread and devastation.

Keywords: Coronavirus, COVID-19; Vaccine, Vaccination; Public Policy; Herd Immunity; Delta Variant; Covid America vs China; Government Covid Regulation; Vaccine Policy in America; Vaccine Policy in China

Introduction

Coronaviruses are a group of viruses covered in spike proteins resembling the shape of a crown, or "corona" in Latin. Specifically, coronavirus disease 2019 (COVID-19) is a contagious, airborne disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) first identified in Wuhan, China in December 2019. In an attempt to protect their citizens from infectious disease, countries like the United States of America and China have launched mass immunization programs intending to vaccinate large numbers of people at different locations over a short period to establish herd immunity. Currently, the three main vaccines being administered in the United States are Pfizer, Moderna, and Johnson & Johnson, and in China, there are two main vaccines, the CoronaVac and Sinopharm.

Methodology

Information in this article comes from secondary sources published through online articles that use both quantitative and qualitative data accessed either through Google Scholar or Jstor with keyword searches including "coronavirus," "herd immunity," "covid vaccine efficacy," "American vaccine policy," and "Chinese vaccine policy." Certain issues arose during the research process. For example, there was difficulty finding information from the Chinese perspective, specifically a lack of official government publications, when compared to the American, which may lead to certain gaps of knowledge. However, this was largely substituted with anecdotal accounts from Chinese citizens and the analysis of laws enforced by their government. By using both government and citizen accounts from the United States of America and the People's Republic of China, a better understanding of public policy and the people's response may be achieved.

[1] https://wwwnc.cdc.gov/eid/article/25/1/17-1901_article

Herd Immunity

The main reason why vaccines are an important public health tool is due to their ability to provide herd immunity. Herd immunity refers to resistance against the spread of a virus within a population based on pre-existing immunity in a high proportion of individuals either by vaccination or previous infection. The basic reproduction number, or R_0, is an indicator of how infectious disease may be and what percentage of the population must be vaccinated for herd immunity to be achieved. A higher basic reproduction number would mean that an illness is highly transmittable while a lower number would indicate a lower risk of transmission. It is generally accepted that once 75-80% of the population is immune, herd immunity will be preserved[2]. Herd immunity is most vital for those who are the most vulnerable as certain individuals may be barred from getting a vaccine due to their age or previous health conditions, increasing their susceptibility to infection. This idea is also applicable to the current pandemic since certain groups remain prone to COVID-19. To effectively protect the entire population, governments in the United States and China need to be strategic with the distribution and administration of vaccines to prevent another wave of the virus from spreading.

2020-2021 Vaccines

The Pfizer-BioNTech COVID-19 vaccine (BNT162B1) and Moderna COVID-19 vaccine (mRNA-1273) were approved through an emergency use authorization (EUA) in the United States on December 11, 2020, and December 18, 2020[3], respectively. The Pfizer vaccine is recommended for individuals 12 years of age and older and has an efficacy rate, or a measure of how much the vaccine lowered the chance of an individual getting sick, of 52% after the first dose and 95% after the second dose, while the Moderna vaccine is recommended for individuals 18 years of age and older and has an efficacy rate of 50.8% after the first dose and 94.1%

[2] https://www.sciencedirect.com/science/article/pii/S1074761320301709

[3] https://www.cdc.gov/coronavirus/2019-ncov/vaccines/different-vaccines/Pfizer-BioNTech.html

[4] https://www.healthline.com/health/vaccinations/pfizer-vaccine-efficacy#real-world-efficacy

after the second dose[4]. Both vaccines rely on recently developed mRNA technology to provide immunity to patients. Unlike the flu shot which introduces a weakened version of the virus to the immune system, mRNA vaccines contain messenger RNA that encodes the antigen. Once introduced to the body, the mRNA is translated into antigen by the same process that cells use to create their proteins. The antigen is then moved to the cell membrane where it is exposed to immune cells which then analyze these harmless covid spikes and begin the formation of antibodies that will be required to combat the virus during infection. These antibodies have the unique ability to tag the coronavirus for degradation through a process known as phagocytosis, which involves immune cells ingesting the pathogen[5]. Certain mRNA vaccines also contain additional mRNA which codes for an enzyme that once replicated, can generate copies of the antigen-encoding mRNA, amplifying the production of antigen, increasing the efficiency and effectiveness of the vaccine. mRNA vaccines have proved to be promising for the future of immunology because they are developed in a cell-free system that is easier and safer to use, reduces the risk of contamination from other infectious pathogens, and lowers the risk of integration in the host genome. Initially, requirements for the transportation of mRNA vaccines made it difficult to guarantee effective administration of the vaccine as it required uninterrupted refrigeration at extreme temperatures. Originally it was believed that the Pfizer vaccine needed to be stored at -70°C and the Moderna vaccine at -20°C; however, recent studies have proved that the vaccines may be stored using regular refrigeration for over twenty days and maintain their properties. While both the Pfizer and Moderna vaccines have been effective, other vaccines with emergency authorization have faced allegations of serious side effects of which the general public was unaware.

The Jansen COVID-19 vaccine (JNJ-78436735) was approved through an emergency use authorization (EUA) on February 21, 2021, and is recommended for individuals 18 years of age and older though distribution has since been halted due to claims that the vaccine has led to blood clots in women under 50. After an investigation was launched, it was found that there was no connection between the clotting and vaccine, which led to the resumption of its administration on April 27, 2021[6]. Unlike the Pfizer and Moderna vaccines which rely on mRNA technology, the Johnson & Johnson, also known as Jansen, vaccine relies on adenosine virus viral vectors with the original viral matter removed to hinder its ability to replicate and cause disease in the patient. The coronavirus gene is then added to the viral vector and injected into the body and once injected, the virus will deliver the spike gene into the nucleus directly where antigen will be coded for. The newly coded antigens are then presented to Helper T cells which direct the immune system to create the antibodies needed to fight the virus after infection. Similar to mRNA technology, the viral vector technology uses the body's cells to create antigens rather than having them be injected artificially, reducing the risk of the patient falling ill with the disease after the administration of the vaccine.

The two leading coronavirus vaccines in China are the Sinovac COVID-19 vaccine (CoronaVac) developed by the Chinese company Sinovac Biotech and Sinopharm COVID-19 vaccine (BBIBP-CorV) developed by Sinopharm's Beijing Institute of Biological Products. Unlike the vaccines being distributed in the United States, both are conventional vaccines that use inactivated viral matter delivered in two doses[7]. These inactive vaccines contain viral particles that have been grown in culture then killed to bar their ability to replicate but the inactive particles still encourage the production of T cells. Once the proper T cells have been produced, they are stored as immune cells and can be used during an actual infection of the coronavirus. Since these vaccines contain inactive viruses, they cannot cause disease, hence they do not have to be stored at cooler temperatures, easing transportation and distribution[8]. Recently, a new vaccine produced in China trade-named Convidecia (AD5-nCOV) that may be administered using only one dosage has been introduced. Like the Johnson & Johnson vaccine, Convidecia is a viral vector vaccine that has a 65.7% efficacy rate in preventing COVID-19 symptoms and a 91% efficacy rate in preventing other types of serious illnesses[9]. One of the major benefits of single-dose

[5] https://www.future-science.com/doi/full/10.4155/tde-2020-0129

[6] https://www.cdc.gov/coronavirus/2019-ncov/vaccines/different-vaccines/janssen.html

[7] https://www.bmj.com/content/373/bmj.n912

[8] https://www.sciencedirect.com/science/article/pii/B9780128021743000023

[9] https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1885-642X2021000100016

vaccines is their relief on the administration process as it effectively prevents the situation of individuals receiving only partial immunity if they received only one of the two doses. The promising benefits of the Convidecia vaccine have granted it a EUA in certain countries outside of China in Asia, Europe, North America, and South America.

While vaccines administered in both America and China have made tremendous strides towards decreasing COVID transmissions, the American vaccines have had higher efficacy rates, meaning the Pzifer, Moderna, and Johnson & Johnson vaccines are more effective compared to CoronaVac and Sinopharm. This is likely due to the extended period of time used to study and develop the American vaccines. Additionally, this may mean that Pzifer, Moderna, and Johnson & Johnson are more effective at preventing another outbreak caused by COVID-19 variants.

Though the aforementioned vaccines have played a key role in the decreased transmission of COVID-19, new variants of the disease have been detected that cannot yet be fully protected against. On November 5, 2020, the UK had enforced a lockdown to control rising coronavirus cases, which seemed to work. However, there was still an increasing rate of infection in Kent, an area outside of London. It was later discovered that a new, more transmittable and deadly strain of the coronavirus, B. 1. 1. 7, also known as the Alpha variant, had mutated in Kent and has since become the dominant form of COVID-19. Since the discovery of this strain, other variants of the SARS-CoV-2 virus named the Variants of Concern by the World Health Organization, have been identified: B. 1. 1. 7 (Alpha variant), B. 1. 351 (Beta variant), P. 1 (Gamma variant), and B. 1. 617. 2 (Delta variant)[10]. Since these variants all have spike protein mutations, it is more likely that they can be transmitted and reinfect those who have already had COVID-19. Recent studies have proved the Delta variant to the most dangerous as it is twice as infectious as the other variants and is now the leading COVID-19 virus in the United States, with those infected showing the most severe symptoms[11]. Though it is unlikely that any current form of the virus will be able to completely evade vaccines, it is not impossible[12]. Given enough time and replicative cycles, the virus may evolve to bypass current medical technology. It is important to note that as long as the virus is still spreading, new strains will be forming, making it crucial that vaccines are administered to decrease the amount of time that SARS-CoV-2 has to mutate.

Vaccine Policy in the United States of America

Although being vaccinated is highly recommended in America, the Centers for Disease Control and Prevention (CDC) claim that vaccines approved under a EUA cannot be mandated as they are technically still in experimental trials. However, this claim has yet to be tested in court making the recent COVID vaccines, like many other vaccines in the United States, voluntary. This is not true in all cases as it is important to note that private-sector employers are generally allowed to use whatever criteria they deem fit for hiring, including an individual's vaccination status. For example, many essential workers were highly recommended by their employers to get vaccinated and due to their employer, they were considered first for the vaccine, followed by the elderly. Currently, this has yet to pose a threat to many as most employers are expected to find "reasonable accommodations," even in a high-risk job such as healthcare, unless it causes "significant difficulty or expense." Additionally, Title VII of the Civil Rights Act of 1964 prohibits any discrimination in the workplace based solely on religion, which may protect some individuals from vaccinations[13]. While the individual may not be a member of a major religious group like Christianity or Buddhism, there must be a religious reason rather than personal beliefs to be protected by Title VII[14]. Generally, citizens are encouraged to educate themselves on the vaccination process and make the decision that is right for them, though this idea has led to the clashing of anti-vax and pro-vax groups.

Due to the freedom granted to the American people in terms of optional vaccines, there has been the formation of various groups that either support

[10] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7825912/

[11] https://www.cdc.gov/coronavirus/2019-ncov/variants/delta-variant.html

[12] https://www.nature.com/articles/s41577-021-00556-5

[13] https://heinonline.org/HOL/LandingPage?handle=hein.journals/umijlr12&div=21&id=&page=

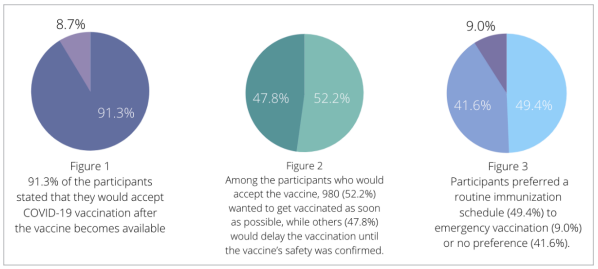
[14] https://ajph.aphapublications.org/doi/full/10.2105/AJPH.2020.306166

the vaccine, pro-vax, or oppose the vaccine, anti-vax. Despite pro-vax groups being larger in number, there are still advantages in anti-vax groups as there is generally a greater presence of anti-vax individuals in groups who are undecided, allowing them to easily spread their influence. Qualitative analyses have also shown that anti-vax groups tend to offer a wide range of potentially attractive narratives regarding safety concerns, conspiracy theories, alternative health and medicine, and potential substitutes to traditional vaccines, and these narratives have garnered large amounts of support, particularly during times of crisis[15]. Yet, despite these supposed advantages, over half of the adult population in the United States is fully vaccinated and many others are in the process, calling into question the influence that anti-vax groups have over the general population. When comparing the opinions of the American people with the opinions of the Chinese on optional vaccinations, there has been a large disparity.

Vaccine Policy in the People's Republic of China

China seems to be generally less divided regarding its attitudes towards vaccines and, beginning in the early 2000s, there have been many advancements in vaccination. With the introduction of the diphtheria, tetanus, and acellular pertussis vaccine, pneumococcal conjugate vaccine, Hib vaccine, Enterovirus 71 vaccine, and HPV vaccine, the standards of healthcare have also been expanded to now meet World Health Organization requirements[16]. The vaccine has also become more accessible to the general population with the passing of laws such as the "Vaccine Distribution and Vaccination Regulation" which stipulates that all national-level EPI vaccines be paid for by the government and the new essential public health service package includes 11 service categories including vaccinations, making them more accessible[17]. Based on a study conducted by Peking University, a member of the elite C9 League of Chinese universities showed overwhelming support for vaccines. 91.3% (Figure 1) said they would accept the COVID-19 vaccination after it was made available to the public. Of this 91.3%, 52.2% (Figure 2) wanted to get vaccinated as soon as possible, demonstrating the positive attitude of the Chinese toward vaccine improvements[18]. In fact, 49.4% are in favor of scheduled vaccine appointments and have already begun the application process. Chinese Citizens on the Coronavirus Vaccines: Statistics Found in a Study by Peking University.

Figure 1. Chinese Citizens Accepting Vaccination (DL, 2021).
Figure 2. Chinese Citizens Preferred Vaccination Time (DL, 2021).
Figure 3. Chinese Citizens Preferred Method of Vaccine Administration (DL, 2021).



On December 15, 2020, China officially launched a vaccination program for the winter-spring period targeting several key groups, including essential workers who handle cold-chain products, customs officers, medical workers, transportation drivers, and market workers[19]. The second round of vaccines was offered to young, working-age people who were more able to contribute to society than the elderly. While the general trust in the vaccine and the vaccination process was ranked of high importance, with 81.9% agreeing or strongly agreeing that the vaccine is safe and 90% of respondents agreeing or strongly agreeing that vaccination was very important, there have been individuals who showed hesitancy towards getting the vaccine. To convince these individuals to get vaccinated, the Chinese government has implemented a system of benefits and punishments around the vaccine. They have included a variety of perks such as cash bonuses and store credits and in some cases, eggs and fresh produce were also gifted[20]. On the other

[15] https://www.nature.com/articles/s41586-020-2281-1?fbclid=IwAR3Z_Q2x2Xc--1zJ9-W6v5A-VIwaeWUZGU7wWIbMB07_ZEi8NdnaeENHHks

[16] https://link.springer.com/article/10.1007/s42533-021-00066-y

[17] https://www.sciencedirect.com/science/article/pii/S0264410X18313574

[18] https://www.mdpi.com/2076-393X/8/3/482

[19] https://link.springer.com/article/10.1186/s12916-021-01923-8

[20] https://www.proquest.com/docview/2508215061/56E880EC928C4C4CPQ/1?accountid=36166

hand, those who choose to not get vaccinated may face serious repercussions. For example, employers may choose to fire those who are unvaccinated and individuals may lose a variety of social privileges. Therefore, despite there being no specific law mandating vaccines for Chinese citizens, there is heavy coercion for them to do so[21].

Differences Between Policies from American and China

Though it has long been apparent that the United States of America and the People's Republic of China are distinct nations with varying forms of government, this difference has been additionally highlighted by the coronavirus situation along with vaccine administration and distribution, namely with reference to the different groups who are prioritized in both nations for vaccines and the role that the government plays in the vaccination process. In the United States, the elderly were prioritized for vaccines because they tended to exhibit more severe symptoms after infections and have the highest mortality rate. On the other hand, China prioritizes the working-age people to better keep the economy afloat and allow those who are required to be in social situations to do so. This was also an attempt to protect the elderly as those who are younger tend to travel more, which may increase the rate of transmission. The decision from the governments on which groups should be vaccinated underscores the unique way in which each government operates. The United States of America is a more idealistic country where morals and abstract ideas are highly valued, even when it contradicts reason. China, on the other hand, is a more pragmatic nation where realism is favored.

There is also a distinction between the two nations regarding the level of government involvement during the vaccination process. In America, individual states are allowed to make unique decisions regarding the intensity of their immunization programs. Even when vaccinations are available, officials tend to take a hands-off approach with vaccines where civilians are allowed to analyze the ethics and effectiveness of the vaccine before deciding to either get vaccinated or not. In China, however, there is a larger presence of the government throughout the vaccination process and individuals are heavily influenced by government incentives to receive the vaccine. This process parallels the government structure in the two nations: the United States of America has a two-party system where debate and conversation are encouraged while China has a one-party system where people are expected to be supportive of government decisions. Regardless of these differences, both nations have taken large strides in producing vaccines and getting large percentages of their population vaccinated.

In addition to the distribution of vaccines, the research and developmental stage for the vaccines also differs in the two counties. The USA focused primarily on a market-oriented model. While there were various government agencies involved in Operation Warp Speed (OWS), a public-private partnership initiated by the United States government to expedite the development and distribution of the COVID-19 vaccine, the project was largely upheld by private corporations. For example, binding agreements like Advanced Purchase Commitments, which guarantee a viable market for a product prior to its creation, generated monetary incentives for companies like Pfizer and Moderna to create a product that consumers desire. In contrast, China focused on a state-driven collaborative model. This method partitions roles to both market and the government during the vaccine development period. Like the market-oriented model, private companies under the state-driven model also have monetary incentives. Due to potential market failures created during public health crises, vaccine companies are more likely to act immediately to minimize its effects and maximize profits. The government in China also has a large role in the development process by strategically funding and concentrating resources to any particular policies deemed important. In the case of the Beijing municipal government, they agreed to fund Sinovac's manufacturing plant unconditionally to ensure that the company can produce at least 300 million doses of the COVID-19 vaccine annually. By working together, government and market are able to encourage collaboration across private and public sectors[22].

[21] https://www.ctvnews.ca/health/coronavirus/china-ramps-up-vaccination-drive-with-incentives-such-as-free-eggs-other-goods-1.5389907

[22] https://academic.oup.com/jtm/article/28/4/taab026/6149495?login=true

Over the course of the coronavirus pandemic, several countries have been competing to offer international leadership in vaccine diplomacy. Notably, China and the United States have been actively donating millions of vaccines to developing countries. Throughout the past few years, the international view of China has been extremely volatile ranging from scrutiny due to the belief that the virus was leaked from a lab in Wuhan to sympathy for the large number of casualties. However, this resentful and pitiful attitude towards the nation largely changed in April 2020 when China launched an international campaign to brand itself as a global health leader, which began by sending masks, medical teams, and testing kits overseas[23]. This strengthened China's public image as the "Health Silk Road" was able to provide aid to countless countries in need. The majority of these donations went to countries in Central and South America, Africa, South Asia, the Middle East, and Eastern Europe. Now, Chinese president, Xi Jinping, vows to provide the world with 2 billion doses of the COVID-19 vaccine and donate $100 million to COVAX by the end of 2021[24]. Likewise, President Joe Biden has made a pledge that the United States will donate 600 million vaccines to developing countries by the end of 2021 through boosting the rate of vaccine production in the United States, India, and South Africa. There has also been a proposed deal to purchase 500 million additional doses of the COVID-19 vaccine from[25]. Currently, the countries receiving the most vaccine donations from America include Pakistan, Bangladesh, the Philippines, Colombia, South Africa, Vietnam, Indonesia, Guatemala, Uzbekistan, and Nigeria. Though it is true that vaccine diplomacy is crucial for a nation's international perception, hence why the United States and China are likely competing to be leaders in this field, it is important to realize the impact that donated vaccines will have, especially for those living in impoverished conditions. It is therefore crucial for these wealthy nations to work together in order to produce the estimated 11 billion vaccines still needed to ensure global immunity.

[23] https://link.springer.com/article/10.1057/s41254-021-00224-4

[24] https://www.nytimes.com/2021/08/06/world/china-vaccine-donations-2-billion.html

[25] https://www.nytimes.com/2021/09/20/us/politics/biden-covid-19-vaccines.html

[26] https://www.pennmedicine.org/coronavirus/frequently-asked-questions-about-covid-19/additional-covid-questions#tab-1o

Additional Methods to Prevent the Spread of COVID-19

Over the past few months, COVID-19 infection rates have been on a downward trend, with upward trends of vaccinations making large contributions to this. However, there are additional steps that could be taken to further stop the spread of the virus such as the investment of better cleansing of public transportation. Specifically, the inclusion of UV lights as a part of the standard disinfectant procedure for air purification has shown promising results for inactivating the virus[26]. Like most viruses, COVID-19 spreads more efficiently in heavily populated areas, meaning that those who live in metropolitan areas of both the United States and China are at increased rates of infection. As public transportation tends to be a hotspot for infection, better cleansing would likely reduce transmission rates and with additional investments in antimicrobial goods, transmission rates should continue on a steep decline. With the investment in antimicrobial goods and surfaces in public spaces, the spreading of viruses may be reduced as people tend to unknowingly often be in contact with surfaces that may be riddled with germs and viruses.

Conclusion

While the emergence of the COVID-19 virus has led to many casualties, it has also forced many governments to reevaluate their attitudes towards public healthcare and their handling of public health matters. The initial difficulties faced during the distribution of vaccines have allowed both the United States and China to rethink healthcare to make it more effective, while also demonstrating the importance of public health and vaccine awareness. Though both nations have been inflicted with similar situations, the variation in governance has led the two to handle the situation extremely differently. Additionally, the massive funds provided for vaccines have pushed scientific breakthroughs such as mRNA technology which was previously difficult to actualize. With all of these advancements in vaccine research and distribution as well as the growing awareness of the general public towards the importance of public health, both the United States and China will be making large strides to better protect their citizens if future pandemics were to arise.

References

1. Complexity of the Basic Reproduction Number (R0) - Volume 25, Number 1-January 2019 - Emerging Infectious Diseases journal - CDC. (n.d.). Retrieved from https://wwwnc.cdc.gov/eid/article/25/1/17-1901_article

2. Randolph, H. E., & Barreiro, L. B. (2020, May 19). Herd Immunity: Understanding COVID-19. Retrieved from https://www.sciencedirect.com/science/article/pii/S1074761320301709

3. Pfizer-BioNTech COVID-19 Vaccine Overview and Safety. (n.d.). Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/vaccines/different-vaccines/Pfizer-BioNTech.html

4. Herndon, J. (2021, May 28). Pfizer Vaccine Efficacy: Research, Real World, and More. Retrieved from https://www.healthline.com/health/vaccinations/pfizer-vaccine-efficacy#real-world-efficacy

5. Shahcheraghi SH;Ayatollahi J;Aljabali AA;Shastri MD;Shukla SD;Chellappan DK;Jha NK;Anand K;Katari NK;Mehta M;Satija S;Dureja H;Mishra V;Almutary AG;Alnuqaydan AM;Charbe N;Prasher P;Gupta G;Dua K;Lotfi M;Bakshi HA;Tambuwala MM;. (n.d.). An overview of vaccine development for COVID-19. Retrieved from https://pubmed.ncbi.nlm.nih.gov/33624533/

6. Johnson & Johnson's Janssen COVID-19 Vaccine Overview and Safety. (n.d.). Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/vaccines/different-vaccines/janssen.html

7. Baraniuk, C. (2021, April 09). What do we know about China's covid-19 vaccines? Retrieved from https://www.bmj.com/content/373/bmj.n912

8. Siegrist, C., & Lambert, P. (2016, July 15). How Vaccines Work. Retrieved from https://www.sciencedirect.com/science/article/pii/B9780128021743000023

9. Tran, V. D., Pak, T. V., Gribkova, E. I., Galkina, G. A., Loskutova, E. E., Dorofeeva, V. V., . . . Pham., D. T. (n.d.). Determinants of COVID-19 vaccine acceptance in a high infection-rate country: A cross-sectional study in Russia. Retrieved from https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1885-642X2021000100016

10. Le Page, M. (2021, January 23). What are the new coronavirus variants? Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7825912/

11. Delta Variant: What We Know About the Science. (n.d.). Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/variants/delta-variant.html

12. Gupta, R. K. (2021, April 29). Will SARS-CoV-2 variants of concern affect the promise of vaccines? Retrieved from https://www.nature.com/articles/s41577-021-00556-5

13. About. (2021, March 08). Retrieved from https://heinonline.org/HOL/LandingPage?handle=hein.journals/umijlr12&div=21&id=&page=

14. The American Journal of Public Health (AJPH) from the American Public Health Association (APHA) publications. (n.d.). Retrieved from https://ajph.aphapublications.org/doi/full/10.2105/AJPH.2020.306166

15. Johnson, N. F., Velásquez, N., Restrepo, N. J., Leahy, R., Gabriel, N., El Oud, S., . . . Lupu, Y. (2020, May 13). The online competition between pro- and anti-vaccination views. Retrieved from https://www.nature.com/articles/s41586-020-2281-1?fbclid=IwAR3Z_Q2x2Xc--1zJ9-W6v5A-VIwaeWUZGU7wWIbMB07_ZEi8NdnaeENHHks

16. Zha, D. (2021, March 24). China and the global search for health security: History, vaccines, and governance. Retrieved from https://link.springer.com/article/10.1007/s42533-021-00066-y

17. Yu, W., Lee, L. A., Liu, Y., Scherpbier, R. W., Wen, N., Zhang, G., . . . Wang, H. (2018, November 26). Preventable disease control in the People's Republic of China: 1949–2016. Retrieved from https://www.sciencedirect.com/science/article/pii/S0264410X18313574

18. Wang, J., Jing, R., Lai, X., Zhang, H., Lyu, Y., Knoll, M. D., & Fang, H. (2020, August 27). Acceptance of COVID-19 Vaccination during the COVID-19 Pandemic in China. Retrieved from https://www.mdpi.com/2076-393X/8/3/482

19. Yang, J., Zheng, W., Shi, H., Yan, X., Dong, K., You, Q., . . . Yu, H. (2021, February 10). Who should be prioritized for COVID-19 vaccination in China? A descriptive study. Retrieved from https://link.springer.com/article/10.1186/s12916-021-01923-8

20. Beijing ramps up vaccine drive with incentives from eggs to shopping coupons. (2021, April 04). Retrieved from https://www.scmp.com/economy/china-economy/article/3128157/chinas-covid-19-vaccine-drive-ramps-beijing-rolls-out

21. Wu, H. (2021, April 16). China ramps up vaccination drive with incentives such as free eggs, other goods. Retrieved from https://www.ctvnews.ca/health/coronavirus/china-ramps-up-vaccination-drive-with-incentives-such-as-free-eggs-other-goods-1.5389907

22. Validate user. (n.d.). Retrieved October 21, 2021, from https://academic.oup.com/jtm/article/28/4/taab026/6149495?login=true.

23. Lee, S. (2021, July 06). Vaccine diplomacy: Nation branding and China's COVID-19 soft power play. Retrieved October 21, 2021, from https://link.springer.com/article/10.1057/s41254-021-00224-4

24. Wee, S. (2021, August 06). China says it will provide 2 billion vaccine doses to the world. Retrieved October 21, 2021, from https://www.nytimes.com/2021/08/06/world/china-vaccine-donations-2-billion.html

25. Jakes, L., & Stolberg, S. (2021, September 20). Biden to push global plan to battle Covid as national gaps widen. Retrieved October 21, 2021, from https://www.nytimes.com/2021/09/20/us/politics/biden-covid-19-vaccines.html

26. *Additional covid-19 questions*. Penn Medicine. (n.d.). Retrieved September 19, 2021, from https://www.pennmedicine.org/coronavirus/frequently-asked-questions-about-covid-19/additional-covid-questions#tab-1o.

The Arousal-Mood Theory: Current Research and Limitations

By Jingzhi Zhuang

Author Bio:

Jingzhi Zhuang is a senior at the Beijing National Day School. She hopes to study music cognition and music therapy and become an academic researcher in the future. She is the Student Rights Department vice president and co-founder of the STEP INTO Music Division (volunteering). She is also a member of Deep Space Choir, a top five children's choir in China.

Abstract

The idea that certain types of music can enhance performance on cognitive tasks has fascinated psychologists for decades. One study that has gained heightened attention is the study on the Mozart Effect by Rauscher and others, which demonstrated a positive effect of Mozart's music on performance on a spatial reasoning task. Later research that replicated the study, however, achieved mixed findings, with some supporting the Mozart Effect while others not. One of the most widely cited alternative explanations is the arousal-mood theory developed by Thompson and others in 2001. Yet, though improving on the previous theory, there are also inconsistent findings regarding this model. Consequently, the present review aims to investigate why contradictory results exist by analyzing experiments conducted in the past decade. Studies are discussed in terms of their level of support for the arousal-mood theory and why this might be so, evoking various pre-existing models. It was found that the results largely rely on task type and individual differences. Research gaps and future research directions are also proposed.

Keywords: Arousal-mood theory, cognitive functioning, music, arousal, mood, valence, Eysenck's theory of personality, Mozart Effect.

Introduction

The role of music in cognitive functioning has long been a topic of interest among psychologists. One of the most cited experiments in this field is Rauscher et al's 1993 study on the Mozart Effect (Rauscher et al., 1993). Rauscher was interested in whether Mozart's music could lead to an enhancement in spatial reasoning, meaning the ability to envision and manipulate three-dimensional objects in one's mind, as measured by the spatial-reasoning sub-task of the Stanford-Binet IQ Test. Participants in the experimental group were asked to first listen to Mozart's Sonata for Two Pianos in D Major, K.448 for 10 minutes and then complete a test assessing spatial reasoning skills. Participants from another two groups either received verbal relaxation instructions or sat in silence before completing the same task. Rauscher discovered that the group that listened to Mozart had a temporary (approximately 15-minute) improvement in spatial intelligence compared with the other groups. The study was later published in Nature, and ever since it has intrigued a large number of scientists and even policymakers, one famous example being Georgia's governor Zell Miller, who decided to offer free classical music CDs to every child born in the state in 1998 to make the babies smarter.

As researchers attempted to replicate the results, though, inconsistencies arose. While some studies supported the original findings (Wilson & Brown, 2010), others demonstrated either an absence of effects or short-term effects associated with increased arousal and mood (Bridgett & Cuevas, 2000). Furthermore, studies reveal that the "Mozart Effect" might not be related solely to Mozart, but that other types of music—even non-musical stimuli—can achieve comparable results. For example, in a study conducted by Schellenberg and Hallam (2005), British schoolchildren were assigned to listen to either music by Mozart or Blur, a famous rock band at the time. Results show that kids who listened to Blur performed better than kids who listened to Mozart on spatial tasks. Additionally, in a study by Nantais and Schellenberg, 28 college students either listened to Mozart or heard a short story by Stephen King. Researchers found that participants demonstrated better performance on whichever activity they preferred—in other words, Mozart's music did not necessarily give all of them a boost in visual-spatial ability (Nantais & Schellenberg, 1999)

With the Mozart Effect called into question, researchers started to seek alternative models. One theory proposed is the arousal-mood theory, developed by Thompson et al. in 2001 (Thompson et al., 2001). The arousal-mood theory suggests that it is not Mozart's music that enhances performance on cognitive tasks. Rather, mediators—in this case, arousal and mood—instead of the music itself produce the effect. Arousal refers to the level of psychological and physical activation (which can be interpreted as the intensity of the felt emotion induced by music), and mood refers to the valence of the emotion perceived, which can be either positive or negative (Husain et al., 2002; Schellenberg et al., 2007). Specifically, Thompson hypothesized that high (but not excessively high) arousal and the positive mood of music contribute to a boost in cognitive functioning.

Nevertheless, though effective at explaining a wide range of cases, this model did not put an end to controversies. For example, in a study that investigated music's effects on memory tasks and processing speed in older adults by Bottiroli et al. (2014), although results for the processing speed task did support the arousal-mood theory, free-recall did not necessarily benefit from upbeat music—conversely, performance was also in the condition with the music of more negative mood (Bottiroli et al., 2014). Since the methodologies across different studies vary, and as the arousal-mood theory does not specify the types of cognitive tasks with which it is relevant, the current literature review aims to explore in a more exhaustive approach the cognitive tasks which the hypothesis applies to and gain insight into why results vary across studies. The research question proposed is: under what conditions is the arousal-mood theory applicable? It is hypothesized that the reasons for the discrepancies among experiments include: (1) the differences between the nature of different cognitive tasks (2) the methodological differences across studies and (3) certain studies' failure to control for confounding variables.

Methodology

Regarding population, all studies of healthy adults, defined as all of those aged 18 or above, are included. The exclusion criterion is diagnosis with any type of mental disorder. The type of intervention examined is music-accompanied cognitive tasks. Control conditions include silence groups and opposite mood/arousal groups. Concerning musical material, the present review only incorporated studies that used non-lyrical musical material, as it has already been shown that music with lyrics impairs cognitive performance (Oliver et al., 2021). Another reason why only studies with non-lyrical music are considered is that the original study proposing the arousal-mood theory (Thompson et al., 2001) employed instrumental music.

Concerning the time when music is played during the task, both studies that used music prior-task and those that used music as background were taken into account, as even though the original study used music prior-task, authors of later studies believed that music played before the task and during task has comparable effects on cognition (Bottiroli et al., 2014; Cloutier et al., 2020; Gonzalez & Aiello, 2019; Lemaire et al., 2019; Marti-Marca et al., 2020; Nguyen & Grahn, 2017; Proverbio et al., 2015; Proverbio et al., 2018; Woloszyn & Ewert, 2012), and there is already evidence suggesting such comparability for verbal episodic memory encoding (Cardona et al., 2020). In terms of the publication time range of the included studies, research from the past decade—specifically, research that has a publication date between August 1st, 2011 and August 1st, 2021—is examined. All articles analyzed are journal articles. Regarding search strategy, the researcher searched for relevant literature from three perspectives: arousal-mood theory/hypothesis, arousal's impact on cognitive task performance, and mood's impact on cognitive task performance. The above three dimensions were taken into consideration because while some studies were done to comprehensively explore both aspects of the theory—namely, both arousal and mood—other experiments only focus on either one of the two. To ensure thorough inclusion of pertinent studies, therefore, the researcher conducted her search with three types of search formulae.

Literature from four databases (ProQuest, Web of Science, PsycNet, and PubMed) was thoroughly screened for eligibility, and 19 articles were incorporated. The three types of searches, as illustrated by an example from ProQuest, applied are as follows: (1) ab(music) AND ab(arousal OR intensity OR activation) AND ab(affect cogni*) (2) ab(music) AND ab(mood OR valence) AND ab(affect cogni*) (3) ab(arousal-mood theory) OR ab(arousal and mood hypothesis) OR ab(arousal mood theory) OR ab(arousal-mood hypothesis) OR ab(arousal and mood theory). A table (see Table 1) summarizing key information of the studies is provided on the next page.

Discussion

Arousal and Task Performance

Results regarding arousal's effects on cognitive functioning are relatively consistent across studies. Among the 9 studies that investigated arousal's effects on cognitive task performance, 7 are in line with the arousal-mood theory—that the higher the arousal, the better the performance. This holds for a variety of cognitive tasks, which include tasks on arithmetic calculations, declarative memory, processing speed, visuospatial attentional control, episodic memory, creativity, and (alleviation of) mind-wandering (Bottiroli et al., 2014; Cloutier et al., 2020; Lemaire et al., 2019; Proverbio et al., 2015; Proverbio et al., 2018; Ritter & Ferguson, 2017; Taruffi et al., 2017).

There are differences in ways in which high arousal enhances performance in the above-mentioned tasks depending on the complexity of the tasks. For simple tasks, study results can be explained by the Distraction-Conflict Theory (Baron, 1986). This theory posits that for each task, there is a state of optimal arousal, a state in which an individual is just mentally activated enough to perform a certain task. Both arousal levels below and above the optimum impair performance. When an individual is inadequately aroused, they are prone to boredom and mind-wandering. In this case, if another task that requires a desirable amount of mental effort can be introduced, one can adjust their level of arousal upwards so that the optimum is reached. On the other hand, when an individual is excessively aroused, it would be harmful to introduce a task that necessitates further attentional

resources, and this is when the additional task becomes not a bonus, but a "distraction". The "simple tasks" referred to in this review include all of the previously mentioned cognitive tasks except creativity tasks. A common feature among the simple tasks, for which six studies are relevant, is that they are all conducted in the intelligence domain. For such tasks, performance was enhanced because the music served as a desirable "distraction"—that they occupied the superfluous cognitive capacity and prevented mind-wandering.

Conversely, for creativity tasks (on which one study is relevant), which are more complex, performance enhancement could be attributed to the Dual Pathway to Creativity Model (De Dreu et al., 2008). According to De Dreu, optimally high arousal is the "necessary precondition" to more creative thinking, and valence determines the "pathway" through which creativity for a particular task is achieved.

In terms of the two studies that are inconsistent with the arousal-mood theory, the hidden cause might be flaws in the study design. Regarding the study by Nguyen and Grahn, as suggested by the authors, low-arousal music might have been more beneficial than high-arousal music due to the possibility that the high-arousal music is excessively arousing (Nguyen & Grahn, 2017), thereby jeopardizing participants' performance on recognition memory tasks. In the other study by Oliver and others, music arousal was varied with a 10dB difference in sound intensity (Oliver et al., 2021). As a 10dB amplitude difference is not very significant, and increased arousal is more often induced by faster tempo (Carpentier & Potter, 2007) and added instrumental layers (Gonzalez & Aiello, 2019), it could be argued that the experimental and control conditions are not perceptually different. Thus, it can still be concluded that cognitive performance on simple, intelligence-related tasks and creativity tasks can be improved by moderately arousing music.

Valence/Mood and Task Performance

Regarding the mood dimension of the arousal-mood theory, the findings are murkier. Among the 16 relevant studies, 5 indicate that positive valence improves task performance, 1 show that positive valence is detrimental, 5 demonstrate that positive valence has no effects on performance, and another 4 studies investigating the effects of mood congruence (which refers to whether the valence of the musical stimuli matches the valence of the task stimuli) on task performance reveal that task-valence-congruent music is the most effective. Among studies that are consistent with the arousal-mood theory, no particular patterns regarding the type of tasks employed can be perceived: performance enhanced for processing speed, verbal episodic memory, visual attention, (divergent) creativity, and mind-wandering was alleviated (Bottiroli et al., 2014; Cardona et al., 2020; Marti-Marca et al., 2020; Ritter & Ferguson, 2017; Taruffi et al., 2017). For the study showing the detrimental effect of positive mood, the task assessed selective attention (Putkinen et al., 2017). For studies in which positive valence has null effects, again no patterns can be found: more positive mood does not enhance performance on arithmetic calculations, working memory, declarative memory, recall memory, recognition memory, or associative memory (Borella et al., 2014; Bottiroli et al., 2014; Isarida et al., 2017; Nguyen & Grahn, 2017; Proverbio et al., 2018). Patterns appear once more among research investigating the effects of musical congruence: all four studies are done on memory tasks such as working memory tasks (Franco et al., 2014; Tesoriero & Rickard, 2012; Ward et al., 2021; Woloszyn & Ewert, 2012).

Among studies demonstrating a beneficial effect of positive valence, all studies indicate that the effect applies to all participants except one study, which specifies a particular type of participants who tend to benefit more from positive valence (Cardona et al., 2020). According to Cardona (2020) and others, for highly hedonic participants, meaning individuals who are highly sensitive to musical pleasure, the more pleasant the musical stimuli, the more the words they can encode in their memory during/after music listening. On the contrary, for less hedonic participants, higher pleasantness of the music was associated with worse episodic memory encoding.

Concerning the study (Putkinen et al., 2017) showing impaired cognitive performance in the face of positive valence, such a discovery could be ascribed to the nature of the task employed. One possible explanation, similar to the idea of the Distraction-Conflict Theory, is that performance is affected by both the level of the pleasantness of the music and the requirement of the task. In other words, pleasant

music might be effective at enhancing performance for certain types of tasks but not others depending on the nature of the task in question. As the author mentions, previous studies indicate that positive valence "broadens the scope of visual attention, which can manifest as heightened distractibility". For a task that necessitates highly focused attention (like the one in this study), positive valence would likely be harmful. In this study, participants were given a fairly complex dichotic listening task: different kinds of auditory stimuli were presented to two ears in an alternate pattern; in one ear, target (which include trumpet and piano sounds that the participants were told to concentrate on), non-target, and cue sounds were presented simultaneously; in the other ear, human novel sounds, non-human novel sounds, and standard tones were simultaneously presented. With a whole mixture of sounds alternating between ears from which to extract target sounds, participants would likely need to exert considerable attentional resources, thereby rendering music a distractor.

It might be confusing, then, as to why positive-mood music facilitated visual attention performance in another study (Marti-Marca et al., 2020). A possible explanation might lie in the difference in nature between visual and auditory attention tasks. According to a study by Chun and others, positive valence does not lead to the reallocation of processing resources from foveal stimuli to peripheral stimuli. Instead, to accommodate the processing of both stimuli, additional attentional resources are employed (Chun et al., 2011). In other words, the processing of the peripheral stimuli does not occur at the expense of the processing of the foveal stimuli (Vanlessen et al., 2016). The current study suggests the opposite—processing of the auditory stimuli occurred at the expense of performance on the dichotic listening task. Judging from the above, it also seems that positive valence does not boost cognitive task performance, as it is only non-distractive in the visual attention task and is distractive in the auditory attention task. Such a postulation might also be supported by the fact that the largest number of studies (6 studies) demonstrate no effect of positive valence on task performance. Rather, it could be that a confounding variable is lurking behind the valence variable. Mood and arousal might not be completely separable. According to a study by Proverbio (2015), participants' heart rates (an indicator of arousal) increased for both emotionally touching and pleasant music. Thus, it could be argued that for studies showing a beneficial effect of the positive valence of music, two possible reasons might account for this phenomenon: (1) for the task (especially with tasks not tapping into attention) employed, broadened attention does not matter and might even be beneficial (2) the induced positive mood caused a concurrent enhancement in arousal, which, at a desirable level, boosts cognitive performance.

Finally, the fourth group of studies, which were the four studies on congruence, yield a uniform conclusion: congruent music enhances memory. In other words, happy music facilitates memory for happy stimuli, and the same mechanism applies to stimuli of other moods. It is shown that mood-mismatching music makes recall error-prone. Therefore, it could be argued that positive-mood music has no significant benefits except when it comes to memory tasks in which valence-congruent music is played.

In all, it appears that when it comes to the valence aspect of the arousal-mood theory, applicability largely depends on the type of tasks considered, with tasks concerning memory for positively/negatively-valenced information being one of the few conditions in which the theory holds. Concerning other types of tasks, it might be that the assumption that music with a positive mood enhances cognitive functioning is simply false.

Differences Across Individuals

Although there are certain general ways (that are common to everyone) in which people benefit from music differing in arousal and valence, as are elaborated on above, there still are slight differences in terms of music's effects on cognitive task performance that are dependent on differences in individual characteristics. Among the studies included in the present review, 3 discuss mechanisms through which different kinds of music influence different kinds of people differently. The three mechanisms concern music-specific hedonia (which has already been discussed), introversion/extraversion, and preference for external stimulation (Cardona et al., 2020; Gonzalez & Aiello, 2019; Proverbio et al., 2018;).

Regarding the effects of an individual's introversion/extraversion on cognitive task

performance, it was found that more extroverted participants tend to benefit more from music with high arousal, while those with more introversion often experienced impaired cognitive task performance (Gonzalez & Aiello, 2019; Proverbio et al., 2018;). An explanation concerns Eysenck's Theory of Personality. According to this theory, introverts and extraverts have different amounts of basal cortical arousal levels, with introverts having higher cortical arousal and extraverts having lower cortical arousal. Cortical arousal refers to the "activation of the reticular formation of the brain" (Oxford Reference, 2021), and it is associated with wakefulness and vigilance. The main idea is that there is a point of optimal cortical arousal when performing a particular cognitive task and that both cortical arousal levels that are too low or too high impair performance. Performance as a function of arousal is said to be in an "inverted U" shape (Oxford Reference, 2021). Thus, with already high resting-state cortical arousal, to not jeopardize performance introverts have less capacity for listening to arousing music than extroverts. If high-arousal music is introduced at this time, introverts' arousal will likely exceed the optimum level. This also raises the question of when introverts might be able to best benefit from arousing music (if at all). Possibly, introverts might benefit from arousing music when they are less aroused than usual—for instance when they are drowsy or bored. Additionally, when it comes to general cases, it might be that music with a medium instead of a high level of arousal works best for introverts.

The mechanism concerning the preference for external stimulation builds on the Distraction-Conflict Theory. Specifically, different individuals differ in terms of their "propensity to attend to stimulating elements of their environment" (Gonzalez & Aiello, 2019). Those who have a greater preference for external stimulation are more likely to be distracted by music, especially when it is played in the background.

Research Gaps and Future Directions

One of the most noticeable research gaps is that all studies within discussion were completed on non-musicians. This is possibly due to the consideration that musicians and non-musicians perceive music differently, with musicians preferring different types of music and attending more to artistic minutiae than non-musicians. It would be worthy, therefore, to investigate to what extent the arousal-mood theory applies to musicians and how this particular group of people benefit from music listening cognitively in the future, as the mechanisms would likely be different.

Another possible future research direction is to unravel the common neurological mechanisms among cognitive tasks on which performance benefits from a certain type of arousal or mood (for instance, why "simple tasks" benefit from increased arousal). With such knowledge, it might be possible to then narrow down completely what tasks benefit from what types of music (which the current review is still unable to exhaustively explore) and even predict what kinds of music can facilitate new types of tasks.

Conclusion

In all, this review synthesizes findings among studies conducted in the past decade tapping into the impact of music's arousal and mood on cognitive functioning to explore the causes of the inconsistencies of results of previous studies on the arousal-mood hypothesis. This review provides examples as to how performance on different tasks is enhanced, impaired, or unaffected by music varying in arousal levels and mood. It also examines the effects of individual characteristics on task performance. It was found that: (1) moderately high musical arousal enhances performance on simple intelligence tasks and creativity (divergent thinking) tasks. (2) positive mood does not seem to enhance performance on cognitive tasks except when it comes to memory tasks in which mood congruence is advantageous. (3) different types of individuals benefit from music differently. The reasons for the inconsistencies are as predicted: they are mainly due to the different nature of cognitive tasks, methodological differences, and failure to control for confounding variables. Future research can be conducted to investigate the effects of music with different arousal and mood on cognitive functioning on musicians and to delve into the biological aspects of cognitive tasks on which the performance is enhanced by certain types of music.

References

Baron, R. S. (1986). Distraction-Conflict Theory: Progress and Problems. In L. Berkowitz (Ed.), Advances in Experimental Social Psychology (Vol. 19, pp. 1-40). Academic Press. https://doi.org/https://doi.org/10.1016/S0065-2601(08)60211-7

Borella, E., Carretti, B., Grassi, M., Nucci, M., & Sciore, R. (2014). Are age-related differences between young and older adults in an affective working memory test sensitive to the music effects? Frontiers in Aging Neuroscience, 6, 298-298. https://doi.org/10.3389/fnagi.2014.00298

Bottiroli, S., Rosi, A., Russo, R., Vecchi, T., & Cavallini, E. (2014). The cognitive effects of listening to background music on older adults: Processing speed improves with upbeat music, while memory seems to benefit from both upbeat and downbeat music. Frontiers in Aging Neuroscience, 6, 284-284. https://doi.org/10.3389/fnagi.2014.00284

Bridgett, D. J., & Cuevas, J. (2000). Effects of Listening to Mozart and Bach on the Performance of a Mathematical Test. Perceptual and Motor Skills, 90(3_suppl), 1171-1175. https://doi.org/10.2466/pms.2000.90.3c.1171

Cardona, G., Rodriguez-Fornells, A., Nye, H., Rifà-Ros, X., & Ferreri, L. (2020). The impact of musical pleasure and musical hedonia on verbal episodic memory. Scientific Reports, 10(1), 16113. https://doi.org/10.1038/s41598-020-72772-3

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. Annu Rev Psychol, 62, 73-101. https://doi.org/10.1146/annurev.psych.093008.100427

Cloutier, A., Fernandez, N. B., Houde-Archambault, C., & Gosselin, N. (2020). Effect of Background Music on Attentional Control in Older and Young Adults. Frontiers in Psychology, 11, 557225-557225. https://doi.org/10.3389/fpsyg.2020.557225

Cortical arousal. Oxford Reference. Retrieved 21 Sep. 2021, from https://www.oxfordreference.com/view/10.1093/oi/authority.20110803095640881.

Cournoyer Lemaire, E. (2019). The effect of background music on episodic memory. [doi:10.1037/pmu0000234]. Educational Publishing Foundation.

De Dreu, C. K., Baas, M., & Nijstad, B. A. (2008). Hedonic tone and activation level in the mood-creativity link: Toward a dual pathway to creativity model. J Pers Soc Psychol, 94(5), 739-756. https://doi.org/10.1037/0022-3514.94.5.739

Dillman Carpentier, F. R., & Potter, R. F. (2007). Effects of Music on Physiological Arousal: Explorations into Tempo and Genre. Media Psychology, 10(3), 339-363. https://doi.org/10.1080/15213260701533045

Franco, F., Swaine, J. S., Israni, S., Zaborowska, K. A., Kaloko, F., Kesavarajan, I., & Majek, J. A. (2014). Affect-matching music improves cognitive performance in adults and young children for both positive and negative emotions. [doi:10.1177/0305735614548500]. Psychology of Music. Sage Publications.

Gonzalez, M. F., & Aiello, J. R. (2019). More than meets the ear: Investigating how music affects cognitive task performance. [doi:10.1037/xap0000202]. American Psychological Association.

He, W.-J., Wong, W.-C., & Hui, A. N.-N. (2017). Emotional Reactions Mediate the Effect of Music Listening on Creative Thinking: Perspective of the Arousal-and-Mood Hypothesis [Original Research]. Frontiers in Psychology, 8(1680). https://doi.org/10.3389/fpsyg.2017.01680

Husain, G., Thompson, W. F., & Schellenberg, E. G. (2002). Effects of Musical Tempo and Mode on Arousal, Mood, and Spatial Abilities. Music Perception: An Interdisciplinary Journal, 20(2), 151-171. https://doi.org/10.1525/mp.2002.20.2.151

Inverted U hypothesis. Oxford Reference. Retrieved 21 Sep. 2021, from https://www.oxfordreference.com/view/10.1093/oi/authority.20110803100009722.

Isarida, T. K., Kubota, T., Nakajima, S., & Isarida, T. (2017). Reexamination of mood-mediation hypothesis of background-music-dependent effects in free recall. Q J Exp Psychol (Hove), 70(3), 533-543. https://doi.org/10.1080/17470218.2016.1138975

Mado Proverbio, A., Lozano Nasi, V., Alessandra Arcari, L., De Benedetto, F., Guardamagna, M., Gazzola, M., & Zani, A. (2015). The effect of background music on episodic memory and autonomic responses: Listening to emotionally touching music enhances facial memory capacity. Scientific Reports, 5(1), 15219. https://doi.org/10.1038/srep15219

Marti-Marca, A., Nguyen, T., & Grahn, J. A. (2020). Keep Calm and Pump Up the Jams: How Musical Mood and Arousal Affect Visual Attention. Music & Science, 3, 2059204320922737. https://doi.org/10.1177/2059204320922737

Nantais, K. M., & Schellenberg, E. G. (1999). The Mozart effect: An artifact of preference. [doi:10.1111/1467-9280.00170]. Psychological Science, Blackwell Publishing.

Nguyen, T., & Grahn, J. A. (2017). Mind your music: The effects of music-induced mood and arousal across different memory tasks. [doi:10.1037/pmu0000178]. Educational Publishing Foundation.

Oliver, M. D., Levy, J. J., & Baldwin, D. R. (2020). Examining the effects of musical type and intensity in performing the flanker task: A test of attentional control theory applied to non-emotional distractions. Psychology of Music, 49(4), 1017-1026. https://doi.org/10.1177/0305735620922599

Proverbio, A. M., De Benedetto, F., Ferrari, M. V., & Ferrarini, G. (2018). When listening to rain sounds boosts arithmetic ability. PLoS One, 13(2), e0192296. https://doi.org/10.1371/journal.pone.0192296

Putkinen, V., Makkonen, T., & Eerola, T. (2017). Music-induced positive mood broadens the scope of auditory attention. Social Cognitive and Affective Neuroscience, 12(7), 1159-1168. https://doi.org/10.1093/scan/nsx038

Rauscher, F. H., Shaw, G. L., & Ky, K. N. (1993). Music and spatial task performance. Nature, 365(6447), 611. https://doi.org/10.1038/365611a0

Ritter, S. M., & Ferguson, S. (2017). Happy creativity: Listening to happy music facilitates divergent thinking. PLoS One, 12(9), e0182210. https://doi.org/10.1371/journal.pone.0182210

Schellenberg, E. G., & Hallam, S. (2005). Music listening and cognitive abilities in 10- and 11-year-olds: The Blur effect. Ann N Y Acad Sci, 1060, 202-209. https://doi.org/10.1196/annals.1360.013

Schellenberg, E. G., Nakata, T., Hunter, P. G., & Tamoto, S. (2007). Exposure to music and cognitive performance: Tests of children and adults. Psychology of Music, 35(1), 5-19. https://doi.org/10.1177/0305735607068885

Taruffi, L., Pehrs, C., Skouras, S., & Koelsch, S. (2017). Effects of Sad and Happy Music on Mind-Wandering and the Default Mode Network. Scientific Reports 7, 14396). https://doi.org/10.1038/s41598-017-14849-0

Tesoriero, M., & Rickard, N. S. (2012). Music-enhanced recall: An effect of mood congruence, emotion arousal or emotion function? Musicae Scientiae, 16(3), 340-356. https://doi.org/10.1177/1029864912459046

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, Mood, and The Mozart Effect. Psychological Science, 12(3), 248-251. https://doi.org/10.1111/1467-9280.00345

Vanlessen, N., De Raedt, R., Koster, E. H. W., & Pourtois, G. (2016). Happy heart, smiling eyes: A systematic review of positive mood effects on broadening of visuospatial attention. Neurosci Biobehav Rev, 68, 816-837. https://doi.org/10.1016/j.neubiorev.2016.07.001

Ward, E. V., Isac, A., Donnelly, M., Van Puyvelde, M., & Franco, F. (2021). Memory improvement in aging as a function of exposure to mood-matching music. Acta Psychol (Amst), 212, 103206. https://doi.org/10.1016/j.actpsy.2020.103206

Wilson, T. L., & Brown, T. L. (1997). Reexamination of the Effect of Mozart's Music on Spatial-Task Performance. The Journal of Psychology, 131(4), 365-370. https://doi.org/10.1080/00223989709603522

Woloszyn, M. R., & Ewert, L. (2012). Memory for facial expression is influenced by the background music playing during study. Adv Cogn Psychol, 8(3), 226-233. https://doi.org/10.2478/v10053-008-0118-9

The Economics and Political Constraints of a Green Jobs Guarantee

By Valerie Xu

Author Bio:

Valerie Xu is a high school junior at Greenhill School in Addison, TX. She is currently an editor on her school newspaper as well as a long-term member of the debate team. She was selected among 11 Americans to participate in the Benjamin Franklin Transatlantic Fellowship, which was sponsored by the Department of State and Purdue University. In the future, she hopes to study interdisciplinary topics related to government, public policy, and economics.

Abstract

Fossil fuel consumption has exacerbated the effects of climate change, and despite these imminent dangers to the environment, the US economy still remains immensely dependent on fossil fuels and other pollutants as a main source for energy because of the high production cost and lack of labor among the renewable energy sector. To solve this economic problem, the US Federal Government has drafted the Green New Deal, which includes a Federal Jobs Guarantee (FJG) mostly directed towards the renewable energy industry. This job guarantee would seek to give temporary employment support while also generating labor towards renewable energy production.

In this paper, I will first outline the economic obstacles that currently prevent renewable energy from being America's main energy source before introducing what the design and implementation of a FJG program would look like. I will then address the political constraints, rooted in campaign contributions and lobbying from fossil fuel corporations, which prevents an FJG from being politically feasible. Lastly, to address the political constraints, I will propose an updated version of the current FJG program that might be more politically viable.

Keywords: Federal Jobs Guarantee; Renewable Energy; Green New Deal; Market Failure; Employment; Optimal Policy; Climate Change; Fossil Fuel Corporations.

The Economics and Political Constraints of a Green Jobs Guarantee

From raging wildfires to flooded cities, the impact of global warming poses a threat to national and global security. Its implications have the capacity to destroy national infrastructures, cause political unrest, and devastate global economies, a reality so daunting that the Red Cross has proclaimed global warming a bigger threat than the COVID-19 pandemic (Goering & Rodriguez, 2020). Already, the cost and financial burden of climate change and extreme weather patterns have been felt by the US government, which last year spent a record $210 billion on extreme weather and natural disaster related expenditures (Newburger, 2021). The human cost of a planet that is 1.5-2□ warmer could be even more disastrous, and according to a recent IPCC model is estimated to surpass one hundred million, twice the death toll of World War II and three times the death toll of the Great Leap Forward (Wallace-Wells, 2020).

Carbon dioxide emissions, largely from the burning of fossil fuels, make up the majority of greenhouse gases and are the main perpetrators behind climate change. As a result, without a sustained multilateral promise from countries, reaching zero emissions and reversing the effects of climate change cannot be achieved. France and Germany have already proposed policies like the carbon tax in the hopes of switching to a more sustainable economy. Total economic decarbonization, however, can only happen if countries drastically shift their energy consumption from fossil fuels to renewable energy sectors, which is especially challenging in the US where there is a high economic and consumer dependence on fossil fuels.

In a recent study, U.S. Energy Information Administration (EIA 2020) reported that 80% of all US domestic energy consumption comes from fossil fuel. The current economic problem is the high cost and low production quantity of renewable energy, especially in wind and solar energy sectors. The US Federal Government proposes tackling the issue of cost and renewable energy through the Green New Deal (GND), which contains a Federal Jobs Guarantee (FJG) program. Considering the high reliance on fossil fuels, current proponents of a Green Jobs Guarantee and GND have said that implementation should start as soon as possible and last about a decade. This ambitious time frame is instrumental to ensure that the US reaches President Joe Biden's goal of net zero emissions by 2050. A recent special report from the International Energy Agency (IEA 2021) found that quick action needs to be taken to limit emissions and prevent warming from reaching 2 degrees Celsius to prevent irreversible effects of climate change. The question I will answer in this paper is has the FJG under the Green New Deal been effective in helping to decrease renewable energy prices and increase production of renewable energy to a level that allows for a shift to a decarbonized economy? I argue that a Federal Jobs Guarantee benefits both unemployment and the environment. The issue, however, is the political feasibility of the current FJG and GND as a result of strong opposition from conservative politicians, many of whom have been backed by influential fossil fuel corporations and lobbyists. One possible alternative is to eliminate all other tenants of the GND and only implement a green FJG program, which would cut down on costs and other parts of the GND that GOP politicians might view as too radical.

Economic Problem

Many progressive climate activists agree that greenhouse gases constitute a negative externality to our society and economy, and although renewable energy is the best alternative to fossil fuels, big barriers to reaching a renewable energy economy remain. One barrier is production of renewable energy, which, despite recent progress, remains lower than that of fossil fuels. In the US, renewable energy constitutes 20% of all energy generated (IEA 2021). The International Renewable Energy Agency (IRENA 2018) suggests that in order for a shift away from emissions, 60% of energy consumption should come from renewable energy, which is significantly higher than current levels. The main reason for the low production level is because of high costs, which means that a transition to a renewable energy economy would require unpopular large national investments. In a study by the Institute of Energy Research (IER 2019), the cost of creating a 100% renewable economy by 2030 would be approximately $4.5 trillion. Although this estimation might seem daunting, a recent Rewired America report calculated

that a green economy could save America about $321 bn in energy costs (Milman 2020). Nevertheless, many hesitate to invest in renewable energy as they fear the uncertain proposition of spending vast amounts of money that could not produce profit for more than ten years. Therefore, it is important to continue to analyze reasons behind the high costs of renewable energy and explore ways to cut down to increase production efficiency.

One reason why renewable energy is expensive is because of its unpredictable supply. Unlike with fossil fuel, day-to-day weather can determine the production quantity of renewable energy production lines. Especially in wind and solar sectors, production quantity is dependent on natural constraints as it is impossible to produce wind and solar energy without wind or sunshine. According to experts in geology Souvik Sen and Sourav Ganguly in the *Renewable and Sustainable Energy Reviews*, "Unlike fossil fuels RE is site specific. As for example, wind turbine output depends on the wind speed and other related properties and a wind-based RE system makes sense only in those particularly favorable regions" (Sen and Ganguly, 2017, pp. 6).

With this level of natural dependency, storage systems, which vary based on the type of renewable energy, act as a safety net so that renewable energy plants can still produce enough energy and meet market demands regardless of weather patterns. Building storage systems, however, are not only important in wind and solar sectors. Hydropower, which constitutes the majority of the renewable sector, uses mass storage systems called Pumped-Storage Hydropower (PSH) to store the water necessary to generate electricity; therefore, in times of high demand, water is released from the storage system to generate electricity, ensuring that energy can be produced despite natural barriers (National Hydropower Association 2019). But renewable energy storage systems are comparably scarcer than that of fossil fuels, which poses an obstacle to increasing production of renewable energy. This scarcity is largely due to the high costs of renewable energy storage systems, which deters companies from investing in storage and renewable energy altogether. In fact, a recent Vox article found that renewable energy storage would have to be an average of $20 per kilowatt hour, which is a 90% decrease from today's cost, to sustain a total renewable economy (Roberts 2019). Therefore, mass investment and intervention are necessary to lower prices and build more storage systems across the country to sustain a level of renewable energy production that can transition the US economy away from fossil fuels.

Another problem that prevents renewable energy production is a lack of labor. The labor problem among clean energy sectors stems from how new the sector inherently is. Workers are vital to ramp up production, and the development of innovation and technology that could possibly increase the efficiency of the energy produced, cuts down on the cost. When laying out barriers to a renewable energy economy, Sen and Ganguly wrote, "Requirement of skilled human resources with specific training in RE is another issue here. Skills to operate and maintain RE hardware [are] very important to make a RE project running successfully," (Sen and Ganguly, 2017, pp. 6). The lack of adequate labor prevents technological innovation essential to helping to create valuable growth in clean energy. The FJG gives such labor support while also giving workers the opportunity to develop skills and training during their temporary employment. In her paper, 'The Job Guarantee: Design, Jobs, and Implementation,' Pavlina R. Tcherneva, Associate Professor of Economics at Bard College and a Research Scholar at the Levy Economics Institute, said that under a FJG training and education services would be offered "at the macro level" and serve to temporarily shift people out of unemployment (Tcherneva, 2018, pp. 15). And especially in newer renewable energy sectors, where skills and technology are not as highly developed, governmental intervention like the Production Tax Credit (PTC) and Investment Tax Credit (ITC) are crucial to incentivizing investment in renewable technology, research, and labor. Without governmental assistance, corporations and workers have no incentive to invest in something as unpredictable and new as green technology. Therefore, a policy proposal that boosts jobs in the renewable energy sector is essential to lowering the cost and increasing the production of renewable energy.

To address the issues of climate change, Congresswoman Alexandria Ocasio-Cortez of New York and Massachusetts Senator Ed Markey have proposed the Green New Deal (GND), which, once implemented, would work to help the US reach a zero-emission economy within a decade. Some of

the key pieces of the GND include "upgrading" national infrastructure in an energy efficient manner, working with farmers to limit agriculturally generated pollution, investing in reducing emissions among the transportation system (expanding electric car manufacturing and high speed rails), and guaranteeing a federal job that provides a "living wage" for every American working in renewable energy (Kurtzleben 2019). This jobs plan, also known as a Federal Jobs Guarantee program (FJG), seeks to address two key problems: unemployment and climate change. In her paper, Tcherneva frames what a jobs guarantee program would look like:

> The Job Guarantee (JG) is a public option for jobs. It is a permanent, federally funded, and locally administered program that supplies voluntary employment opportunities on demand for all who are ready and willing to work at a living wage. While it is first and foremost a jobs program, it has the potential to be transformative by advancing the public purpose and improving working conditions, people's everyday lives, and the economy as a whole (Tcherneva, 2018, pp. 2).

Through a unique incorporation of an FJG concentrated in the renewable energy sector, production capacity of renewable energy would increase, and prices of renewable energy would theoretically drop, allowing for an expansion of renewable energy development and competition (Tcherneva, 2018, pp. 4-7). The financial incentive of a guaranteed job could also garner the attention of high-skilled workers like engineers from the private sector or alternative energy sectors that could be crucial to advancing research and development of production methods. Any developments or advancements in renewable energy garnered from a federal investment in green infrastructure will not only help the public sector but will also spill over to benefit the productivity of private renewable energy companies (Fogarty 2015). Additionally, the FJG also acts as a safety net for individuals most susceptible to unemployment during the transition to a green economy; this includes fossil fuel workers and engineers whose skills coincide largely with those of renewable energy workers (Tcherneva 2020). Moreover, for individuals with no prior knowledge or skills, this program acts as a job training and transition system that allows many to gain adequate skills to adapt and find employment in a green economy.

In recent years, climate change has gotten exponentially worse and every year since 2015 has been the warmest on record (World Meteorological Organization 2021). So, if comprehensive actions are not taken, the effects of climate change will continue to get exponentially worse, and the financial burden on taxpayers will only increase. Therefore, to minimize costs and the worsening effects of global warming, a GND must be passed as soon as possible. Although many progressives have applauded this aggressive approach and timeline, the GND and FJG has come under fire by many conservatives and moderates like Senators Joe Manchin from West Virginia and Kyrsten Sinema from Arizona, who have publicly declared their opposition to the current proposed version of the GND and FJG out of fear of inflation and private sector crowd out. Especially in the wake of a record $1.9 trillion spending bill from the American Rescue Plan and the compromised $1 trillion for Biden's hard-fought infrastructure bill, many fiscal conservatives hesitate to spend the estimated $543 billion to maintain a FJG (Darity and Hamilton 2017). Despite these costs, proponents argue that the benefits the bill offers to our economy compensates the cost of the program and adds an additional $2 trillion to our economy within the next ten years (Friedman 2019). Therefore, the question for climate activists is whether the mass federal investment necessary to maintain a FJG program under a Green New Deal is effective enough to garner a decrease in renewable energy prices and create a shift to a decarbonized economy.

Political Constraints and Optimal Policy Recommendations

Although many conservative politicians have explained their opposition to the GND as a rejection of big government, it is also important to analyze the role that fossil fuel corporations and lobbyists have on influencing climate policy and lawmakers. According to a 2020 article from The Guardian, oil and gas companies alone spent over $84 billion for congressional campaigns (overwhelmingly Republican) during the 2018 election cycle (Holden 2020). Most notably, Senator Jim Inhofe, from Oklahoma, who tried to bring a snowball onto the Congressional floor to disprove the effects of climate change, has received over $1.5 million from the fossil fuel industry alone during his career (Geary 2019). As seen with Senator Inhofe, these campaign contributions have real-world impacts in

shifting congressional decision making and changing perceptions on climate change. According to a study from the Proceedings of the National Academy of Sciences (PNAS), lawmakers were estimated to receive an additional $1,700 in campaign contributions from fossil fuel corporations every time, during an election cycle, their League of Conservation Voters (a non-profit environmental activist group) score decreased by 10%; the study concludes that, "The more a given member of Congress votes against environmental policies, the more contributions they receive from oil and gas companies supporting their reelection" (Goldberg et al. 2020). And because green infrastructure and GND directly oppose the interests of fossil fuel corporations, it is very likely that fossil fuel corporations will funnel big money into candidates and lobbyists to promote anti-environmentalism and kill the bill.

Fossil fuel companies also use campaign contributions to leverage politicians to give them large sums of aid in the form of tax exemptions and subsidies, allowing fossil fuel corporations to get away with almost no costs when polluting the environment. In fact, the fossil fuel sector receives a total of $20 billion of subsidies from the federal government each year, which is far more than the subsidies provided in any other energy industry in the US (Urpelainen and George 2021). Fossil fuel subsidies not only suck up the federal budget but also incentivize the production of emissions, making it harder to transition towards alternative energy methods. With these subsidies, fossil fuel corporations generate fuel and pollute the Earth with greenhouse gases with taxpayer dollars at an extremely low cost (Irfan 2019). But even though there is a direct correlation between fossil fuel subsidization and the worsening of global warming, the need for campaign contributions from fossil fuel companies deters many politicians from actually doing anything to end these subsidies.

But fossil fuel lobbying and campaign contributions do not only impact Republican politicians. Although 85% of the lawmakers who receive campaign contributions from fossil fuel industries are Republicans, in the wake of the 2020 election, where the Democratic nominee Joe Biden announced his position against a ban on fracking, fossil fuel corporations like ExxonMobil started dramatically increasing their support and contributions to Democratic candidates and campaigns. During the 2020 presidential election, ExxonMobil increased its political contribution towards Democrats to 41%, a significant increase compared to 32.6% during the 2016 presidential election (Hampton 2020). This shift shows that the influence of fossil fuel corporations is prevalent, not only among Republican candidates but also among many moderate Democrats from coal mining or rural states like West Virginia. It is not only the GOP but also the Democratic Party that have prevented the ability for a comprehensive climate policy like the GND from passing. The "quid pro quo" between fossil fuel corporations and politicians causes gridlock and our inability to combat climate change.

When responding to discussions about Biden's climate policy, Senator John Barrasso of Wyoming commented, "Proposals that impose a cost on carbon will hurt American families," and Senator Tim Scott from South Carolina said bluntly that "Our best future won't come from Washington schemes or socialist dreams" (Gross 2021). Scott's argument has some merit. Thermal storage and hydrogen storage, one of the most commonly used types of storage for renewable energy, is far more expensive than fossil fuel storage, yet many Americans seem to be in favor of climate policies that seek to lower prices of renewable energy. In fact, a staggering 83% of Americans favor tax breaks for utilities that develop renewable energy, and 62% of Americans favor taxing companies for emitting greenhouse gasses, showing a disconnect between the American people and Washington (Gross 2021). Despite these poll numbers, politicians tend to vote in ways that guarantee reelection, which means avoiding tax increases and policies that require big budgeting.

So, although an FJG would address many of the issues like labor shortages and high production cost plaguing the renewable energy sector, it is unlikely to pass the U.S. Senate with many moderate Democrats like Senator Manchin already voicing their objection, especially considering the possibility that Republicans will take control of the Senate after the 2022 midterm election. Therefore, a plan that increases development in renewable energy technology and labor while minimizing cost to a level that would generate bipartisan support is key to allowing for a practical solution to the current issues of renewable energy. An optimal solution to the current version of the GND is to cut the rest of the GND and purely implement the FJG portion of the bill, which will cause much

less controversy as it excludes most of the social aspects of the plan like healthcare expansion that many Republican politicians have deemed "socialist." Through this budget cut, cost will also be minimized to $500 billion per year, which is far cheaper than the multi trillion-dollar bill originally proposed. Although $500 billion seems like a lot at first glance, it is not abnormal for federal budgeting bills to exceed multi billion or trillion dollars in funding (as seen with the American Rescue Plan and infrastructure bill); although there still might be some hesitancy from ultra conservative politicians to invest, a $500 billion bill under America's current political limit is not totally out of the norm.

These proposed budget cuts would create a more targeted solution to the labor shortages in the renewable energy market. In fact, an FJG was designed with the specific intent to increase labor among the renewable energy sector, so it is not necessary to add additional programs to solve a problem that only requires a simple solution that an FJG can solve. Moreover, an FJG offers economic benefits to local economies (it is estimated that a FJG would employ over 10.7 million people) and would raise the quality of living for millions of Americans. Therefore, fewer people would qualify or need assistance from current governmental welfare policies, and budgeting for these programs can be cut substantially to fund the FJG. These cuts should not cause much outrage from Republican politicians as many of these social programs are not programs that Republicans have special interest in. Therefore, an updated GND, which only contains a green jobs guarantee, would largely minimize the burden on taxpayers, make Republican opposition on the grounds of cost less warranted, and increase the likelihood of some level of bipartisan agreement necessary to overcome the political gridlock that has prevented climate bills from being implemented in the past.

To name a few positive externalities, a politically feasible FJG would offer better public health because with an increased income, people would be more capable of investing in better healthcare. With the optimal version of the policy, there would also be no new CO2 emission within 10 years of enactment. As a result, air pollution related diseases like heart attack and stroke would decrease significantly. Therefore, this bill would create a healthier and larger working population that would contribute to the economy and would prevent a large number of premature deaths caused by the factors related to environmental pollution.

Another positive externality of the FJG is that it would create new jobs for the working class and also provide easier ways of transportation. For example, building greener infrastructure such as zero emissions transportation systems around the country would allow for easier movement and accessibility. An increased accessibility through an upgraded infrastructure system allows for more employment opportunities, which helps to break down socioeconomic inequalities facing marginalized groups as a result of redlining or other forms of exclusion.

Nevertheless, there are still many concerns surrounding an FJG. One of the greatest concerns is that an FJG would crowd out the private renewable energy industry. Proponents of the FJG like Tcherneva have dismissed those claims, saying that any potential harm to the economy from private sector crowd out would be offset by the increased employment that an FJG would bring, and the development of new technologies and production methods that would eventually spill over to benefit the private renewable energy sector. For example, workers who set up Electric Vehicle (EV) charging stations across the country under a FJG program benefit private companies like Tesla and GM that seek to expand their outreach of EV. It is also important to recognize that an FJG is a temporary 10-year plan, meaning that the purpose of the plan is simply to phase out America's current fossil fuel dependent economy. It is not meant to act as a permanent federal replacement of the current energy system or to replace the private renewable energy sector, but rather it seeks to support the private market by building a framework for sustainable infrastructure and energy production systems that would eventually be regulated by the private sector at the end of the proposed 10 years.

Furthermore, another concern highlighted by sceptics of the bill is that the probability of the GND happening is low insofar as fossil fuel corporations continue to funnel large amounts of money to manipulate politicians and ordinary people. So, when discussing climate policy, it is important to incorporate legislative action that will address the enormous power fossil fuel industries have in preventing necessary and bipartisan efforts to tackle climate change. By

getting rid of all other tenets of the GND and solely implementing an FJG, the problem of political feasibility becomes less pronounced. Although this version of the GND will continue to stimulate renewable energy sector growth, it does not directly crash or impact the fossil fuel sectors. Moreover, this version of the bill gives corporations like ExxonMobil sufficient time to phase out and develop new technologies necessary to adapt to a green economy, and it is more likely that fossil fuel industries will support this bill rather than more ambitious forms of climate legislation that directly disincentivize the purchasing and selling of fossil fuel like the original GND or a carbon tax.

Through hurricanes and raging wildfires, Mother Earth has warned humanity that we cannot afford the cost of negligence in the face of such an existential threat, and especially in an age where our economy depends predominantly on fossil fuels, it is important for drastic action to be initiated to reverse the course of an impending global disaster. Therefore, an FJG would be an optimal solution to phasing away from a fossil fuel-based economy while ensuring economic stability and steady employment rates during the transition process.

References

Climate change indicators and impacts worsened in 2020. World Meteorological Organization. (2021, April 20). Retrieved September 14, 2021, from https://public.wmo.int/en/media/press-release/climate-change-indicators-and-impacts-worsened-2020.

Cost of transitioning to 100-percent renewable energy. IER. (2019, July 15). Retrieved September 13, 2021, from https://www.instituteforenergyresearch.org/renewable/cost-of-transitioning-to-100-percent-renewable-energy/.

Darity, W., & Hamilton, D. (2017). Full employment and the job Guarantee: An All-american Idea. *Full Employment and Social Justice*, 195–204. https://doi.org/10.1007/978-3-319-66376-0_9

Fogarty, W. (n.d.). *Federally funded, locally organized: An affordable job guarantee program*. Global Institute for Sustainable Prosperity. Retrieved September 14, 2021, from http://www.global-isp.org/working-paper-no-108/.

Frequently asked Questions (faqs) - U.S. Energy Information Administration (EIA). Frequently Asked Questions (FAQs) - U.S. Energy Information Administration (EIA). (n.d.). Retrieved September 13, 2021, from https://www.eia.gov/tools/faqs/faq.php?id=427&t=3.

Friedman, L. (2019, February 21). *What is the green New DEAL? A Climate Proposal, Explained*. The New York Times. Retrieved September 14, 2021, from https://www.nytimes.com/2019/02/21/climate/green-new-deal-questions-answers.html.

Geary, J. (n.d.). *The dark money of climate change*. DigitalCommons@COD. Retrieved September 14, 2021, from https://dc.cod.edu/essai/vol17/iss1/17.

Global energy transformation: A roadmap to 2050 (2019 edition). International Renewable Energy Agency. (IRENA). (n.d.). Retrieved September 13, 2021, from https://www.irena.org/publications/2019/Apr/Global-energy-transformation-A-roadmap-to-2050-2019Edition.

Goering, L., & Rodriguez, S. (2020, November 17). *Global warming bigger threat than coronavirus: Red Cross*. Climate Crisis News | Al Jazeera. Retrieved November 5, 2021, from https://www.aljazeera.com/news/2020/11/17/climate-change-bigger-threat-than-covid-red-cross.

Goldberg, M. H., Marlon, J. R., Wang, X., van der Linden, S., & Leiserowitz, A. (2020). Oil and gas companies invest in legislators that vote against the environment. *Proceedings of the National Academy of Sciences, 117*(10), 5111–5112. https://doi.org/10.1073/pnas.1922175117

Gross, S. (2021, June 16). *Republicans in Congress are out of step with the American public on climate*. Brookings. Retrieved September 14, 2021, from https://www.brookings.edu/blog/planetpolicy/2021/05/10/republicans-in-congress-are-out-of-step-with-the-american-public-on-climate/.

Hampton, L. (2020, October 16). *U.S. oil MAJORS pitch more campaign cash to Democrats As frack BATTLE looms*. Reuters. Retrieved September 14,

2021, from *https://www.reuters.com/article/us-usa-election-oil-donors/u-s-oil-majors-pitch-more-campaign-cash-to-democrats-as-frack-battle-looms-idUSKBN27116P.*

Holden, E. (2020, February 24). *Oil and gas industry rewards US lawmakers who OPPOSE environmental Protections – study*. The Guardian. Retrieved September 14, 2021, from https://www.theguardian.com/environment/2020/feb/24/oil-gas-industry-us-lawmakers-campaign-donations-analysis.

IEA. (2021, May 1). *Pathway to critical and FORMIDABLE goal of net-zero emissions by 2050 is narrow but brings huge BENEFITS*, according to Iea special report - news. IEA. Retrieved September 13, 2021, from https://www.iea.org/news/pathway-to-critical-and-formidable-goal-of-net-zero-emissions-by-2050-is-narrow-but-brings-huge-benefits.

Irfan, U. (2019, May 17). *Fossil fuels are underpriced by a whopping $5.2 TRILLION*. Vox. Retrieved September 14, 2021, from https://www.vox.com/2019/5/17/18624740/fossil-fuel-subsidies-climate-imf.

Kurtzleben, D. (2019, February 7). *Rep. Alexandria OCASIO-CORTEZ Releases green New Deal Outline*. NPR. Retrieved September 14, 2021, from https://www.npr.org/2019/02/07/691997301/rep-alexandria-ocasio-cortez-releases-green-new-deal-outline.

Milman, O. (2020, October 22). *Aggressive push to 100% renewable energy could save Americans billions – study*. The Guardian. Retrieved September 13, 2021, from https://www.theguardian.com/environment/2020/oct/22/us-renewable-energy-costs-savings-study-report.

Newburger, E. (2021, January 7). *Disasters caused $210 billion in damage in 2020, showing growing cost of climate change*. CNBC. Retrieved November 5, 2021, from https://www.cnbc.com/2021/01/07/climate-change-disasters-cause-210-billion-in-damage-in-2020.html.

Pumped storage. National Hydropower Association. (2019, September 12). Retrieved September 14, 2021, from https://www.hydro.org/policy/technology/pumped-storage.

Roberts, D. (2019, August 9). *Getting to 100% renewables requires cheap energy Storage. but How cheap?* Vox. Retrieved September 14, 2021, from https://www.vox.com/energy-and-environment/2019/8/9/20767886/renewable-energy-storage-cost-electricity.

Sen, S., & Ganguly, S. (2017). Opportunities, barriers and issues with renewable energy development – a discussion. *Renewable and Sustainable Energy Reviews, 69*, 1170–1181. https://doi.org/10.1016/j.rser.2016.09.137

Tcherneva, P. R. (2018). The Job Guarantee: Design, jobs, and implementation. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3155289

Tcherneva, P. R. (2020). *The Case for a Job Guarantee*. Polity Press.

U.S. Energy Information Administration - EIA - independent statistics and analysis. Fossil fuels account for the largest share of U.S. energy production and consumption - Today in Energy - U.S. Energy Information Administration (EIA). (n.d.). Retrieved September 13, 2021, from https://www.eia.gov/todayinenergy/detail.php?id=45096.

Urpelainen, J., & George, E. (2021, July 21). *Reforming global fossil FUEL subsidies: How the United states can restart international cooperation*. Brookings. Retrieved September 14, 2021, from https://www.brookings.edu/research/reforming-global-fossil-fuel-subsidies-how-the-united-states-can-restart-international-cooperation/.

Wallace-Wells, D. (2020). Cascades. In *The Uninhabitable Earth: Life After Warming*. essay, Tim Duggan Books.

Pharmaceutical Market Failure, Subsequent Policy, and its Obstacles in China

By Walter Liu

Author Bio:

Yaning (Walter) Liu is a 10th grade student at the Episcopal High School in Alexandria, Virginia. He has a great passion for history and social studies, being the founder and president of his school's history club. He is also a correspondent in the Spectrum Club of his school, where they discuss social issues such as race and gender. Additionally, he is part of the government club, which enriched his knowledge of the U.S government while staying well-informed, civic, open-minded, and empathetic. His research paper on "The Battle of Macau in 1622" is currently under review at the *Concord Review*. He was cited in a *Washington Post* article about the Critical Race Theory Controversy, in which he contributed a personal perspective informed by his family history.

Abstract

There is a perception today among many Chinese people that domestically produced drugs are inferior in quality to foreign ones, which is partly true considering that China is not researching and developing its own medicine at a high enough rate. The economic issue of the pharmaceutical market in China is that foreign drugs require high-cost and out-of-pocket payment because the medical insurance from the government only pays for domestic medicine, which doesn't cover certain diseases due to issues such as a slow drug approval period and low R&D investment. In order to deal with this problem, in October of 2017, the Chinese government implemented a policy that allowed importing of foreign medication and medical tools that are urgently needed by Chinese people through an easier procedure. However, this policy is hard to implement since it disadvantages domestic pharmaceutical producers and regulatory authorities, generating voices of opposition.

Considering the obstacles, this paper argues that the optimal policy is to reimburse the cost of more foreign drugs until the quality of Chinese drugs improves through R&D. The money for R&D will come from government tax breaks, especially for private companies, which constitute 70% of the industry. In order to make up for the money lost in tax breaks, the government should spend less in other sectors, considering the importance of healthcare to citizens. With a larger healthcare budget, more money can be invested in state-owned pharmaceutical companies in China, who are also responsible for producing domestic medicine.

Keywords: Pharmaceutical market, pharmaceutics in China, healthcare, research and development, healthcare cost

Introduction

During the 2000s, there were many counterfeit pharmaceutical drugs produced in China, which resulted in harm or even death of patients who took them. Consequently, there is a perception today among Chinese people that domestically produced drugs are inferior in quality to foreign ones, which is partly true considering that China is not researching and developing its own medicine at a high enough rate. The economic issue[1] (See endnotes) of the pharmaceutical market in China is that foreign drugs require high-cost out-of-pocket payment because medical insurance from the government only pays for domestic medicine, which still doesn't cover certain diseases due to issues such as slow drug approval period and low R&D investment.[2]

The Golden Horse award-winning movie *Dying to Survive* is exemplary of this issue and talks about a shopkeeper in China who illegally smuggled medicine for leukemia from India. This medicine costs 5000 USD per bottle in the hospital, which is unaffordable for most Chinese citizens.[3] The shopkeeper sells it at a much lower price and receives plenty of customers. In the end, this shopkeeper gets caught by the government but the hospitals end up recognizing the high cost of this leukemia medicine and lowering their prices. This movie successfully portrays how China's pharmaceutical system is not able to provide medicines for certain diseases because it hasn't opened up its drug industry for foreign enterprises and that its domestic industry is underdeveloped in some areas.

In order to deal with this problem, in October 2017, the Chinese government implemented a policy that allowed importing of foreign medication and medical tools that are urgently needed by Chinese people through an easier procedure. Part of the policy includes relying on foreign clinical trial data to decide if a drug should be approved and used in the domestic market. According to the state-run *People's Daily* newspaper, this policy change significantly expedited the process of reviewing a medication before it comes out; previously, overseas medications required more than six years for Chinese regulators to test and approve, thus delaying the time for patients to access them.[4]

However, this policy is hard to implement since it disadvantages domestic pharmaceutical producers and regulatory authorities, creating opposition. For example, the Chinese government provided special economic zones (SEZs) for its domestic manufacturers and companies in order to economically favor them over foreign operations. But by permitting overseas influence entering the Chinese market, the original advantage for the domestic industry will diminish. Domestic drugs will be in competition with multinational ones, and Chinese consumers would prefer foreign drugs due to the belief that they are higher in effectiveness and safety. Domestic enterprises made this policya challenge to implement, even though the policy is there to fix China's pharmaceutical market.[5]

China's Pharmaceutical Market

It is important to acknowledge China's market because of its substantial size given its 1.4 billion potential consumers. China's drug market is an ever-growing one, with the trend beginning in recent times. Estimates suggest that by 2023, it will reach $161.8 billion and account for 30 percent of the global market. In 2018, China's pharmaceutical market stood at $134.6 billion, ranking second in the world.[6] Of the total market value, generic drugs (out of patent) take up $85.3 billion, the largest fraction of all. Over-the-counter (OTC) medicines in pharmacies add up to 18.4 billion, constituting 13.6% of the whole. This shows that because of cultural norms, only a small fraction of Chinese consumers choose self-medication (buying medicines from a pharmacy) instead of going to the doctor. Finally, patented drugs, which are much more common in Western countries, created $30 billion of revenue in 2018, making up 22.9% of the total portion.[7]

Distribution of China's pharmaceutical market, as mentioned above: Table 1[8]

As a result, many of the largest transnational pharmaceutical businesses have been increasingly locking their eyes on China and establishing R&D facilities, as they see a lot of potential in its market. China's 1.4 billion population provides many consumers. This large demand for drugs greatly incentivizes domestic and international pharmaceutical companies to innovate and produce medicines. Because of China's aging population and rapid urbanization, there will be more diseases and public health needs appearing, leading to a growth in the drug market. On top of its sizable market, China is also a very strong manufacturer. Manufacturing drugs in China is cheaper than in developed countries because China has a large workforce to perform cheap labor. Actually, China's reputation for being a secondary economic sector exceeds pharmaceutical goods, as it is common to see different products made in China.[9]

Currently, there is a policy in China called "zero markup", stating that public hospitals should obtain all drugs on the essential drug list composed by the government. These essential drugs are to be sold without a price difference between the producer and the hospital, so that it can be more affordable for consumers. Drugs that do not make the list (mainly foreign ones) are not required to be in stock. Patented medications are mostly used by China's upper class since it is more expensive, and its cost is rarely covered by the government.[10] Considering both the large pharmaceutical market and manufacturing ability of China, more transnational pharmaceutical companies will enter China if the government allows, and in turn, people's healthcare would improve significantly, not just the upper class.

The Exclusion of Foreign Medicines in the National Drug List

The National Reimbursement Drug List is the same as the Essential Drug List mentioned above. Before more western drugs were added to the list in the past five years, most drugs on the list were domestic. The two main considerations by the Chinese government to do so is:

1. Build a competitive domestic pharmaceutical industry and not depend on foreign imports

2. High levels of distrust in overseas industries entering China[11]

The reason why China's pharmaceutical system favors its domestic players so much is because it is highly controlled by the government. The role of a government is usually to support and balance its domestic enterprises with foreign competitors. And in China's case, the government has a strong will to develop its domestic drug industry into a world-leading one. President Xi Jinping once said to make China "masters of its own technologies", which includes the pharmaceutical industry. The promotion of domestic pharmaceutics is also part of a national plan called "Made in China 2025", meaning to make China self-sustainable in every industry by 2025. Therefore, it can be concluded that the main motives for China to only reimburse its domestic medicines are strengthening its own technologies and improving its national image.[12]

Distrust in international companies is a main motive for the Chinese government to reimburse foreign drugs. Because there have been cases of bribery happening in multinational pharmaceutical companies, the Chinese government decided to place stricter surveillance on them. The anti-corruption department in China has taken action on these overseas industries, adding regulations such as anti-commercial bribing and anti-unfair competition law. As a result, foreign enterprises are working under heavy scrutiny in China, with many restrictions to expand its influence.[13]

Market Failure

Even though China's pharmaceutical market is growing, many citizens are still unable to obtain the medications they need at an acceptable price, just like in the movie Dying to Survive. Three key reasons for this problem are:

1. Harsh restrictions on foreign medicine

2. The beneficial results of China's research and development (R&D) investment in its domestic drugs are not evident enough

3. Overly-strict regulations for reviewing new medicines

There is a paradox in China's approach to improving its pharmaceutical system. The goal of the government is to expand its own pharmaceutical industry and make China self-sustainable, not relying on foreign products. On one hand, foreign medicines are deliberately set at a sky-high price by the government, so most patients will choose domestic ones instead due to economic concerns. On the other hand, however, China's pharmaceutical industry still comes with flaws, the most notable of which are stated in points 2 and 3 above.

While the Chinese government wants to minimize foreign influences in its pharmaceutical market, their R&D sector requires refining in order to match the level of foreign medicines. "Although the Chinese pharmaceutical industry has been developing fast in terms of market size and revenue volumes, the scale of Chinese pharmaceutical companies remains relatively small with a low market concentration. Therefore, local pharmaceutical companies with higher Research and Development (R&D) input are generally less profitable. Although there have been increases in the number of patented drugs in the pharmaceutical industry in China, patents have made relatively low contributions to the industrial values."[14] Additionally, the intellectual property (IP) of pharmaceutical firms in China are not able to compete with foreign enterprises. This is because Chinese drugs are mainly generic and out of patent, creating less incentive to produce better medicine.

Apart from weaknesses in R&D investment, the process for assessing new drugs is also too extensive, lengthening the time for patients to access them, and increasing the cost. For example, the government set new regulations for the manufacturing process of medicine and for handling chemical waste in drug factories. As a result, this improved the quality of drugs, but also increased its cost for patients because there are more steps involved in making the medicine. Secondly, because there are strict standards for registering new drugs and licensing existing ones, many new drug developments have been hindered by the government. As Table 2 shows, different drugs (y-axis) have undergone longer review times over the years (x-axis).[15] To sum up, China is overly circumspect in regulating its domestic pharmaceutics, leading to delays.

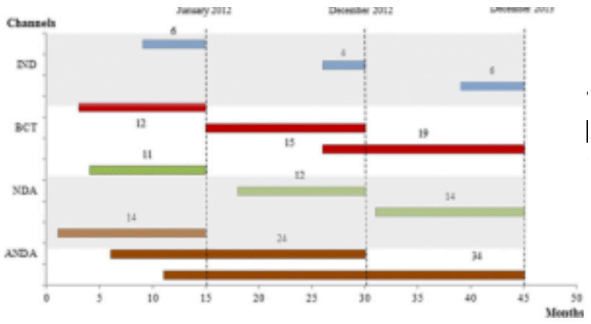


Table 2: Average waiting time for technical review of chemical drugs [16]

Other evident problems are in areas like price caps and market entry. The government is restricting drug spendings on companies by using price caps and profit-margin regulations. With these restrictions, companies in the innovation, manufacturing, and distributing sector will have little market incentive to produce higher-quality medicine, because they wouldn't profit enough from consumers. Even with these regulations, medicine still makes up for half of the total health spending by the government, 43 percent of which are used for patients staying in hospitals (inpatient), and 51 percent for those who leave after checking with the doctor (outpatient). The greatest problem is that China currently lacks the technology to produce certain medicines and has many medicines in the market whose effectiveness are unknown, because not enough investment is made by the government.[17]

Table 3: An increase in R&D government spending [18]

In general, there are many areas to improve in China's innovation in the pharmaceutical industry. Limitations in industrial, financial, institutional, and academic sectors are the main sources that impede the progress of domestic medicine innovation. However, China's strengths lie in its ever-growing drug market and increasing R&D fundings (Table 3). But for

China's pharmaceutical system to be well-rounded and stronger, the sector needs to invest more in developing new drugs and at the same time, embrace the "global value-chain" of drug production by contributing their own strengths.[19]

Government Intervention Through Policy

Recognizing the problems with its pharmaceutical system, the Chinese government implemented a policy to solve them. This policy loosened the grip on foreign medications and sped up the process for drug regulators. In terms of foreign pharmaceutics, there is an upward trend of foreign pharmaceutical investments and overseas medicines entering Chinese markets. After joining the International Council for Harmonization in 2017, China announced plans to integrate international regulatory guidelines into its own industry, regulating system, and R&D. In consequence, these implementations helped China better connect with the global community in terms of regulation standards, and also allowed Chinese people to better access foreign medicine.[20]

In 2017, China followed its plan and added more than 100 foreign drugs such as AstraZeneca's Brilinta (ticagrelor), and cancer treatments like Roche's Herceptin (trastuzumab), MabThera/Rituxan (rituximab), Avastin (bevacizumab) and Tarceva (erlotinib), so people can use these drugs without paying because it's covered by the government. Before these medicines were added, some diseases were untreatable because the domestic medicines for them were underdeveloped. By adding western medicines to the market, which are generally higher in quality, people could finally access the treatments they need. In 2018, new cancer treatment methods also became part of the healthcare insurance list.[21]

In addition to the policy's goal of making foreign medication more accessible, China has been making efforts to match its vast array of drug demands by updating its regulation and licensing process for overseas enterprises. In 2019, the Center for Drug Evaluation and the Chinese Food and Drug Administration (CFDA) decreased the time for reviewing domestic and international medicines, making the addition of new drugs quicker than in the past (Table 4). Furthermore, in order to speed up the time for drugs to be released, authorities also removed the need for clinical trials prior to the release of medicines. What used to slow down the availability of Chinese market drugs is now eliminated. China now accepts foreign drugs that have been approved in their respective countries, and these drugs will only need to undergo fast-track review in China without conducting further testing. Hence, the NationalMedical Products Administration (NMPA) changed its testing obligations concerning the quality of non-patent drugs (generic) so that industry standards are increased and foreign drugs can be reviewed faster.[22] Since people are getting the medicines they need more [23] quickly, this policy is helpful for China's citizens.

Table 4: Number of approved new medications by the CFDA, 2009-2012.[24]

Category 1.1 (blue): Medicines that have not been approved in any country

Category 3 (Orange): Medicines that are only sold in foreign countries

Category 4 (Green): Medicine with changes in its component but is already sold in China

Category 5 (Green): Medicine already being sold in China but with a changed dose method

Category 6 (Green): Medicines that completely follow Chinese standards

China has been working on a plan to provide better healthcare for every citizen, originally to be completed by 2020. The main goals of the plan are making sure drugs are safe to take, effective at fighting the disease, easily accessible, and within the affordable range. All of these goals are part of the larger National Pharmaceutical Policy. Actions such as regulating the production process of drugs (manufacturing, distribution, safety monitoring) have been taken by the China Food and Drug Administration. Because the prices for medicine are heavily regulated in hospitals, there is less economic incentive for the sellers (Table 5). To solve this problem, the National Development and Reform Commission eliminated price caps in June 2015. Additionally, there will be more drugs supplied to retail pharmacies instead of hospital pharmacies, in order to decrease the burden in hospitals and help people with common illnesses get their medicines faster and easier. Currently, pharmacists in China are not playing an important role in healthcare. However, this possible transition to retail pharmacies will require a change in lifestyle, as most Chinese citizens are used to going to the hospital to treat issues with their body, no matter big or small.[25]

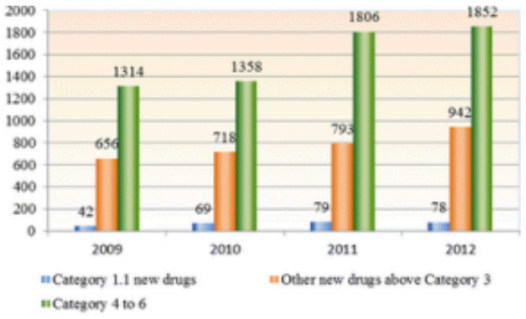


Table 5: The number of patented medicines and licensing percentage throughout the years[26]

Obstacles to the Policy

China is a one-party nation with a government-dominated pharmaceutical industry, where instructions from central authorities are mostly followed without much explicit political resistance. Even so, there are still barriers for the government to implement a certain policy. In this case, the two main challenges for the government are:

1. Dissatisfaction from domestic companies since the policy benefitted foreign products.

2. The problem of domestic R&D inefficiency lingers due to lack of cooperation between research and industrial sectors; importing more foreign drugs is a short-term solution, but China still needs to improve its innovation sector in order to fulfill the "Made in China 2025" plan, and make China more self-sustaining in the long run.

China's domestic pharmaceutical industry is mainly composed of individual distributors and producers scattered all around the country, large in number but varying in size. In 2012, the total number of manufacturers was 4500, with 14,000 distributors. Out of the drugs they produce, 32% are traditional Chinese medicine, which foreign companies do not make. This means that because of the implementation of this policy, 68% of domestic drugs will be in competition with foreign medications, where foreign ones have a better reputation among patients of being more refined in quality. This poses a direct threat to domestic Chinese companies, who have gotten used to developing under a privileged environment created by the Chinese government, as the government needs to boost its pharmaceutical economy.[27] This is an example of politics interfering in the economy.

Among domestic Chinese companies, the ones that will be harmed the most from this policy are the small-scale ones. As of 2012, 70% of China's domestic medicine distributors were small-scale (Table 6) who supplied to their local region. These distributors usually only had under 300 employees with around $3 million US of market value.[28] Wholesalers (large retail merchants), on the other hand, only constitute one third of the total distribution market. However, more and more wholesalers, some of which are foreign, are starting to merge with these

small-scale distributors and acquire them under the wholesaler's name. The reason behind this is that the Chinese government has been encouraging these mergers, with the goal of improving distribution efficiency, compressing supply lines, making regulation easier, and raising the market value of pharmaceutics in different regions. Thus, the losers in this situation are clearly the small-scaled distributors, who are losing their companies and getting eaten by bigger fish.[29] As this policy opens the door for more overseas companies, who are obviously not small-scaled local distributors but instead wholesalers and multinational corporations, the domestic small-scaled distributors will be unhappy and stand up to oppose this policy.

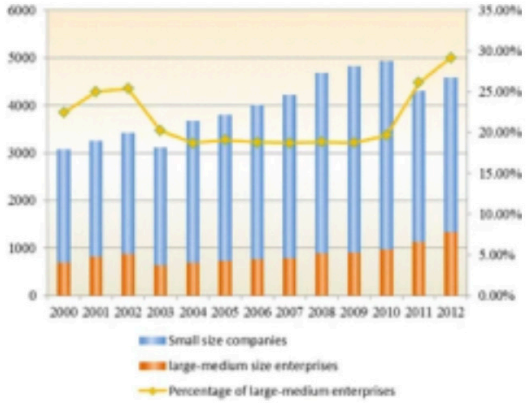


Table 6: The total number of Chinese drug companies, differentiated by size.[30]

On top of the obstacles created by domestic small-scale businesses, R&D in pharmaceutics remains a weakness for China. This policy hasn't addressed the fact that China's own medicines need to be of higher effectiveness and safety. The current approach to conducting R&D is a combined effort between colleges, research institutes, and enterprises, in order to ensure maximum commitment and results. Nevertheless, this three-way partnership is hard to manage, because it crosses between academic and industrial sectors.

While experts from universities and other research institutions are focusing on the scientific nuances of pharmaceutics, companies are considering how to present them to consumers in society. There is constantly a gap between these multiple parties, as the research done in laboratories "does not usually take into consideration the overall development of the pharmaceutical industry. Consequently, the research work may not fully address and respond to the challenges and changing demands of the industry." To sum up, universities and businesses[31] companies have different purposes and goals in society, and these two entities have clashing interests in China.

Another problem in the domestic research sector that this policy has not addressed is the quality of scientific papers and patents in Chinese universities. The main motivation for Chinese scholars is to publish on the Science Citation Index (SCI), so that they receive promotions in work. Incentivised by personal achievement, many scholars publish dozens of SCI papers a year and apply for patents non-stop thus lowering the quality of each piece of research. Hence, most Chinese patents don't last very long due to their low quality. With fewer long-lived patents, the incentive for industries and enterprises will decrease, widening the gap between research and development. All of these factors result in China's innovation output falling behind Western countries who have both a strong foundation and effective method in pharmaceutical R&D. Therefore, the Chinese government should set a new policy to improve the domestic R&D situation and catch up with foreign medications.[32]

Conclusion

Looking at the bigger picture, the policy implemented by the Chinese government in October 2017 is mostly positive. The main beneficiary are patients in China, who had foreign medicines added to their healthcare insurance list, and who can now increase their chances of surviving with more effective medicines. Moreover, as part of the policy required the shortening of the regulation process, patients can receive new drugs faster, thus treating their diseases sooner. However, the risk of deregulation is that some imported medicines might be crude in quality and worsen the users' health. Nevertheless, these occurrences are rare because patented medicines are usually thoroughly reviewed in their respective companies and countries, so they are mostly safe to use. Therefore, this policy benefits the mass population of China who are common patients that need appropriate medications. While Chinese citizens benefited from the reimbursement of more foreign drugs, it undermined the importance of domestic

companies, especially for local-scale ones, and the policy did not propose solutions to this problem. Domestic companies would only benefit if they had a large share of the domestic market, which creates the main source of profit. Once they get in competition with foreign drugs which are better in quality, fewer people will purchase their products.

Hence, the current policy is considered Pareto efficient. The definition of Pareto efficiency is that no entity can be better off without making another entity worse. In this case, the government took away part of the advantage held by domestic pharmaceutical companies and gave it to foreign enterprises. Specifically, the government gave foreign companies a greater share of the Chinese drug market, as an increasing number of Chinese consumers will choose foreign over domestic products in the reimbursement list knowing that domestic medicines lack the ability to treat a range of illnesses. This in turn will decrease the profit gained by the domestic industry, which is damaging. With less money, people working in domestic companies may lose their jobs, making it a negative externality of the current policy.

The plan I believe the government should follow is to reimburse more foreign drugs until the quality of Chinese drugs improves through R&D. While the current policy only addresses the former, it doesn't explain how China could improve its domestic technology in order to reach the "Made in China 2025" goal. The job of the government is to give patients the most effective medicines possible, and for the foreseeable future, foreign medicines are generally higher in quality, so hospitals should use them more until they can be replaced by domestic ones. But in order to reimburse foreign drugs and conduct R&D for domestic drugs at the same time, the national healthcare expenditure ought to be increased. In 2017, only 13.7 out of 2959.5 billion dollars from China's national spending were on healthcare, which is around 0.4%. Although it is a 50.3% increase from the previous year (2016), more money should be allocated for public health considering the significance of this sector.

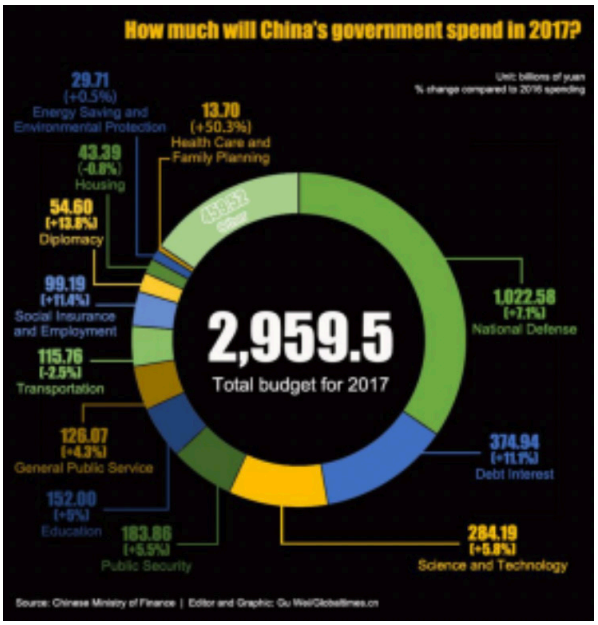


Table 7: Pie chart of government budget in 2017[33]

As of today, 4 years after the policy has been implemented, some foreign drugs are still challenging to access for Chinese patients. My grandma, for example, lives in Chengdu and has had Alzheimer's for 3 years already. In the case of Alzheimer's, only domestic medicines are reimbursed by the government. There are German-made Alzheimer medicines available in the hospital, but they require out-of-pocket payment and are too expensive for my grandma to afford in the long term. Therefore, my grandma decided to take the domestically produced one, which even the doctor admitted was less effective. As a result, her condition gets worse year by year, and our family has no choice but to trust the development of Chinese drugs.

Endnotes

[1]Jingyun Ni, Junrui Zhao, Carolina Oi Lam Ung, Yuanjia Hu, Hao Hu and Yitao Wang, "Obstacles and opportunities in Chinese pharmaceutical innovation", *Research Gate*, March 2017, https://globalizationandhealth.biomedcentral.com/articles/10.1186/s12992-017-0244-6#:~:text=Universities%20 Statistics%20Yearbook-,Pharmaceutical%20industry,new%20drug%20development%20in%20China.

[2]Jiayue Huang, "Clinical trials, drug approvals slow in China as COVID-19 cases grow globally", *S&P Global*, 8 Apr, 2020, https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/clinical-trials-drug-approv als-slow-in-china-as-covid-19-cases-grow-globally-57915442.

[3] Muye Wen, "Dying to Survive," *Beijing Joy Leader Culture Communication Co.*, 2018.

[4] "China prioritizes drug innovation while increasing access to drugs approved abroad," *Generics and Biosimilar Initiative*, last modified November 20, 2020, https://www.gabionline.net/generics/general/China-prioritizes-drug-innovation-while-increasing-access-to-drugs -approved-abroad.

[5] "China policies to promote local production of pharmaceutical products and protect public health," *World Health Organization*, 2017, https://www.who.int/phi/publications/2081China020517.pdf.

[6] "China's Pharmaceutical Industry will be the world's largest in under 10 years", *Daxue Consulting*, February 9, 2021, https://daxueconsulting.com/pharmaceutical-industry-china/.

[7] Ibid.

[8] Ibid.

[9] Huileng Tan, "China's pharmaceutical industry is poised for major growth", CNBC, 2018, https://www.cnbc.com/2018/04/19/chinas-pharmaceutical-industry-is-poised-for-major-growth.html.

[10] Jielin Du, Jiajia Xie, Yan Qian, Mingyue Wu, Wenjing Huang, Jin Yin, Xin Peng and Dan Deng, "Impact of China's zero mark-up drug policy on drug cost of NCDs' outpatients: an interrupted time series analysis", *BMC Health Services Research*, April 29 2021, https://webcache.googleusercontent.com/search?q=cache:bFoLybNawY4J:https://bmchealthservres.biomedcentr al.com/articles/10.1186/s12913-021-06414-3+&cd=1&hl=en&ct=clnk.

[11] "China Issues 2020 National Reimbursement Drug List", *Eversana*, 2021, https://www.eversana.com/2021/01/04/china-2020-national-reimbursement-drug-list/

[12] Robert D. Atkinson, "China's Biopharmaceutical Strategy: Challenge or Complement to U.S. Industry Competitiveness?", *Information Technology and Innovation Foundation*, 2019, https://itif.org/publications/2019/08/12/chinas-biopharmaceutical-strategy-challenge-or-complement-us-industry .

[13] "Corruption Risk in the Chinese Pharmaceutical Market", *Deloitte*, 2013, 2, https://www2.deloitte.com/content/dam/Deloitte/cn/Documents/finance/deloitte-cn-fas-corruption-risk-chinese pharmaceutical-mkt-en-150813.pdf.

[14] "China's Pharmaceutical Industry will be the world's largest in under 10 years", https://daxueconsulting.com/pharmaceutical-industry-china/.

[15] Jingxi Ding, Yajiong Xue, Huigang Liang, Rong Shao and Yongfa Chen, "From Imitation to Innovation: A Study of China's Drug R&D and Relevant National Policies", SciELO, 2011, https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-27242011000200001.

[16] "China Drug Review Annual Report", 2013.

[17] Qiang Sun, Michael A. Santoro, Qingyue Meng, Caitlin Liu, and Karen Eggleston, "Pharmaceutical Policy in China", *Health Af airs*, August 2008, https://www.healthaffairs.org/doi/10.1377/hlthaff.27.4.1042#:~:text=China's%20key%20pharmaceutical%20pol icy%20goals,medical%20services%20with%20drug%20sales.

[18] "China Drug Review Annual Report", 2012.

[19] Ni, Zhao, Oi Lam Ung, Hu, Hu and Wang, "Obstacles and opportunities in Chinese pharmaceutical innovation", https://globalizationandhealth.biomedcentral.com/articles/10.1186/s12992-017-0244-6#:~:text=Universities%20 Statistics%20Yearbook-,Pharmaceutical%20industry,new%20drug%20development%20in%20China.

[20] Jessica Zhou, "China Focusing Innovation Through ICH Global Regulatory Vision", *DIA Global Forum*, 2021, https://globalforum.diaglobal.org/issue/august-2021/china-focusing-innovation-through-ich-global-regulatory-vi sion/.

[21] "China's Pharmaceutical Industry will be the world's largest in under 10 years", https://daxueconsulting.com/pharmaceutical-industry-china/.

[22] "What's New for the Administration of the Drug Post-approval Changes in China?", *Tigermed*, 2021, https://www.tigermed.net/whats-new-for-the-administration-of-the-drug-post-approval-changes-in-china/.

[23] Binxin Li, Caitlin Hou and Elva Yao, "Further Updates on Patent Linkage and Patent Term Extension in China", *Kluwer Patent Blog*, 2021, http://patentblog.kluweriplaw.com/2021/07/21/further-updates-on-patent-linkage-and-patent-term-extension-in-c hina/.

[24] "China Drug Review Annual Report", 2013.

[25] Yu Fang, "Pharmaceutical Policy in China," in *Pharmaceutical Policy in Countries with Developing Healthcare Systems*, comp. Zaheer-Ud-Din Babar (n.p.: Springer, 2017), 169, https://www.researchgate.net/publication/315846396_Pharmaceutical_Policy_in_China. 26China Universities Statistics Yearbook, Date N/A.

[27] Ni, Zhao, Oi Lam Ung, Hu, Hu and Wang, "Obstacles and opportunities in Chinese pharmaceutical innovation", https://globalizationandhealth.biomedcentral.com/articles/10.1186/s12992-017-0244-6#:~:text=Universities%20 Statistics%20Yearbook-,Pharmaceutical%20industry,new%20drug%20development%20in%20China.

[28] Ibid.

[29] flowerhouse, "How to navigate the challenges of a changing Chinese market", *Inceptua*, 2019, https://www.inceptua.com/how-to-navigate-the-challenges-of-a-changing-chinese-market/.

[30] China High-tech Industry Statistics Yearbook, Date N/A.

[31] Ni, Zhao, Oi Lam Ung, Hu, Hu and Wang, "Obstacles and opportunities in Chinese pharmaceutical innovation", https://globalizationandhealth.biomedcentral.com/articles/10.1186/s12992-017-0244-6#:~:text=Universities%20 Statistics%20Yearbook-,Pharmaceutical%20industry,new%20drug%20development%20in%20China.

[32] Fisch, Christian & Block, Joern & Sandner, Philipp. "Chinese university patents: quantity, quality, and the role of subsidy programs", *ResearchGate*, 2014, https://www.researchgate.net/publication/284833645_Chinese_university_patents_quantity_quality_and_the_rol e_of_subsidy_programs.

[33] "How much will China's government spend in 2017?", *Global Times*, 2017, https://www.globaltimes.cn/page/201704/1040910.shtml.

References

(2012). China Drug Review Annual Report.

(2013). China Drug Review Annual Report.

(Date N/A). China High-tech Industry Statistics Yearbook.

(Date N/A). China Universities Statistics Yearbook.

(Date N/A). China Issues 2020 National Reimbursement Drug List. *Eversana*. https://www.eversana.com/2021/01/04/china-2020-national-reimbursement-drug-list/.

(Date N/A). China's Pharmaceutical Industry will be the world's largest in under 10 years. *Daxue Consulting*. https://daxueconsulting.com/pharmaceutical-industry-china/.

(2020). China prioritizes drug innovation while increasing access to drugs approved abroad. *Generics and Biosimilar Initiative*. https://www.gabionline.net/generics/general/China-prioritizes-drug-innovation-while-increasing-access-to-drugs -approved-abroad.

(2021). What's New for the Administration of the Drug Post-approval Changes in China?. *Tigermed*. https://www.tigermed.net/whats-new-for-the-administration-of-the-drug-post-approval-changes-in-china/.

(2013). Corruption Risk in the Chinese Pharmaceutical Market. *Deloitte*. https://www2.deloitte.com/content/dam/Deloitte/cn/Documents/finance/deloitte-cn-fas-corruption-risk-chinese pharmaceutical-mkt-en-150813.pdf.

Atkinson, R.D. (2019). China's Biopharmaceutical Strategy: Challenge or Complement to U.S. Industry Competitiveness?. *Information Technology and Innovation Foundation*. https://itif.org/publications/2019/08/12/chinas-biopharmaceutical-strategy-challenge-or-complement-us-industry.

Du, J.L., Xie, J.J., Yan, Qian., Wu, M.Y., Huang, W.J., Yin, Jin., Peng, Xin., & Deng, Dan. (2021). Impact of China's zero mark-up drug policy on drug cost of NCDs' outpatients: an interrupted time series analysis. *BMC Health Services Research*. https://webcache.googleusercontent.com/search?q=cache:bFoLybNawY4J:https://bmchealthservres.biomedcentr al.com/articles/10.1186/s12913-021-06414-3+&cd=1&hl=en&ct=clnk.

Ding, J.X., Xue, Y.J., Liang, H.G., Shao, R., & Chen, Y.F. (2011). From Imitation to Innovation: A Study of China's Drug R&D and Relevant National Policies. SciELO. https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0718-27242011000200001.

Fang, Y. (2017). Pharmaceutical Policy in China, in *Pharmaceutical Policy in Countries with Developing Healthcare Systems*. (169). https://www.researchgate.net/publication/315846396_Pharmaceutical_Policy_in_China.

Fisch, C., Block, J., & Sandner, P. (2014). Chinese university patents: quantity, quality, and the role of subsidy programs. *ResearchGate*. https://www.researchgate.net/publication/284833645_Chinese_university_patents_quantity_quality_and_the_role_of_subsidy_programs.

Flowerhouse. (2019). How to navigate the challenges of a changing Chinese market. Inceptua. https://www.inceptua.com/how-to-navigate-the-challenges-of-a-changing-chinese-market/. (2017). "How much will China's government spend in 2017?. *Global Times*. https://www.globaltimes.cn/page/201704/1040910.shtml.

Huang, J.Y. (2020). Clinical trials, drug approvals slow in China as COVID-19 cases grow globally. S&P Global. https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/clinical-trials-drug-approv als-slow-in-china-as-covid-19-cases-grow-globally-57915442.

Li, B.X., Hou, C., & Yao, E. (2021). Further Updates on Patent Linkage and Patent Term Extension in China", *Kluwer Patent Blog*. http://patentblog.kluweriplaw.com/2021/07/21/further-updates-on-patent-linkage-and-patent-term-extension-in-c hina/.

Ni, J.Y., Zhao, J.R., Ung, C.O.L., Hu, Y.J., Hu, Hao., & Wang, Y.T. (2017). Obstacles and opportunities in Chinese pharmaceutical innovation. *Research Gate*. https://globalizationandhealth.biomedcentral.com/articles/10.1186/s12992-017-0244-6#:~:text=Universities%20 Statistics%20Yearbook-,Pharmaceutical%20industry,new%20drug%20development%20in%20China.

Sun, Q., Santoro, M.A., Meng, Q.Y., Liu, C.L., & Eggleston, K. (2008). Pharmaceutical Policy in China. *Health Af airs*. https://www.healthaffairs.org/doi/10.1377/hlthaff.27.4.1042#:~:text=China's%20key%20pharmaceutical%20pol icy%20goals,medical%20services%20with%20drug%20sales.

Tan, H.L. China's pharmaceutical industry is poised for major growth. *CNBC*. https://www.cnbc.com/2018/04/19/chinas-pharmaceutical-industry-is-poised-for-major-growth.html.

Wen, M.Y. (2018). Dying to Survive. *Beijing Joy Leader Culture Communication Co*.

Zhao, Ni., Ung, O.L., Hu, & Wang. Obstacles and opportunities in Chinese pharmaceutical innovation. https://globalizationandhealth.biomedcentral.com/articles/10.1186/s12992-017-0244-6#:~:text=Universities%20 Statistics%20Yearbook-,Pharmaceutical%20industry,new%20drug%20development%20in%20China.

Zhou, J. (2021). China Focusing Innovation Through ICH Global Regulatory Vision. *DIA Global Forum*. https://globalforum.diaglobal.org/issue/august-2021/china-focusing-innovation-through-ich-global-regulatory-vi sion/.

Bridging the Broadband Gap: Leveraging Public-Private Partnerships to Level Out Internet Inequality

By Zach Wu

Author Bio:

Zachary (Zach) Wu is a junior at Naperville North High School in Naperville, Illinois. Academically, he is interested in the social studies and humanities, having completed coursework in AP classes, including AP US Government and AP US History. Outside of school, Zach is a captain of the debate team, plays varsity golf, and leads Youth Voices, a 501(c)(3) nonprofit that offers speech and debate programs as a method of empowering young leaders.

Abstract

As a result of huge fixed costs to enter the Internet market, providers of reliable broadband connections are often overly hesitant to expand their service to rural areas. Over the past decade, the chasmic separation between those who have high-speed connections and those who do not has only widened, leading to economic inequality. Without access to broadband, millions of Americans are left unable to participate in the economy, which is now more digital than ever. Internet resources are becoming increasingly critical in telework, job searching, education, and more. As such, it is particularly damning when a large subsection of the population faces severe limitations in the quality of their Internet service. Although many solutions have been explored by the government, most have failed to deliver long-lasting improvements to the broadband gap.

This paper argues that in order for broadband infrastructure expansion to be effective, the optimal policy solution must harness the strengths of both public and private stakeholders in unison. Whereas private actors like Internet providers have little to no organic incentives to build cables in untapped markets where there is no sufficient guarantee of return, and government actors like federal agencies cannot directly finance these projects for fear of exposing taxpayers to risk, neither privatization nor government intervention is normatively preferable. Rather, when governments and private firms share the responsibility, and by extension, the associated risks, of broadband projects, this framework of public-private partnership offers the most potent solution to status quo inequities.

Keywords: Internet, Broadband, Infrastructure, Public-Private Partnership, Economic Growth, Fiber Optic Cables, ISPs, Subsidies, Rural

Economic Problem

The internet market is characterized by returns to scale, wherein companies must spend tons of money to build and install cables to facilitate high-speed internet, but cannot make any money until thousands of people use and pay for them. As a result, there is a functional monopoly in the Internet Service Provider (ISP) market. While there are options on paper, the companies that are involved have formed an economic cartel wherein they only serve certain communities. 13 of the largest providers serve over 80 percent of the consumer market.[27] The dominance of Comcast, Charter, and AT&T deprive millions of a choice in their provider. Monopolization allows a handful of large telecom companies to turn their backs on rural ratepayers without experiencing any repercussions. This is as the sheer volume of consumption that ISPs receive in densely-populated areas more than makes up for what they might lose in a more sparsely-populated consumer market.

Rural Broadband Inequality

Moreover, rural areas also suffer lower quality internet speeds. Because the cables in these areas lack the technical capacity to provide high-speed internet, thus, there are lower speeds at comparatively higher prices. This issue deprives rural areas, particularly in the American South, of a suitable environment for businesses or other economic development because firms would be at a marked technical handicap. In fact, during the last year when the COVID-19 pandemic forced a lot of jobs online, firms moved their business out of the South, because their infrastructure was not conducive to industry.[28] This transition en masse out of this region highlighted the disparities in broadband across the United States. As telework ostensibly becomes more normal moving forward, it is prescient to address these troubling disparities and potentially close them before they become irreparably acute.

Importantly, when identifying the reasons underlying this inequality, there are collective action issues that deter providers. Given that building cables and setting up the appropriate physical infrastructure requires a massive cost, there needs to be a suitable guarantee of return to justify it. However, untapped markets tend to involve risk. Rural areas may not be conducive to certain industries and there is no guarantee that, just because there is internet access, apprehensive businesses will move in and that ISPs will see a sufficient return on their investment. In urban areas, the higher population densities enable a telecom company to serve a greater number of customers from a single switch.

Additionally, internet infrastructure requires continual use and maintenance over time, such as updating, to be viable. Ergo, no ISP wants to be the first company to build the cables and bite the losses. This paper will explore how the government can address the incentive problem on the part of internet providers effectively and examine a policy of public-private partnerships (PPPs) as a solution for the internet inequality that underdeveloped areas face. Different ISPs can offer competing solutions to the market problem that the government is seeking to resolve, and subsidies are issued to the best prospective options.[29] Under a public-private partnership framework, ISPs have enhanced incentives to offer fair solutions to expand service to rural areas, because government subsidies are finite. The government, with the mandate to provide these services at reasonable cost, is unlikely to structure partnerships to facilitate competition. Having multiple ISPs competing over the network, given the economics, is inefficient for companies, who want stronger guarantees of profit in order to buy into the model. However, public investments into broadband should be seen as long-term plays to leverage more robust competition through changing market demand.[30] The use of a PPP structure may provide a novel and innovative outlet to facilitate the nationwide expansion of broadband Internet.

[27] Federal Communications Commission. (2021, February 11). Measuring Broadband America. https://www.fcc.gov/general/measuring-broadband-america.

[28] Britt B.K. (2021, January 15). A Lack of Broadband Access Can Hurt Small Business Growth. *US Chamber of Commerce*. https://www.uschamber.com/on-demand/technology/a-lack-of-broadband-access-can-hurt-small-business-growth

[29] Hovis, J. (2017 May). The Emerging World of Broadband Public–Private Partnerships: A Business Strategy and Legal Guide. *Benton Foundation* 10. https://www.benton.org/sites/default/files/partnerships.pdf

[30] Ibid., 21.

Broadband and Economic Opportunity

Chiefly, a lack of access to the Internet leads to fewer economic opportunities. In the digital age, online resources and connectivity have become key enablers of economic mobility through employment. Across the board, those with reliable broadband service can more readily access the tools necessary to seek out and gain employment. From 2013 to 2015, nearly 80 percent of job-seeking Americans reported utilizing resources and information online as a part of their most recent employment search.[31] Indeed, Internet resources play an increasingly crucial role in job searches for millions. Between 60 and 70 percent of job openings are now posted online, with up to 90 percent of openings that require at least a Bachelor's degree being characterized as such.[32] To find a new job, let alone advance upwards in a current one, the Internet has become essential. Yet, in reciprocal fashion, a lack of access to the Internet condemns an equally sizable fraction of the American population to economic stagnation and the inability to attain employment.

Specifically, using online tools to search for job openings, obtain various skill and competency-based certifications, and apply for positions yields a far faster return for unemployed individuals. One month after initial unemployment, 19 percent of individuals in households with Internet entered or re-entered the workforce, compared to 15 percent of individuals in households without Internet. Over the span of 12 full months, this gap magnifies to 46 percent versus 36 percent for unemployed individuals in households who do or do not use the Internet.[33] Effectively, the proliferation of online resources in the job-seeking process has enabled faster re-employment, but for those without access to a broadband connection, taking advantage of those resources becomes necessarily impossible.

By this token, it can be said that broadband infrastructure and connectivity are a basic prerequisite to upward economic mobility and growth. First, increased broadband penetration gives disadvantaged and underdeveloped areas a path to greater prosperity, by opening up previously-closed opportunities on the Internet. For instance, Addie Maiden, a seventh grade special-needs teacher in Dahlonega, Georgia, has been forced to share her already-limited bandwidth with three sons, making it difficult for the family to work simultaneously. In an interview with CNN, Maiden expressed that in order to do so, the family disconnects the TV, turns off the WiFi on their phones, and "stagger[s]" work "throughout the day."[34] Many rural American families are often forced to use Digital Subscriber Line (DSL), which, while a step above dial-up internet, seldom qualifies as broadband. When the COVID-19 pandemic made telework the new norm, rural families like the Maidens were left in the dust. Furthermore, the direct consequences of inequitable broadband deployment apply to more than just individuals and families. It also holds back economic growth wholesale. Broadband makes areas far more conducive for business, as companies can reliably telework or operate components of their business electronically. Beginning with the idea of greater broadband penetration, between 2008 and 2011, broadband adoption in rural areas was associated with several key measures of growth relating to jobs, such as the total number of businesses and employees.[35]

Availability vs. Adoption

However, it is important to distinguish adoption from availability. On the one hand, availability deals with supply, and the presence of cable lines and Digital Subscriber Line (DSL) networks. On the other hand, adoption deals with demand, and the number of people willing and able to routinely utilize Internet technologies. To further

[31] Smith, A. (2015, November 19). Searching for Work in the Digital Era. *Pew Research Center*. https://www.pewresearch.org/internet/2015/11/19/searching-for-work-in-the-digital-era/.

[32] Carnevale A., Jayasundera T., Repnikov, D. (2014, April 11). Understanding Online Job Ads Data: A Technical Report. *Georgetown University Center on Education and the Workforce*. https://cew.georgetown.edu/wp-content/uploads/2014/11/OCLM.Tech_.Web_.pdf.

[33] Council of Economic Advisers Issue Brief. (2016, March 8). The Digital Divide and Economic Benefits of Broadband Access. *Council of Economic Advisers*. https://obamawhitehouse.archives.gov/sites/default/files/page/files/20160308_broadband_cea_issue_brief.pdf.

[34] Kaur, H. (2020, April 29). Why Rural Americans Are Having A Hard Time Working From Home. *CNN*. https://www.cnn.com/2020/04/29/us/rural-broadband-access-coronavirus-trnd/index.html.

[35] Whitacre, B., Gallardo, R., Strover, S. (2014, September 9) Does rural broadband impact jobs and income? Evidence from spatial and first-differenced regressions. *Annals of Regional Science*, 16-18. https://link.springer.com/article/10.1007/s00168-014-0637-x.

contextualize this, the Federal Communication Commission (FCC) Form 477, which the federal government uses for data on broadband deployment, revealed in 2005 that 96.2 percent of the least-populated ZIP codes had broadband access (there were reported subscribers).[36] In many cases, the mere availability of broadband does not positively track with greater entrepreneurship or job growth. Rather intuitively, simply having more of a good does not automatically necessitate that local businesses immediately know how to use that good or that they will choose to actively adopt it. For the adoption of broadband service, on the other hand, there exists a stronger relationship with job growth and economic opportunity. As more individuals adopt broadband, emergent job and business opportunities become more abundant, due to increased access to markets, networks, and ideas. The salient question is not whether most areas will ultimately have some form of broadband service, but rather that in rural areas, deployment will occur *well after* such services are available in more densely populated areas or that the technology options and/or performance will be drastically different in rural areas.

Government Intervention

Highlighting this availability-adoption distinction is not to argue that broadband should not be expanded and investments avoided unless there is a surefire guarantee of sufficient adoption, but rather, that there are two complementary approaches to the problem. In situations where even basic broadband capabilities are not present, policies that can establish this infrastructure are categorically good, because they raise the baseline capacity of an area for economic growth. However, in order to best capitalize on broadband deployment, more fine-grained and targeted policies can and should be used to encourage the use of Internet resources, especially amongst specific demographics.[37] By nature, communities are not concerned with low levels of adoption if the basic level of availability is insufficient to sustain jobs or businesses. These policy perspectives are not competing, but may exist in tandem. Policies that aim to increase the availability of infrastructure are a prerequisite to widespread adoption, and policies that aim to incentivize regular use of the Internet are what generate sufficient returns on investment.

Failure of FCC Subsidization

Since 2013, the federal government has provided upwards of $22 billion in subsidies and grants to extend and improve broadband capabilities in rural America.[38] Despite this, tens of millions of Americans remain unable to access the Internet, and in turn, today's economy. Notwithstanding the substantial efforts that have been undertaken by federal agencies to level out the broadband gap, there exist multiple barriers that undermine the success of these measures. Most notably, the FCC's own definition of "broadband" greenlights the construction of slow networks far behind the national technological standard. Per their 2015 definition, internet service 25 megabits per second (mbps) to the home and 3 mbps out of the home could qualify as broadband. Having arrived at these numbers as a baseline, the FCC's standards have not been raised to match the technological developments of the present day. In 2020, 90% of Americans had connections of 100 mbps, four times the required 25 mbps downstream standard of service.[39]

However, federal funds, paid for by ratepayers, are distributed to companies at the level of the outdated definition, meaning that the FCC pays providers billions to build networks that are markedly worse than the current national standard. This discrepancy in definition largely explains the failure of the FCC's approach and has entrenched the broadband inequality and widened the chasmic separation between those who have Internet and those who do not. Furthermore, other failed programs and allocations to the broadband issue have run into similar problems. For instance, the CARES ACT meted out $100 million in funds to the Department of Agriculture's Rural Utilities Service, specifically for

[36] Whitacre, B. (2010) Rural Broadband Availability And Adoption In Oklahoma. Choices Magazine 1. https://www.choicesmagazine.org/UserFiles/file/article 160.pdf.

[37] Ibid., 19.

[38] Wheeler, T. (2020, May 27). 5 Steps To Get The Internet To All Americans. *Brookings Institution*. https://www.brookings.edu/research/5-steps-to-get-the-internet-to-all-americans/.

[39] Federal Communications Commission. (2020, April 24). 2020 Broadband Deployment Report. https://docs.fcc.gov/public/attachments/FCC-20-50A1.pdf.

rural broadband service. However, that law expressly prohibits the government to help upgrade and update substandard connections below the FCC's accepted definitions to meet modern benchmarks.[40] As a result, rural Americans that the program targets are left without the means to access quality Internet service.

Public-Private Negotiation Problem

Perhaps the greatest limitation of the aforementioned model, wherein the FCC subsidizes the costs of telecom companies, is that the agency must be in a state of negotiation with the private companies in perpetuity. Essentially, providers will only expand service in proportion to the amount of money that the government is willing to subsidize. One attempted solution to this issue is for the FCC itself to fund the rollouts of infrastructure, and skip the inartful arbitration process with the private sector. In 2009, a Great Recession stimulus bill created the Broadband Technology Opportunities Program (BTOP), which provided $4.7 billion in direct funds to 233 high-speed fiber construction projects.[41] One such project, the Idaho Commission for Libraries grant project under BTOP, exemplifies the success of direct funding. BTOP increased connectivity for 55 Idahoan public libraries, by doubling the computing capacity in 40% of the state's public libraries and increasing average bandwidth by a factor of ten.[42]

While programs like BTOP have had measurable success in a number of cases, the overarching theory of using government interventions to provide direct funding to rural areas remains especially difficult to justify. Characteristically, these areas have fewer subscribers and users of the Internet, which means that high capital costs are often being spent by the federal government, for little perceived gain. From a policymaking standpoint, the aggregate cost of undertaking an intervention does not stack up to the immediate benefits. This approximate calculation frequently deters programs like BTOP from becoming politically feasible.

Optimal Policy

The design of policy related to the promotion of broadband technologies has become more politicized as it involves a jostling scrum of actors and rules at the federal, state, and local levels. Alongside public policy, the broadband debate also features industry associations and consumer advocacy organizations with considerable incentives to influence legislation and bureaucratic rule-making. From the National Cable and Telecommunications Association to the Center for Digital Democracy, the existence of misaligned and often competing incentives seems a near-truism. Among the issues at stake are questions of pricing, service standards, rules for the use of public resources, and privacy. Differences in incentives inform the calculi that the key stakeholders use to make decisions about broadband deployment. Take ISPs, for instance, who lack the incentive to invest in underserved areas with low population density. In the absence of sufficient government intervention where supply does not meet demand, the free market is empirically incapable or unwilling to do so. However, the solution to this incentive issue is not to excise private ISPs from the equation entirely, as conferring the onus of broadband deployment onto government directly is much more difficult to justify normatively. Similar to the cost calculation that private telecom firms make, governments need to treat tax dollars with a measured dose of caution, because if they expend an exorbitant amount on the expansion of a public good for a small subsection of the population, it is difficult to rationalize to the entire electorate. Politicians need to cater to citizens whose policy preferences will dictate re-election odds, and more often than not, rural

[40] Fechner, H., Shapanka, M., Chung, Eric. (2020, March 27). CARES Act Will Support Internet Connectivity for Remote Education, Healthcare, and Work. *Global Policy Watch*. https://www.globalpolicywatch.com/2020/03/cares-act-will-support-internet-connectivity-for-remote-education-healthcare-and-work/.

[41] U.S. Department of Commerce. (2020 January). Broadband Technology Opportunities Program 43rd Quarterly Report to Congress. *National Telecommunications and Information Administration*. https://www.ntia.doc.gov/files/ntia/publications/ntia_btop_43rd_quarterly_congressional_report_01162020.pdf.

[42] U.S. Department of Commerce (2011, December 7). High Quality Public Computing in Idaho Libraries. *National Telecommunications and Information Administration*, December 7 2011. https://www2.ntia.doc.gov/grantee/idaho-commission-for-libraries

[29] Hovis, J. (2017 May). The Emerging World of Broadband Public–Private Partnerships: A Business Strategy and Legal Guide. *Benton Foundation* 10. https://www.benton.org/sites/default/files/partnerships.pdf

Americans who most sorely need better Internet do not fit that mold. Governments must minimize taxpayer exposure to risk, whereas profit-minded companies are financed by investors who are willing to accept the risk of losses.

To this end, neither privatization or government intervention is normatively preferable or efficient in the absolute. Rather, a creative amalgam of, or otherwise collaborative relationship between, public and private actors in broadband deployment supplies the best opportunity for overcoming the restraints associated with each alone. Importantly, there are a few policy levers that facilitate this partnership. First, governments must recognize their own historic failures and shortcomings by reducing the bureaucratic impediments that impede market forces from being most effective. Specifically, governments provide the legal and regulatory regimes to telecom companies that enable them to install physical infrastructure. Secondarily, on the issue of risk-sharing and the allocation of public funds to the project, it is not possible to hide all of the risk of ventures like this. However, there are ways to manage and share risk that advantage both parties. For instance, in Santa Cruz, a 2015 public-private partnership gave the city ownership of the new networking, while the ISP constructed and exclusively operated it for the first several years of its completion.[43] In this scenario, the partnership grants the company an initial period of exclusivity, as a lucrative opportunity to achieve a better return on investment, whereas the city retains ultimate control over the network, giving it discretion over future expansion against unmet demand. Suboptimal solutions are when one side disproportionately bears responsibility. Instead, more meticulous design of PPPs will enable both sides to depend on each other for mutual success. Lastly, in dealing with the incentives and cost calculation of private firms, it is possible to create a stronger assurance of profit. Lackluster internet adoption prohibits firms from seeing enough users in certain regions to justify expanding service. By extension, increasing adoption rates may adequately resolve, or at the very least, mitigate part of the justified hesitancy. In order to do so, policies can engage parties like libraries and schools in relevant areas to create broader support for broadband and Internet technologies. Local governments especially must communicate the measurable impact of broadband on other governance objectives like economic growth or health outcomes.

[43] Lucey, P., and Mitchell, C. (2016 July). Successful Strategies for Broadband Public-Private Partnerships. Institute for Local Self-Reliance. https://ilsr.org/wp-content/uploads/downloads/2016/08/PPP-Report-2016-1.pdf

References

Britt B.K. (2021, January 15). A Lack of Broadband Access Can Hurt Small Business Growth. *US Chamber of Commerce*. https://www.uschamber.com/on-demand/technology/a-lack-of-broadband-access-can-hurt-small-business-growth

Carnevale A., Jayasundera T., Repnikov, D. (2014, April 11). Understanding Online Job Ads Data: A Technical Report. *Georgetown University Center on Education and the Workforce*. https://cew.georgetown.edu/wp-content/uploads/2014/11/OCLM.Tech_.Web_.pdf

Council of Economic Advisers Issue Brief. (2016, March 8). The Digital Divide and Economic Benefits of Broadband Access. *Council of Economic Advisers*. https://obamawhitehouse.archives.gov/sites/default/files/page/files/20160308_broadband_cea_issue_brief.pdf

Fechner, H., Shapanka, M., Chung, Eric. (2020, March 27). CARES Act Will Support Internet Connectivity for Remote Education, Healthcare, and Work. *Global Policy Watch*. https://www.globalpolicywatch.com/2020/03/cares-act-will-support-internet-connectivity-for-remote-education-healthcare-and-work/.

Federal Communications Commission. (2020, April 24). 2020 Broadband Deployment Report. https://docs.fcc.gov/public/attachments/FCC-20-50A1.pdf.

Federal Communications Commission. (2021, February 11). Measuring Broadband America. https://www.fcc.gov/general/measuring-broadband-america.

Hovis, J. (2017 May). The Emerging World of Broadband Public–Private Partnerships: A Business Strategy and Legal Guide. *Benton Foundation* 10. https://www.benton.org/sites/default/files/partnerships.pdf.

Kaur, H. (2020, April 29). Why Rural Americans Are Having A Hard Time Working From Home. *CNN*. https://www.cnn.com/2020/04/29/us/rural-broadband-access-coronavirus-trnd/index.html.

Lucey, P., and Mitchell, C. (2016 July). Successful Strategies for Broadband Public-Private Partnerships. *Institute for Local Self-Reliance*. https://ilsr.org/wp-content/uploads/downloads/2016/08/PPP-Report-2016-1.pdf

Smith, A. (2015, November 19). Searching for Work in the Digital Era. *Pew Research Center*. https://www.pewresearch.org/internet/2015/11/19/searching-for-work-in-the-digital-era/.

U.S. Department of Commerce (2011, December 7). High Quality Public Computing in Idaho Libraries. *National Telecommunications and Information Administration*, December 7 2011. https://www2.ntia.doc.gov/grantee/idaho-commission-for-libraries

U.S. Department of Commerce. (2020 January). Broadband Technology Opportunities Program 43rd Quarterly Report to Congress. *National Telecommunications and Information Administration*. https://www.ntia.doc.gov/files/ntia/publications/ntia_btop_43rd_quarterly_congressional_report_01162020.pdf.

Wheeler, T. (2020, May 27). 5 Steps To Get The Internet To All Americans. *Brookings Institution*. https://www.brookings.edu/research/5-steps-to-get-the-internet-to-all-americans/.

Whitacre, B. (2010) Rural Broadband Availability And Adoption In Oklahoma. *Choices Magazine* 1. https://www.choicesmagazine.org/UserFiles/file/article_160.pdf.

Whitacre, B., Gallardo, R., Strover, S. (2014, September 9) Does rural broadband impact jobs and income? Evidence from spatial and first-differenced regressions. *Annals of Regional Science*, 16-18. https://link.springer.com/article/10.1007/s00168-014-0637-x.

Mass Repression in the Heart of Brazil: Facing the Brazilian Military Coup in Goiás

By João Victor Ataides Marques

Author Bio:

João Victor is a Gap Year student who graduated from Colégio Expressão in Goiânia, Goiás. He was born and raised in Brazil and enjoys exploring and coming across different cultures. He has yet to decide his future career path, but he has strong interests in the social sciences/humanities/communication analysis fields (i.e. history, psychology, sociology, journalism, archaeology). João has been part of Scholar Launch's research practicum, studying 'Mass Atrocities in the Fog of War." João is also the co-founder and CEO of the Constructive Revolutionary Changes Project, a non-profit organization that educates Brazilian youth and that support the democratization of knowledge. Additionally, he is a Global Citizen Year Academy '21 academician and a writer with published journalistic articles and artistic compositions.

Abstract

This study scrutinizes the repression and resistance movements in the state of Goiás, Brazil, during the Brazilian military dictatorship (1964-1985). Specifically, it explores the establishment of the military government, analyzes its repressive aspects, and highlights significant resistance episodes against the military regime in this state, located right in the heart of Brazil, in its midwest region.

First, I focus on the key episodes of national importance that occurred in Goiás, dividing them into repression and resistance. Against this background, I introduce the testimony of a notable participant, Neso Natal, who experienced arrest, exile, and brutal torture in that barbarous age. Finally, I reflect on the political violence and repression experienced in this state, recognizing the notable marks it left in Goiás and Brazil and observing how the globe might reflect on this dictatorship's implications.

Keywords: Dictatorship; Goiás; Coup; Brazil; Military; Repression; Resistance; Torture; Neso Natal.

Introduction

Almost six decades ago, a coup d'etat established a military dictatorship in Brazil, popularly called "Coup of 64." It was one of the longest anti-democratic regimes established in Latin America, lasting twenty-one years (1964-1985). By investigating the impact of the Brazilian military regime, specifically in my home state Goiás, this study aims to answer the question "How was the Brazilian Military Coup faced in Goiás?". The use of this question prompted findings that showed how Goiás played an essential role during this period, including the massive presence of repression and resistance activity there. Thus, the study highlights these events and is accompanied by oral history through an interview with Neso Natal, a resistor and a persecuted victim of the authoritarian military government. Essentially, the research analyzes the military approach to power in Goiás and its repercussions in Brazil.

The coup wrested power from João Goulart, a democratically elected president who assumed his office in 1961. Themilitary perceived Goulart as sympathetic with the communist ideology. His visit to China— a declared communist country —attracted considerable attention. Therefore, the Brazilian Armed Forces' key commanders did not want Goulart as a president and decided to take power in Brazil. They justified the taking of power by saying there was a communist threat that they had to impede. The coup counted on support from conservative groups (e.g., the conservative party, Catholic church, civic anti-communist movements) and foreign contributions from the United States— in the context of the Cold War. Operation Brother Sam was one of the military regime's greatest sources of support, as the United States sent the Navy and Air Force to support the coup. In President John F. Kennedy's words, the reason for Operation Brother Sam was "preventing Brazil from becoming another China or Cuba." Kennedy believed Goulart was becoming too friendly with anti-American radicals in the Brazilian government.

Throughout the military regime, five military presidents assumed power: Humberto Castelo Branco (1964-1967), Artur da Costa e Silva (1967-1969), Emílio Médici (1969-1974), Ernesto Geisel (1974-1979), and João Figueiredo (1979-1985). Also, there were Seventeen Institutional Acts (AIs) issued, so those in power could manipulate and control the government and civil society through a repressive regime. These Institutional Acts enforced censorship, revoked democratic mandates, and even restricted and eliminated political rights. For instance, AI-1 determined the cessation of mandates, the withdrawal of political rights, and indirect presidential elections, all in the name of the LSN (i.e., National Security Law). But the most infamous AI was AI-5, which marked the radicalization of the dictatorship by solidifying absolute power to the military regime. It was after the establishment of this AI that censorship and torture became more intense. For instance, according to Human Rights Watch (HRW), more than 20,000 people were tortured, including children and pregnant women. Worse still, people started disappearing. At least 434 were killed or are still considered missing today. On the other hand, the cultural, artistic, and educational industries were only allowed to produce materials accepted by the military commission. Many had to flee the country, such as Brazilian singers Caetano Veloso and Gilberto Gil, who were a symbol of resistance for criticizing the military regime in their lyrics.

In brief, the 1964 military coup and its use of mass repression marked a troubled period in the recent past of Latin America. Although affecting the entire nation, the dictatorship more severely impacted some specific regions of Brazil. For example, Goiás, which borders Brazil's national capital, Brasília, was a region that first-hand suffered those impacts. There was rigorous political persecution such as the downfall of Governor Mauro Borges, irregular arrests, torture, and severe educational censorship. Goiás experienced major resistance events such as the Assault on the 'War Shot' Barracks (the first-ever armed movement against the military regime), the most significant rebel and rural guerilla movement in the country, and important student revolts as well. Goiás is, therefore, home to many exiled, arrested, and tortured victims, who experienced authoritarianism and censorship during that period.

This study contributes to understanding how mass violence took place in Brazil, which is often debated only on a national level. It aims to break these boundaries and demonstrate that the repression that

occurred in the heart of Brazil, in this state, is relevant for those studying mass violence episodes or simply seeking to learn more about dictatorships and Brazil's history.

Geographical, Cultural, and Economic Context

The occupation of the Goiás' territory began in the 18th century. In 1937, the capital of Goiás, previously located in the city of Goiás, was transferred to Goiânia. The inauguration of Brasília, in the 1960s, also brought economic and population growth to Goiás, due to its proximity to the capital Goiânia. The state of Goiás is located in the central portion of Brazil, more precisely in the country's midwest geographic region. The territory has a wide diversity of fauna and flora, characterized by the tropical climate and the Cerrado vegetation. The state's culture is strongly influenced by the religiosity and traditions of the countryside. The history of the settlement of Goiás is linked to gold mining. The period of gold extraction led to the foundation of historic cities in the state, such as Cidade de Goiás and Pirenópolis. The state has important urban centers, a good infrastructure of services, and an economy based on the primary sector. Tourism is also an important activity in the inland cities of the state.

Under the military regime, the changes in the territorial organization of the states were in charge of the government and were governed by political guidelines. Strong examples of this were the merger of the State of Guanabara, by Rio de Janeiro (1975), and the dismemberment of the South of Mato Grosso (1977). In this context, Federal Deputy Siqueira Campos started a campaign in the Lower House calling for the territorial redistribution of Amazônia Legal (with emphasis on the north of Goiás), since even with investments from projects such as Polocentro and Polamazônia, the north of the state still had poor economic performance and with strong popular movements against the military regime. This way, in 1988, Goiás had its territory divided, and its northern part became the state of Tocantins. Hence, Goiás played a major role during the Brazilian Military Dictatorship, from its geographical localization to its constant urban growth.

Goiânia, capital of Goiás, in aerial view photo, 1960 - Colombo Company

1. Histories of Repression and Resistance

The military takeover meant a democratic rupture, resulting in persecutions, annulment of political rights, imprisonment, repression of students, torture, and exile. Goiás felt the effects of this in many spheres and repeatedly. In this way, this chapter highlights episodes that show the two sides of the military coup: repression and resistance.

1.1. Histories of Repression

The repression in Goiás occurred in several ways, but the most common was throughMilitary Police Investigations (IPMs). The first IPM instituted in Goiás by the military dictatorship occurred soon after the coup with the support of conservative political parties and organized sectors of Brazilian civil society. It covered several segments, such as colleges, unions, peasant leagues, political parties and leftist organizations, the executive and judicial branches, and student movements. In this way, the IPM had the purpose of investigating, from the military point of view, participation in activities considered communist, or that promoted leftist ideologies, which could go against the regime.

Thus, this section elaborates how this organized repression affected Goiás, beginning with the fall of Governor Mauro Borges, re-assessing the massive political persecution, and ending with the severe educational censorship in Goiás.

1.1.1. The Downfall of Governor Mauro Borges (1964)

Goiás Governor Mauro Borges supported the IPMs— and was their victim. A few months after the military coup, on November 26, 1964, troops advanced along the avenues of Goiânia, the capital of Goiás. Airplanes flew low over the capital, threatening to bomb the seat of the state government. In the Esmeraldas Palace, the center of government, Governor Borges, head of the PSD (Partido Social Democrático), tried to resist his deposition, which had been ordered by the military regime. Mauro was also a military man - a lieutenant colonel - but he mistrusted the conservative government. Essentially, he took action in the Legality Network in 1961, together with the governor of Rio Grande do Sul Leonel Brizola, which made Mauro no longer trusted by the military staff. This Legality Network guaranteed the inauguration of João Goulart as president, and because of this, the military junta did not trust Borges and removed the state power of Goiás from his hands.

Controversially, in the beginning, Borges supported the military coup of '64 and took part in a famous meeting called the "meeting of the governors," which endorsed the nomination of Castelo Branco for the presidency, the first military leader to assume power after the coup. Later, seeing that the dictatorship would not be so advantageous, Mauro made enemies in the new regime, such as the then Minister of War General Costa e Silva, who would be the next president after Castelo Branco. For this reason, after being anti-democratically deposed from the government, Mauro also had his political rights revoked in 1966.

It is crucial to analyze Borge's downfall because it led to a wave of repression in Goiânia and across Goiás. For instance, members of government agencies who were condemned as part of a "communist infiltration of the government palace" began to be persecuted and have their mandates revoked, rights denied, and to be imprisoned. Those politicians, such as the Deputy Chief of the Civil Cabinet João Batista Zacariotti, were preventively arrested, with "habeas corpus" denied. Many people, like Zacariotti himself, ended up being tortured afterward.

1.1.2 Persecution and Censorship

1.1.2.1. Persecution

Arrests, torture, revoked mandates, and censorship became more intense than ever in Goiás after Mauro's downfall. Anyone who opposed the military government or conducted themselves in a manner that did not match the military's ideology was already considered a political enemy. Since it was not a democratic government, the opposition, including communist parties (e.g., Communist Party of Brazil, or PC do B), activists, students, and militants, were not welcome to speak up or even express their opinions.

In the first days of April 1964, the military regime enforced the first political arrests in Goiás. One of the most notable arrests was the student leader Tarzan de Castro, president of the PC do B for his participation in the creation of the Peasant Leagues in Goiás, which had as their objective the struggle for land distribution, based on the model of the Cuban revolution. Such leagues clashed head-on with the ideologies of the military regime. Tarzan was taken to the 10th "Caçadores" Battalion of the Brazilian Army in Goiânia, where he was detained and held incommunicado. He suffered physical and psychological torture.

The ruthless torture methods included but were not restricted to: restraints, firing-simulations, electric shocks, "telephones" (slaps with closed hands on the ears), sticks, tying up of the victim's genitalia, threats against family members, and experiments to cause mental confusion. The military widely used torture to obtain confessions from political prisoners by forcing people to sign anything to end the brutal sufferings. Many militants, politicians, students, or simply civilians suffered from these cruel methods (at least 57 were tortured that year.)

One of the political prisoners arrested with Tarzan, Hugo Brockes, soon after leaving prison went to the Notary Public and registered a Public Declaration Deed. In this act of unprecedented courage, this deed narrated the veracity of the facts of his incarceration at the "10th Caçadores" Battalion Barracks in Goiânia, contradicting the IPM Report regarding his arrest and that of other colleagues. He was one of the very few at that time to expose the reality of the military cruelty, recounting his torture

and other sexperiences in prison.

Political persecution in Goiás and Brazil continued until the end of the dictatorship in 1985.

1.1.2.2. Censorship

Another form of repression in constant use during the military regime was censorship. Censorship affected mainly the educational field in Goiás. The situation of the Federal University of Goiás summarizes how censorship affected the academic boundaries of that state.

On April 25th, 1964, the Center of Brazilian Studies (CEB) of the Federal University of Goiás (UFG) director, Gilberto Mendonça Teles, started yet another collegiate meeting. The agenda of the meeting was to deal with the Circular Letter no. 214, of April 20, 1964, which was concerned with the "purges of subversive elements" from UFG and also the "intervention in the Academic Directory" by determination of the rectory. It was only at the end of the meeting that all those present were informed in more detail about the letter "coming from the rectory." It was supported by Institutional Act no. 1 (AI-1) and requested the removal of a list of educators who were considered communists because of their professional backgrounds. The letter also recommended that the Board of Directors take control of the Academic Directory so that the student body would be restricted only to culture and sports. The effects of AI-1 were already being felt at UFG, even before the military government had completed one month.

The creation of the CEB, in 1963, according to its first and only director, Gilberto Mendonça Teles, was related to the effort of strengthening the larger sense of a university project in midwest Brazil. However, with the rise of the military regime, this project came to be considered subversive. Gilberto recalls the beginning of the physical military intervention in the CEB on a Monday in August 1964:

"At the door of the Center was an Army jeep and, next to it, a military man in uniform. Well," said the soldier, "I came to get the communist books that are here. Maria do Rosário answered: 'I am responsible for the library. Did you bring a warrant? No," replied the arrogant Captain Coutinho, "but I'll take them anyway and you can consider yourself arrested. Soon I arrived and Maria told me: 'They've come for the books. Then the captain introduced himself, saluting: 'Captain Coutinho. I've come to get all the books from the communist library. Maria told me: 'He arrested me. The captain heard her and asked me: 'Is she your wife? Then consider yourself released'".

(Carolina Melo, Journal UFG reported 03/27/19)

As he recalls, the director of the CEB noticed "the wayward manner" of Captain Coutinho to contest the warrantless seizure. But he only managed to gain time because they took all the books from the Center for Brazilian Studies with the warrant in hand. Gilberto Mendonça Teles managed to keep only about 20 books, which today make up his private library donated to UFG.

The censorship directly impacted the future of the Federal University of Goiás, the foremost higher education institution of the state. Nevertheless, fortunately, after 1985, the university recovered from that repressive period. The university rebuilt a safe and democratic place for its students, even though UFG will never recover some materials.

1.2. Histories of Resistance

Following the repression of the military coup, resistance started to erupt throughout Brazil. I present in this section the main resistance movements against the military regime in Goiás, beginning with the "Assalto ao Tiro de Guerra," an activity directly connected to PC do B (Partido Comunista do Brasil) militancy. Then, I discuss key student revolts and the largest rural guerilla movement against the military regime.

1.2.1. Before the Assalto ao Tiro de Guerra - Assault on the "War Shot" Barracks (1964)

After the coup, the PC do B, a communist party, began to prepare for armed resistance in August of that same year in response to the repressive events that resulted from the military coup. The Executive Committee of the Communist Party met and wrote a document to the Supreme Court, which warned that

the coup had come to stay and was not something temporary, as many still thought. To combat the coup and defeat the dictatorship, the Communist Party tried to increase the pace of its work in preparation for the armed resistance.

Thus, the first step was to begin choosing locations where the PC do B would deploy future guerrilla and resistance movements. From the beginning - even before the coup - Goiás had called the attention of PC do B's leaders. In their minds, the experience of armed resistance that occurred in the rural Goiás region of Trombas and Formoso was still fresh, a movement in which peasants defeated an extensive process of land grabbing in Goiás (1954-1957). The communist leaders, like Angelo Arroyo and Dynéas Aguiar, had already made a reconnaissance trip to that area in 1963. From there, the party began to establish a presence in the state. Lectures in educational institutes and public spaces to the general populace encouraged, even more, party expansion. Influential activists that would later take part in resistance movements began to join the party, such as Tarzan de Castro, Élio Cabral de Souza, and the student leader Neso Natal. All the reaction to the PC do B's ideology strengthened the idea that only an armed movement would be able to establish a new regime of popular democracy, being a starting point for the crucial resistance movements after the military coup.

1.2.2. The Assault on the "War Shot" Barracks

In November 1964, news of the military's plans to oust Governor Mauro Borges, who threatened to resist, had already spread among the population. The rhetoric became more radical. There was talk of arming the people, as the PC do B was already planning, and the first volunteers for a possible resistance began to appear. The militants were convinced that a grassroots fight was needed at the time. Therefore, this resulted in the idea of an assault on the "War Shot" Barracks to collect armaments and prepare for worse to come.

The assault was planned by a group of people dissatisfied with the military coup who devised a spectacular plan to confiscate weapons from the Brazilian Army to defend the government of Mauro Borges before his downfall and create conditions for a popular and socialist revolution. It was their objective to execute this revolution, inspired by the USSR and Cuban models. They were all left-wing militants, members of the Communist Party of Brazil - PC do B. The assault took place in Anápolis, which is 67.7 kilometers from the Capital of Goiás, Goiânia. The "Assalto ao Tiro de Guerra" was the first-ever armed resistance movement against the military government to defend democracy.

The assault occurred on the night of November 13th, 1964, and was premeditated and plotted by Neso Natal, Daniel ngelo, Jaime José Mendes, Belmiro Vieira de Rezende e José Mendes Vieira. The operation itself was a success, more so than they expected. Neso Natal had already studied the location and realized that it was practically unprotected against such an action. They did not face any security guards there. This way, their only liability or concern was to load the heavy weaponry into the vehicle. There were sixty-four antique rifles, some grenades, and other armaments.

The following morning it was a mess. Army troops moved into Anápolis and placed barriers on all the roads, closing access to the airport. The governor of Goiás, for his part, declared that "the theft of the guns from Anápolis was nothing more than a farce designed to precipitate persecution by the hard-liners who wanted to see him overthrown as head of the state executive."

Overall, the communist activists' action had a national repercussion. It was covered in newspapers headlines throughout the country, including the *Jornal Última Hora*, from São Paulo. It became an inspiration and symbol of courage to people repressed by the military regime. Flávio Tavares, a journalist who was linked to militancy and resistance, wrote the following in his column in the *Última Hora* newspaper about the Assault on the "War Shot" Barracks: "Everything seems to be a typical provocation, carefully planned and prepared, which would consolidate the idea that Goiás had become a center of guerrilla action." (Source: *Última Hora* Journal Headline about the "Assalto ao Tiro de Guerra". The photo depicts the assaulting militants: Neso Natal, Daniel ngelo, Jaime José Mendes, Belmiro Vieira de Rezende e José Mendes Vieira: November 1964 (ANIGO archives).)

However, the situation of Neso Natal and his companions became more and more difficult. The officials were seeking them, and nothing indicated that the governor intended to resist, much less with weapons in his hands. Everything began to fall apart when frightened by the climate of terror that was being created, the driver of the operation came forward to the state security agencies and denounced the entire scheme. Just two days after the assault, on November 16, Mauro Borges' own police arrested the group and seized the weapons in Belmiro Vieira's house.

After the assault on the "War Shot" Barracks on November 23, the Supreme Court acknowledged the habeas corpus of governor Borges. It confirmed that only a majority of the Legislative Assembly could impeach him. On November 26, General Castelo Branco became determined to intervene in the state and remove Mauro Borges. Thus, as his downfall occurred right after the assault, the group commented that the assault on the 'War Shot' Barracks accelerated the intervention in the state of Goiás. Newspapers and investigators said that this robbery accelerated the process of intervention to depose the Governor of Goiás.

The majority of the group was imprisoned for around four months, but eventually, they were granted habeas corpus and released. Even so, they had to flee to another state or country to avoid being persecuted again. Thus, some flew to Rio de Janeiro and others into exile. Neso Natal, whose personal perspective features prominently in this paper, went to Uruguay, then Rio de Janeiro. Finally, he was exiled in the Soviet Union for almost a decade.

1.2.3. Student Revolts (1968)

In addition to the Assault on the "War Shot" Barracks, which was led primarily by students working with the resistance, Goiás had several student revolts during the military dictatorship.

The first student strike in Brazil after the military coup took place in Goiás. It occurred in the city of Goiás, the old Vila Boa, today declared a World Cultural Heritage Site by UNESCO. In 1965, parents and students of the Colégio Estadual Lyceu de Goiás, who were unhappy with the dismissal of the director of that educational institution, promoted a general strike protesting for the permanence of their director. He was dismissed for "suspicious subversive behavior," by the military staff. This movement was enough to be declared by the dictatorship authorities as a communist focus. Several police officers were sent there to quell the protests.

In 1968, a student revolt stood out among the others, given the degree of extreme violence associated with this revolt. This year brought intense episodes in Goiás and Brazil. It was when the murder of the student Edson Luiz occurred, in the Calabouço restaurant in Rio de Janeiro, due to violent police repression during a student protest. It had repercussions in the entire country, making outraged students go to the streets. This way, in Goiânia, on April 1st, 1968, a protest was held by mainly high school students and student leaders calling for justice for Edson Luiz. In the middle of the movement, the student Ornalino Cândido da Silva was shot in the head. He was 19 years old and also worked as a car washer to financially help his family. The military mistook him as being the student leader Euler Ivo during a chase against the protesting students. Here, it is possible to perceive that the military forces were willing to do anything to eliminate any form of freedom of expression different from their ideology, even if that means killing young students.

For that day and the following day, troops chased the students, culminating in the "Invasion of the Cathedral," when soldiers entered Goiânia's main church to arrest young people who had taken refuge there. The Archbishop of Goiânia, Dom Fernando Gomes, who was present on the day, could not prevent this violence - two students from being wounded by bullets inside the church.

All in all, teenagers and young adults, mainly students just like those participating in the protest of April 1st, were violently repressed simply for opposing the military government. Such resistance was the inspiration for student revolts in states like São Paulo and Rio de Janeiro.

Nowadays, the memories of the student resistance movements against the violence of the military forces at that time stand both as a symbol and an alert for possible future anti-democratic regimes. For this, there is also a monument in Goiânia called "Monumento aos Mortos e Desaparecidos

na Luta Contra a Ditadura Militar" (i.e., Monument to the Killed and Disappeared in the Fight Against the Military Dictatorship) that has the intention of passing these memories of resistance on to the next generations.

1.2.4. Guerrilla of Araguaia (1967-1974)

Another notorious resistance movement was the Guerilla of Araguaia. In 1967, militants linked to the PC do B wanted to imitate resistance inspired by the Cuban Revolution. Moving into remote areas, they tried to set up a rural guerilla movement in the Bico do Papagaio, a region that then belonged to Goiás, on the border with the state of Pará. This region contains a part of the great river Araguaia, hence the name of this guerilla movement.

The militants carried out training, hid fugitives, and planned armed operations there. After the General Médici government (1969-1974) launched intense repression against the insurgents, the guerrilla movement became violent. With operations Sucuri, Papagaio, and Marajoara, the fighting intensified from 1972 onwards. The Commission on Political Deaths and Disappearances (CSMDP) calculated that seventy-nine guerrillas were involved in the guerrilla war. Local peasants also supported the movement. Almost half of the disappearances of opponents in Goiás during the military regime occurred in the Araguaia region. A total of seventy people went missing, sixty-four of them militants and six peasants. Thus, this was the most significant rural guerrilla war against the Brazilian military regime, which had a national repercussion.

2. Memories of Repression and Resistance

This section is primarily based on oral history, complemented by documents provided by prominent participant Neso Natal.

Neso Natal, on August 27th, 2021, during his interview with João Victor A. Marques.

To deeply understand the military coup of 1964 and its impacts on a victim's life, I interviewed Neso Natal. Natal is one of the figures that marked the resistance during the military dictatorship. He was one of those responsible for the Assault on the "War Shot" Barracks, a militant student leader and once part of communist parties, namely the PCB and PCdoB. He was arrested and brutally tortured — physically and mentally.

In this section, I aim to present the statement that Neso Natal provided me during our interview, describing moments that are reflected in his physical scars and precise memories. Natal's history is presented in six sections: Background, Militancy, Arrest, Exile, Torture, and Amnesty.

2.1. Background

Neso Natal was born on November 4, 1942, in Goiandira, a city in the interior of Goiás, where he was raised. "Since I was six years old, in Goiandira, I was heavily influenced by Juca Ferreira, who was from the Communist Party of Brazil (PCB) and worked in the pharmacy with my father," Neso Natal said. Natal told me that the pharmacy where his father worked was a center where people would meet to talk about left-wing ideologies, such as during World War II, when people would comment on Nazism and Fascism.

"Something interesting is that very few people had radios at that time, and in this pharmacy, there was a radio that people listened to, and there was also a map. So Russia's victory in World War II was much discussed. Russia lost more than 20 million people in the fighting, between soldiers and civilians. Some historians claim that this number was even higher. It was people who suffered a lot in that period."

He told me that from a very young age, he could tell that everything revolved around that conversation, praising the ideology of socialism: "I remember that soon after, at that time, they built a huge tower for the 'Oil is Ours' campaign. That was a fight for the people, and the next day the mayor had it torn down, but the persistence of the communists was great and they raised the tower again."

In addition to the towers being torn down, Neso notes that in Goiandira, where he lived, the severity of persecutions was already very high." As Goiandira was the gateway to the communist party in Goiás, from the railroad that passes through Goindira, Catalão, Ipameri and ends in Anápolis, my city was also the focus for persecution and torture."

He specified that they shaved the communists' heads, rubbed tar on them, and pulled their fingernails with pliers. Then, Neso relates that the presence of the communist party in Goiandira was an essential influence on him. "I lived in an environment where there was always talk about misery, poverty, inequality; I always heard about the struggle related to Socialism, the development of this political-economic system, the growth of the Socialist Block (USSR), about Stalin. Stalin was regarded as a god; he was called Papa Stalin, Grandpa Stalin."

Moreover, his father participated in the movement that took Riberão Preto, in 1932, in the Constitutionalist Revolution that was São Paulo's fight against the authoritarian president Getúlio Vargas. Many people in Goiás also participated in this movement against the Paulistas. "Further than being an activist, my father ran for mayor of Goiandira through the communist party. There was even a shootout at one of the political rallies since the other candidate was the extremely conservative Colonel Sinfrônio." As Natal grew up under the influence of political activism inspired by his father and the socio-political context, even when he was still in school, he had already become involved with militancy and left-wing organizations.

Map showing where Goindira is located in Goiás. (Raphael Lorenzeto de Abreu - Image:Goias MesoMicroMunicip.svg, own work)

2.2. Militancy

Within the student body in high school, Neso says he was involved with organizations such as UGES (União Goiana dos Estudantes Secundaristas), which prompted his involvement in future militancy resistance movements. This section provides an overview of Natal's track record as a militant.

"I had a lot of contact with the left-wing people inside UGES. I participated in congresses and, as I already had a tendency, I tried to participate in a revolutionary organization. Later, I entered this revolutionary organization, which was the Brazilian Communist Party (PCB), eager to trigger a revolution, influenced heavily by the Cuban Revolution, the Vietnam War, and the Bolshevik Revolution in Russia in 1917, which are historical events that influenced many democracies and countries in the world."

Neso told me that the knowledge he already had in left-wing ideologies, even though only a theoretical one, helped him to engage relatively early on with PCB. In the 1960s, being part of the Brazilian communist party, he told me he had a sort of "petty-bourgeois tendency" to generate a revolution similar to what the Cubans did. "I thought that by taking the gun, the rifle would make the revolution. I joined the party half dissatisfied with the political position and the fight that should be done through the vote, not through an armed revolution. I was unhappy, but I participated. I even participated in the campaign for the legalization of the Communist Party."

During this time, he studied in the highly reactionary Don Bosco School. The revolutionaries, the people from the left-wing, studied in a public school, in the Lyceu school located in Goiânia. Thus, he transferred to the Lyceu school, and there was a place that supported his involvement in activist movements. "I participated in various student movements in the capital. We had intense activities, like campaigns against bus fare increases and participating in several strikes that UGES promoted."

As if the strikes and student movements were not enough, Neso told me he wanted to do something bigger. "I thought that it was very easy to make a revolution, all you had to do was take the rifle, go up and take the power. I thought that, sometimes for lack of knowledge, even though I had already read some Marxist books, books by Lenin, and Hegel. Soon after that, a workers' congress happened here in Goiânia. It was in this congress that I started to have doubts. I began to realize that the fight for socialism was very peaceful, everything by voting, and I began to think that it wouldn't work that way. However, in this same period, João Amazonas, one of the leaders of the other communist party in Goiás (PCdoB), was here in Goiânia. When he came I began to think that the PCdoB's position was a more revolutionary one. Then, I left PCB and joined the PCdoB."

This change of parties was crucial to Neso's life; the PCdoB adopted the more radical approach Neso was looking for and was the party that led Neso to organize the Assault on the War Shot Barracks. "Inside PCdoB, I met Tarzan de Castro, Elio, Divino, people that were already active as militants. We organized several and constant meetings in Goiânia to put our plans of resistance in action. Then, it was prior

to the coup of 64 that we started planning our first armed resistance movement seeing that a military coup was about to happen. We organized the Assault on the 'War Shot' Barracks."

2.3. First Arrest

After the Assault on the "War Shot" Barracks, the majority of the operations group was imprisoned for around four months. Eventually, they were granted habeas corpus to be released, as was the case of Neso Natal. "We were imprisoned in the 10º BC (a state prison in Goiânia). The police officer took our statement with everyone around, in a swivel chair, each one asking a question wanting to know what was happening and what had happened. After this robbery, we were practically left in the lurch. We were arrested immediately for lack of support because if we had had support these weapons could have gone somewhere else. We would do a sequence; other people would do the distribution."

He also commented that he was not tortured in this first arrest but saw the dictatorship's atrocities directly for the first time there.

"I can honestly say that I was not tortured in this first arrest. It is enough to be arrested for your rights to be curtailed, the methods for taking statements did not let us sleep, but I was not in fact tortured, and I cannot say anything about it in this first arrest. However, I met with people who were already being victims of the experimental and sadist military torture. Leaving Goiânia, we were transferred to the PE - Army Police in Brasília, and there were people with very serious problems, like Pawel Gutko. He was not a member of the party, but he was Polish, his father was Russian, who happens to have been my teacher, and his mother was Polish. They arrested him in Goiânia and started severely torturing him. By the time I was there, you could notice the mental damage the torture caused to him. As a result, he used to burn his whole body with cigarettes himself, drink water from the toilet, and rub his face with his own feces when he defecated. From what I know from the newspapers, he could not stand the torture and did not last much after that. Thus, after being in jail for four months or so, I came out with the habeas corpus that I wrote and knew that if I did not want to end up just like Pawel Gutko, I needed to run away from Goiás."

Neso told me that he and his colleagues started planning to flee to another state or country to avoid being persecuted again because they were already considered a threat and state enemies by the military regime.

2.4. Exile

When the military forces started persecuting the militants again, many of Neso's colleagues went to Rio de Janeiro, but he went to Uruguay. As the political situation in Uruguay was not favorable to foreigners, however, he returned to Rio de Janeiro, hoping to find another way to flee the country. He struggled to find a solution, even entering the Mexican embassy seeking exile, without success and almost being caught. Nevertheless, the PC do B supported him with a ticket and visa to France.

From there, he went to the Soviet Union. Thus, Neso Natal was exiled in Russia for almost a decade and started to build a life there. He studied geology there and met his wife, Maria. In 1975, Neso decided to come back to Brazil. However, the regime had condemned him for political crimes while he was outside of the country, which resulted in Natal being brutally tortured after getting caught by the government upon his return.

"I went to Paris, stayed there for a while. From there I went to Moscow, where I took my higher education course. I was very well received in the Soviet Union. But despite everything you are left with that desire to return to your homeland. But I am sure that if I had stayed here during the Medici Government, which was the most violent government at the time, I would not be here alive and speaking. When I arrived in Moscow, I went to the Party Staff School. I stayed there for a year, studying Marxism. After I finished school, I got an opportunity to transfer to Kyiv to study geology. So I went to become a geology technician, studying in Kyiv where I stayed for almost a year. In this period of one year, I had already met Maria who lived in Moscow and was studying medicine; then, I asked for a transfer to Moscow. In Moscow, I entered the Institute of Geology, now as an undergraduate student. I was studying, and I had a lot of difficulty with mathematics and other subjects. Therefore, Maria got a private tutor, and with a lot of effort, I overcame the difficulties and finished the Geology course. The course lasted

five years. I finished the course that was Geology and Engineering. It was engineering for the construction of mines, dams, and mining galleries. After finishing this course, I returned to Brazil, in 1975, in the Geiser government. Also, during this period in exile, I kept in touch with Gregório Bezerra and with Prestes' family. I arrived in the Soviet Union before Prestes, but he left soon after, and so did I. You may ask me why I made this decision, but the only thing that I had in mind is that Brazil was getting better than it was at that age. I believed I had to come back, I had to continue to fight for my country."

Going through his return to Brazil, Neso told me about what he expected from living in Brazil again and how things went after landing in Brazil.

"When Maria and I came to Brazil, we had a collection of books, paintings, a large number of records of classical music, and we shipped them. I still have some of these books here, but the military took most of them. We arrived in September 1975, and my great joy was to have graduated together with Maria, come back to Brazil, and start working— especially because she graduated in medicine, tropical diseases, and Brazil is a low country. We had a dream of doing something for the country, but when we arrived, still under the military regime, under a rule of persecution, we wanted to go back. Eager to work, I went to São Paulo and published ads in the newspaper O Estado de São Paulo for those looking for a geologist to find me. While not receiving any job proposals, I stayed with Maria and my family.

On November 1st, I received a letter saying that the move from Moscow had arrived and I went to São Paulo. On November 15, 1975; I stayed in São Paulo for three days and went to Santos' harbor. What I did not know is that, during this period, it seemed that the military people were already waiting for me there. They identified the Russian ship and found me. I was kidnapped, arrested, and taken to the former Maritime Police. At the Maritime Police, the torture began. Still, it was lighter than at the DOI-ODI (i.e., one of the most infamous prisons regarding the military hell, located in the state of São Paulo) afterward. They started punching me, asking questions without much interest in my answers, because they wanted to torture me and not get answers. Soon afterward, I was transferred to the DOI-CODI, they took me away handcuffed, hooded, and the torture began."

2.5. Torture

From this point on, Neso related in detail the cruelty he experienced after being kidnapped and tortured. It is essential to mention that while all of this was happening, Maria, Neso's wife, did not know where he was and did not speak Portuguese at all, being pregnant at that time.

Neso told me that it was as if he was entering hell at the DOI-CODI prison, a hell of torture. They took him to a room, stripped him of his clothes, threw water and salt on him, and tied him to a heavy zinc chair. They tied wires to his genitals and fingers. They tortured him with electric shocks; the chair was so heavy that it shook, and his teeth contracted.

"They applied violent shocks, a terrible feeling. Even worse was when they put the wires inside my ear. It was as if you were being hit with an anvil, a noise that made you faint. You would faint and they would come and punch you in the diaphragm so that you would start coming to your senses to continue on and on. In the first room that I was in, which they called the visiting room, they did not want to know any information; not even your name, because they asked and when you were going to answer they came with a shock and you could not even answer because your body was shaking. That was in the beginning. Then they put you in a very dark 1.20-meter cell, with a mattress and a toilet. They put you there and from that point on they would not let you sleep. You spent the whole night taking statements. I realized that there were three teams: teams A, B, and C, that took the statements. You were hooded, but when you came out you could see a little bit of the room, it was a padded room, acoustically protected. The first team would take the statements, these statements would be passed to the other team that would analyze them and ask other questions, and the third team would sometimes have us write up and tell things about the story."

They asked Neso a lot about his participation in the assault on the 'War Shot' Barracks. Also, because he was still speaking a lot of Russian, they usually called him a KGB agent. Because of that, sometimes the torture would go even more intensely. Neso remained arrested there in DOI from the middle of November until January 1976, when he was transferred, and psychological torture began.

"In the hell, at the DOI I was still conscious during the tortures, I still kept my conscience, although I was suffering. I stayed more or less from November until the beginning of January suffering from the physical torture. Then, I was transferred to the DOI in Brasilia, when the psychological violence and mental confusion became part of the torture. There I went crazy, completely out of my mind. They used the closed prison method with heat, cold, lights on, water dripping, hearing things. They divided the day into four. How did they do this? They mess with your psychology: the light would get colder, hotter, they would bring you coffee, food, even if you did not eat. They did this until you had no idea. When we are imprisoned, we look at our fingernails and hair to have an idea of how long we have been there, but after a certain period of time, we lose track. Life no longer has any value, you only think about something else. You think that it does not have any value anymore, you become a lab rat. That's why, when they do the tests, they immediately kill the rats. We were like that. The physical torture, you are conscious and you endure it. Sometimes it gets to a stage where you cannot stand it anymore. In fact, the first time I attempted suicide was in São Paulo during the physical torture. But even so, the psychological torture is absolutely worse."

Neso said it was martyrdom. Sometimes he did not even know his family anymore. He told me he was mute; he was not able to speak.

"I was delirious. They made me become a person without conscience as if I were some kind of animal. A person who does not think nor do anything. It's hard to explain the situation I was in, I became inconsequential in a certain way. When I was transferred from Brasilia to CEPAIGO in Goiânia, if they had placed me with the common prisoners, I think they would not have accepted me. It was the comrades who treated me and took care of me. There, they kept a 24-hour watch on me so that I would not do anything reckless. In CEPAIGO I also tried to commit suicide with a thread, but fortunately, it did not work. Meanwhile, my wife was in Goiânia. She did not speak Portuguese, only Russian. Dr. Rômulo's advice, a lawyer, was that she was not allowed to go out. She always stayed inside the house, and she was pregnant. She could not go out, because they were threatening to deport her. They found out about her somehow. She was so desperate that she went to the Soviet Union embassy, but the ambassador could not interfere in internal Brazilian affairs, the only thing he could offer was a return ticket. She preferred to stay here and wait for the consequences. She did not know what to do, even though she knew about my case in Brazil, about the danger I was in."

Neso said he felt naive about returning to Brazil and not listening to the comrades who said it was better to wait for amnesty. It was a great difficulty with Maria's situation, and she was left on the loose, even though Neso's family gave her excellent assistance.

"My brothers Zequinha and Jehovah, and my mother gave her assistance. She stayed in Brazil and had a very hard time because of the assistance that I could not give her due to my imprisonment. I was imprisoned for almost two years, from November 1975 to June 17, 1977. From the moment my wife visited me in the prison, and I was able to receive visits, I did not recognize the baby in her womb, deliriously thinking it was a pillow. Sometimes I spoke, sometimes not; I was mute, I saw things, I did not recognize people. I thought everyone was a spy, I was upset. It was indeed a very difficult period. It took a long time for improvement to begin. Abrão, a comrade of mine, helped me a lot. He and other colleagues managed to get me out of CEPAIGO and took me to a psychiatric clinic so that they could prescribe me medication. Little by little I got better, but even after I left prison I spent a long time rehabilitating myself. I am not ashamed to tell you this because the reactions that I had were real."

Neso also commented on the consequences of the torture that followed him until nowadays:

"The impressive thing is that even after a long time, you can notice in my house, there are no keys in the doors, I do not lock the doors. I have the greatest dread and fear of closed places. This is because of the prisons. I always have nightmares about the police running after me. When I got out of prison, I could not see any police, even the traffic police I thought were a danger to me. The torture had a very big influence on my irritability. I have marks from the torture; I still have the marks from the dragon chair. They would throw salt and you would move around, even though you were tied up. They threw salt to conduct energy. The seat was made of zinc and when you moved, the salt made wounds, and the wounds with the salt were

an unbearable pain. I have marks on my legs and my teeth are all false because the molar and premolar teeth were crushed because of the strong contractions, with so much force the teeth would break. I have hearing problems, diabetes, I have a lot of irritability, all of this is a consequence of these events."

He was one of the last victims to be released from prison: "I got out, but I was persecuted, I believe not only me but also the other comrades, until 1985."

2.6. Amnesty

Amnesty was granted in 1979. During the Geisel Government, it was already possible to understand that they were heading towards an amnesty. Neso saw that coming: "At the time I was arrested, the DOI was full of people who had already been arrested and were there again. When there was Geisel's immediate reaction, it benefited everyone that was there, even though there was resistance inside from the military forces."

Neso described what it was like receiving amnesty after all the events he suffered:

"Receiving Amnesty was very exciting. Sometimes I was incredulous, wondering if this recognition would really happen. At the Amnesty trial, the Human Rights trial, the Peace Commission trial, I was very distressed. The first trial was a trial that I did not participate in, it was a more ordinary event, and I lost. I became tense and by the second trial, I was already without hope of winning the amnesty, of winning this recognition. There was the trial, they wanted to judge me and I was recognized as a professor because I had been a professor of the Russian language at the Brazil-Soviet Union Cultural Institute - it was listed in ABIN that I was a professor and they wanted to judge me as such. I had a degree in Geology, my profession was not that of a teacher. The reporter had asked that they judge me as a teacher, but Lavonier, who was the president of the Amnesty Commission, did not accept and said that I was a geologist, that I had arrived with a diploma, and that I had been arrested as a geologist. The documents mentioned the arrest of a geologist who studied in the Soviet Union. This thesis was defended and won. I was accepted as a geologist and received a pension as such. It was very emotional. There were several people in the plenary that applauded me and the Commission. Comrades like Marcantônio and João Silva were there. My daughter Maya was there. I was very satisfied and happy with the judgment. This judgment took place at the end of 2004. The important thing is not the pension, not the salary, but the recognition that the State was wrong. That is the most important thing and the main thing."

Neso still thinks that there should have been trials of the dictatorship's oppressive leaders to repair a lack of justice left by the military government.

"I am in favor of the trial of the torturers. For me, torturers should not be given amnesty. There is no amnesty for torturers anywhere in the world. There should be a trial for these people, they killed many, not only Brazilians but also Italians, Argentines, Spaniards who were fighting here and who died. In Argentina and Chile, they were prosecuted and convicted. In Brazil, there is resistance, not only from the Armed Forces but also within the Congress, within the Senate, because there are many people there who have no interest. In my time we had Commander Ustra, who has now been condemned; he was a torturer and has nothing to do with Amnesty. Congress and many people think that both sides should have amnesty, and that is just hard to swallow." "I believe that their domination will not be eternal. It will come to an end like the Roman and Byzantine Empire; like Nazism and Fascism will also come to an end, but I believe it still exists today. The Congress and the Senate are full of people from that time, from the Arena. That is why it is difficult to do something. Why hasn't land reform been done in Brazil yet? Because it is very difficult to do it in this situation, in this capitalist regime; and it is even more difficult because of the people that are in the Congress and in the Senate. There are many democrats, but there is a big resistance. There are people who are against amnesty."

All in all, Neso's legacy of resistance and memories of a violent period should never be forgotten, as a lasting reminder of how harmful anti-democratic governments can be, not only in Brazil but across the globe.

Neso Natal and his wife Maria, in 2018, fifty-five years after the military coup

Final Remarks

I have demonstrated in this paper that Goiás was a region that abundantly experienced the Brazilian military dictatorship, passing through severe repression, facing it with considerable resistance, and leaving many disturbing memories and scars in those who survived the atrocities.

Although the documented materials about the military dictatorship in Goiás helped reconstruct the key events of this period, interviewing Neso Natal provided a more transparent and human approach to the research. It contributed with crucial information about the atrocities committed at that time and left explicit the harm that anti-democratic systems can do to human rights. Essentially, Neso Natal is living proof of how precisely the coup affected Goiás and injured the human rights of many Goianos, victims of military authoritarianism.

I should point out that the experience of Neso Natal, unique and important as it is, is one among the thousands affected by the Brazilian military dictatorship. It is essential to recognize the victims' fight against the repression of that government's ideologies and recall their constant fight. This legacy might be a warning for current Goianos and future generations of Brazilians and people all over the world. A warning to take a stand against any kind of governance which leans toward authoritarianism.

Author's Future Steps

I am aware that the research paper does not cover one hundred percent of the events in Goiás, nor is it a complete piece about the Brazilian military dictatorship, which leaves space to continue to engage in the learning of this period. I aim to continue pursuing this path. Specifically, developing further understanding about the Brazilian military dictatorship and studying other mass violence episodes shall be part of my future as a researcher. Being a researcher is indeed a contribution to my intellectual curiosity, but also holds a cultural responsibility to represent the history of my state, and it is an encouragement to anyone who seeks to study human rights abuses in their surroundings. I hope that I demonstrated that the history of the military regime in Goiás matters, regardless of how the years and decades have passed since the events occurred. Such a record should be kept alive, and mass violence should be continuously explored, as a lasting warning against the abuse of power, in Goiás and worldwide.

Acknowledgments

This research forms part of a history research project funded by Scholar Launch into the 'Global Mass Violence' field. The author would like to thank Dr. Vladimir Petrovic for being the Faculty Advisor and guiding the research process, Dr. Lakmali Jayasinghe for being the Teaching Assistant and contributing to the construction of the research paper and the Scholar Launch team that made this study possible. Also, a special thanks goes to Neso Natal, who so openly contributed to oral history in this research and shared crucial materials that supported the study of the topic.

References

Alves, M, M HM. *Estado e Oposição no Brasil* (1964-1984). Petrópolis: Vozes, 1989.

ANIGO, Federal University of Goiás. *Arquivos Revelados: A Ditadura Militar em Goiás Vol. I.* Goiânia: UFG Cegraf, 2016.

ANIGO, Federal University of Goiás. *Arquivos Revelados: A Ditadura Militar em Goiás Vol. II.* Goiânia: UFG Cegraf, 2016.

Arns, Paulo Evaristo. *Brasil Nunca Mais*. São Paulo: Vozes, 1985.

Assembléia Legislativa do Estado de Goiás. "21 anos de arbítrio no Brasil." ALEGO no. [116099.]. Updated March 31, 2021. https://portal.al.go.leg.br/noticias/116099/21-anos-de-arbitrio-no-brasil

Associação dos Anistiados Políticos de Goiás (ANIGO), UFG. *Memórias Transcritas - Depoimentos*. Goiânia: Cegraf UFG, 2013.

Bobbio, N. *Liberalismo e Democracia*. São Paulo: Brasiliense, 2000.

Borges, R. "Como o *golpe de 64* respingou em *Goiás*." (How the *coup of 64* was spread in *Goiás*.) O Popular. Published March 30, 2019. https://www.opopular.com.br/noticias/politica/como-o-golpe-de-64-respingou-em-goi%C3%A1s-1.1764801

Brasil Nunca Mais."BNM - Digit@l." BNM Project. Updated April 6, 2017. https://bnmdigital.mpf.mp.br/

Campos, Mateus. "Geografia de Goiás." UOL, Mundo Educação, 2010. https://mundoeducacao.uol.com.br/geografia/goias.html

Cerqueira e Francisco, Wagner. "O Brasil na Guerra Fria." Brasil Escola. 2015. https://brasilescola.uol.com.br/geografia/o-brasil-na-guerra-fria.html

Champagne, P. *Formar a opinião: o novo jogo político*. Petrópolis: Vozes, 1998.

Clifford, L. "*Amnesty Group Accuses Brazil Of Torturing Political Prisoners*" The New York Times (September 6, 1972) https://www.nytimes.com/1972/09/07/archives/amnesty-group-accuses-brazil-of-torturing-political-prisoners.html

Codato, A. "*O golpe de 1964 e o regime de 1968: aspectos conjunturais e variáveis históricas.*" História. Questões e Debates 40 (2004). https://revistas.ufpr.br/historia/article/view/2735/2272

Comissão Nacional da Verdade. "Relatório Final da Comissão Nacional da Verdade: Mortos e desaparecidos maio de 1974 - outubro de 1985" National Archive Brasil. Updated December 10, 2014. http://cnv.memoriasreveladas.gov.br/

Federal University of Goiás. "Ditadura militar na UFG: Vigilância, perseguição e expurgo." UFG Journal, March 27, 2019. https://jornal.ufg.br/

Goiânia's Municipal Government. "História de Goiânia." Acervo [MIS|GO]. https://www.goiania.go.gov.br/sobre-goiania/historia-de-goiania/

Gouveia, Marcelo. "Comissão Nacional da Verdade: 15 goianos são listados como vítimas", Opção Journal no. [2057]. Updated November 12, 2014. https://www.jornalopcao.com.br/ultimas-noticias/comissao-nacional-da-verdade-15-goianos-sao-listados-como-vitimas-23046/

IBGE. "Geographical Data, Brasil." Instituto Brasileiro de Geografia e Estatística, 1964. https://biblioteca.ibge.gov.br/visualizacao/periodicos/20/aeb_1964.pdf

Truth Commission. "Brazil: Panel Details 'Dirty War' Atrocities." HRW, 2014: https://www.hrw.org/news/2014/12/10/brazil-panel-details-dirty-war-atrocities

Urt, João Nackle. A participação dos Estados Unidos *no golpe de 1964*, UFGD. (February, 2018) https://repositorio.ufgd.edu.br/jspui/handle/prefix/2743

Vieira Borba, C A. "A CONTRA REVOLUÇÃO ANTES DA REVOLUÇÃO: *no golpe de 1964* em *Goiás*." MS, Federal University of Uberlândia, 2013. https://www.historia.uff.br/estadoepoder/7snep/docs/093.pdf

Combatting Sexual Harassment: The Transformational Impact of the MeToo Movement and Its Future Potential

By Megan O'Sullivan

Author Bio:

Megan O'Sullivan is a senior at St Ignatius College Preparatory in San Francisco, California. She hopes to study political science in college and conduct future research in the field.

Abstract

The MeToo movement was created in 2006, but its massive popularity in 2017 led it to shift away from its original focus. Originally, Tarana Burke, a social worker, created the MeToo movement to create a safe space for survivors of sexual harassment to share their stories and listen to similar experiences while also focusing on changing the American culture which perpetuates the issue. The original goals addressed the ways in which a person's identities could change their experiences and understood that more vulnerable people in society are more likely to be sexually harassed. As the MeToo movement gained popularity in 2017, it shifted to focus much more on the individual cases of sexual harassment of public figures rather than on the prevailing issue as a whole. The movement lost much of its intersectionality and failed to address the ways in which American culture perpetuates sexual harassment and how people could effectively change it. In order to truly address sexual harassment, activists need to reexamine the original foundation of the movement and create goals and programs which combine those ideas with the current movement.

Keywords: MeToo movement, intersectionality, public figures, feminism, sexual harassment, media

Introduction

"You took away my worth, my privacy, my energy, my time, my safety, my intimacy, my confidence, my own voice, until now." Chanel Miller, a woman involved in a highly publicized sexual assault case, wrote these words in a personal statement addressed to her attacker. 81% of women and 43% of men have experienced some form of sexual harassment or assault and often endure these same feelings in silence (Kearl 2018). Just like Miller, many victims of sexual harassment are finally able to share their stories due to the recent popularity of the MeToo Movement online. Millions of women were able to become part of an online support network which aimed to bring awareness to the pervasive issue of sexual harassment. Despite the MeToo Movement's beginnings with a focus on intersectionality and helping vulnerable populations, the media's focus on celebrity involvement in the movement caused it to become more focused on seeking justice in individual cases than changing the culture around sexual harassment in American society. The creation of a safe place for survivors became secondary to the cause of demanding repercussions for powerful individuals, which does not solve the greater institutional issues that perpetuate sexual harassment. This paper seeks to examine how the movement developed and grew into what it is today, while focusing on how the inclusion and focus on celebrities shifted the movement away from its original goals. It also discusses how the effects of this shift could be addressed by future participants in the movement.

The Origins of The Movement

History of Sexual Harassment

Before the recent explosion of interest, sexual harassment was not discussed or understood by most of American society, and many, especially women, accepted it as a necessary evil of having a job. In 2019, only 310 out of every 1000 sexual assaults were reported to police, and even fewer of these perpetrators faced justice (Morgan, Truman 2019). Lin Farley, a professor at Cornell University, coined the term sexual harassment in the 1970s as a way to raise awareness of the way women were treated while at work (Siegel 2003). According to "A Short History of Sexual Harassment," this movement was an important push in creating laws protecting women from sexual harassment. As part of this process, feminists needed to "persuade the American judiciary that sexual harassment is 'discrimination on the basis of sex'" (Siegel 2003). Since the beginning of the term sexual harassment, social movements, specifically ones born out of significant cultural events, have been an instrumental part in changing and creating new laws.

Emergence of the Phrase "MeToo"

Despite the phrase's recent connection to the hashtag, "MeToo" originated with a youth worker named Tarana Burke in 2006 (Ohlheiser 2017). On her official MeToo program website, Burke cites a story about a young girl sharing her story of sexual abuse as helping spark the original ideas of the movement. Through this event, Burke realized that an important factor in healing from sexual assault or harassment is being able to share your story, which has been found to create "improvements in physical and mental health as well as the development of closer social bonds" (Pennebaker 2000). Burke utilized this idea, as well as her own experiences as a Black woman working with many other women of color, to create the goal of her movement as giving "young women, particularly young women of color, a sense of empowerment from the understanding that they are not alone in their circumstances," also known as "empowerment through empathy." Burke's original MeToo movement sought to spread their cause through the creation of "an extended network" of women who could support each other. Burke's ideas for the movement show a deep understanding of the intersectionality of the issue. Intersectionality is the ways in which different parts of a person's identity merge to give them varying aspects of marginalization or advantages in society. Statistically, women of color are more likely to be targeted by a sexual harasser than white women (Hernandez 2000), and gender, race, and class all create inequalities in the workplace, which can lead to greater risks of sexual harassment (Acker 2006). Burke sets out to help victims of sexual harassment by first focusing on those who are the most vulnerable and do not have opportunities to combat or report it. She creates an effective foundation for her movement

because she chooses to address the intersectionality of sexual harassment and create an attainable goal of how to help victims. Burke's MeToo movement continued throughout the 2010s and gained more popularity after cultural events gave people greater insight into the issue of sexual harassment.

Catalysts for Future Discussion

Recently, numerous significant events have occurred in popular culture which have forced American society to evaluate its views on sexual harassment and assault, ultimately offering the MeToo movement the foundation it needed to make change. In 1991, when Anita Hill was brought before Congress to testify about the sexual harassment she faced while working with Supreme Court nominee Clarence Thomas, many people saw no issue with the behavior she described, which included discussing various sexual topics, often specifically aimed at Hill, at work (Totenberg 2018). In fact, after the hearings, 58% of participants in a Gallup/CNN poll said that they supported Thomas' nomination (Sparks 2018). Even many of the lawmakers who were questioning Hill very explicitly denied her story, some even going as far as to refer to it as "perjury" and "fantasy" (Totenberg 2018). Thomas was ultimately confirmed to the Supreme Court, which seemed to prove the public's views about sexual harassment at that time (Totenberg 2018). The backlash Hill received for her testimony exemplifies the ways in which the culture at the time did not perceive sexual harassment as an issue. Even with increased media scrutiny on the topic during the hearings, the culture did not shift its views; however, the increased attention introduced the term to many people and set the stage for future discussions.

One of these discussions occurred during Donald Trump's 2016 presidential campaigns. During the campaign, many instances came to public attention of Trump joking about or describing how he sexually harassed women (Gibson et. al 2019). After Trump's election, many people were enraged that the nation seemed to confirm the idea that sexual harassment was not a big enough deal to stop someone from holding the highest government office, and even though many people reacted with public protests about women's rights, not much of the focus was on sexual harassment itself (Gibson et. al 2019). The events created a culture in which many people knew about sexual harassment, but not many addressed it, which created a perfect opportunity for the MeToo movement to use social media to spread awareness and demand accountability.

MeToo's Resurgence in 2017

Many reference the beginning of MeToo and the topic of sexual harassment in the news as a series of articles published by Jodi Kantor and Megan Twohey in the New York Times in October 2017. The articles put forth a detailed account of how Harvey Weinstein used his power to sexually harass young women and cover it up (Kantor, Twohey 2017). In their book She Said, the pair stated that their investigation began with Kantor reaching out to many famous women and people who worked with Weinstein. Despite finding many victims of Weinstein, these women were either unwilling or unable, due to non-disclosure agreements (NDAs), to go on record. Eventually, Kantor and Twohey, by creating a network of women who agreed to share their stories together, were able to publish their article on October 5, 2017. In it, they establish Weinstein's record of assaulting women and covering it up:

After being confronted with allegations including sexual harassment and unwanted physical contact, Mr. Weinstein has reached at least eight settlements with women, according to two company officials speaking on the condition of anonymity. Among the recipients, The Times found, were a young assistant in New York in 1990, an actress in 1997, an assistant in London in 1998, [and] an Italian model in 2015. (Kantor, Twohey 2017)

Harvey Weinstein's story exemplifies the way the culture of a company can allow its most powerful members to exploit their power to harm workers lower on the corporate ladder. His power within the company also allowed him to continue his cycle of abuse, which spans his decades-long career in the film industry, without fear of anyone speaking out. Weinstein had the power in his company and the film industry to silence those who would stand against him. As stated in one of Kantor and Twohey's later articles about the Weinstein company, the culture of a company like this can often force other employers to become complicit in and normalize this behavior. This article helped expose what many referred to as the "open secret" of the

movie industry and created much more media attention for the issue of sexual harassment. Unfortunately, the media attention seemed to focus only on certain aspects of the issue and the movement to solve it, specifically by putting the focus on highly visible public figures.

After the publication of these highly influential articles, the MeToo movement spread even more, especially on social media websites such as Twitter. The first usage of the words MeToo in 2017 came in October after a tweet from Alyssa Milano, which read "If you've been sexually harassed or assaulted write 'me too' as a reply to this tweet" (Twitter. October 15, 2017). Burke was not immediately connected with the movement, though. After the phrase became popular on Twitter, many news outlets, such as the Washington Post, published articles acknowledging Burke's involvement in the creation of the movement, as well as the attention she was receiving from many Black feminists and news outlets (Ohlheiser 2017). Despite not being credited immediately with the movement, Tarana Burke's vision of "empowerment through empathy" came through as survivors, 12 million in the first 24 hours of the hashtag (CBS 2017), shared their stories and heard from others like them. In addition to people sharing their stories, awareness of the issue rose tremendously, with searches for topics relating to sexual harassment and assault reaching an all-time high at 86% higher than expected (Caputi, et. al 2018). The far-reaching nature of social media platforms ultimately created a path to the ultimate goal of Burke's original MeToo Movement: creating a network of survivors who could share their stories and support one another.

However, media outlets did not focus on this aspect of the movement so much as they focused on the public figures involved on both sides of the movement. Many articles were written about famous men and women coming forward with their stories and about famous men facing consequences for previous sexual harassment claims. The most famous of these cases were Louis CK, Kevin Spacey, and Matt Lauer (Carlsen, et. al 2018). These men subsequently faced repercussions, such as Spacey's loss of the lead role in House of Cards, CK's canceled production deal with FX, and Lauer being fired from his position as an anchor on NBC's "Today" show (Carlsen, et al 2018). After so long, the public began to listen to those accusing famous men, which worked to create a cultural shift and led to less acceptance of sexual harassment. While justice is being served in these cases, the position of each of these men as public figures played a large part in these consequences, and many men who are exhibiting this same behavior in other sectors are not experiencing the same consequences due to not being involved in public-facing fields. Additionally, not much focus was brought to the pervasive issue of sexual harassment to all people and how to shift the movement to meaningful change. The influence of media sites ultimately caused the public to care more about the recognizable people involved and seeking justice against powerful individuals rather than the original goal of the movement: creating a safe space for survivors to share their stories and changing the culture that supported sexual harassment.

Impact of the Movement

As the ideas of the MeToo movement became more widely accepted among the public, many activists began to push for a shift in the laws and the way the legal system views sexual harassment. After MeToo, 15 states enacted various sexual harassment bills which all aim to solve different issues regarding sexual harassment (Johnson et. al 2019). Additionally, due to the origin of the movement focusing heavily on Weinstein and his use of non-disclosure agreements to silence victims, many state legislatures "are considering or have passed prohibitions on certain types of non-disclosure agreements" (Tippett 249). These states include New York, California, and Pennsylvania, but each of these states have different ideas of what a ban on non-disclosure agreements should look like (Tippett 249). There is no data that non-disclosure agreements, despite their recent popularity in higher-profile cases, are used often enough in sexual harassment claims to warrant this much focus. Non-disclosure agreements may be used in other cases, but it seems that the only times they would be necessary is if the perpetrator or company involved was well-known enough to warrant a public response to accusations. Once again, the focus of the movement is being pulled away from creating broader sexual harassment protections in favor of focusing on the specifics of high-profile cases. Despite the many federal bills that were introduced regarding sexual harassment protections and arbitration agreements

(Tippett 235), only one Congressional bill managed to pass: The Congressional Accountability Act (Tully-McManus 2018). The Act focuses on congressional employees and ensuring that lawmakers are held accountable for their actions, including banning congress people who have been accused of sexual harassment from using taxpayer dollars for settlements (Tully-McManus 2018). This bill once again focuses on a very specific section of the population: congress people and congressional employees. While this act was beneficial for federal workers, it fails to apply these same protections to all workers. Even with all of the backing behind the movement for better legislation for sexual harassment, only specific groups of workers have been protected, which may be a positive effect of the MeToo movement, but it does not apply to enough workers to truly address sexual harassment. To improve the issue of sexual harassment, legislation needs to focus on broader protections in order to focus on changing the cultures which support sexual harassment and protecting all levels of workers, not just those in public-facing fields.

The organization Time's Up, which was founded by women working in the entertainment industry, works to change the culture around sexual harassment and how companies deal with such issues (timesupnow.org/work). According to their website, they are looking to create a conversation between important figures in different industries, activists, and lawmakers about the prevention of sexual harassment. Time's Up is also specifically using the power of its members, notable female entertainment industry leaders such as Ava Duvernay and Shonda Rhimes, to encourage the inclusion of female creators in Hollywood. The organization aims to do this by launching a mentorship program to promote diverse stories in media and challenging "key players" in entertainment to include female directors. Ultimately many of these initiatives were successes, with Amazon, Universal, Warner Brothers, and many famous directors and actors joining the pledge to work with female directors. Despite the narrow focus on entertainment, the actions taken by Time's Up can have a lasting impact on our culture. The inclusion of female creatives behind the camera can improve female representation in film and help viewers better understand the stories of different women. Additionally, Time's Up supports women who bring claims against major companies, such as harassment claims at McDonalds, pregnancy discrimination at Nike, and pay disparities at Google (timesupnow.org/work). Even though not much tangible change occurred because of these efforts, they were important steps to shape companies' and the public's views on how the serious repercussions of discrimination and sexual harassment. By believing these employees and using resources to support them, Time's Up is putting pressure on companies to take sexual harassment more seriously and shift the way they handle it, for fear of public outrage over their actions. While Time's Up's work often focuses on smaller sectors, they are utilizing the attention these sectors receive to change our culture by bringing women into power and holding powerful people and corporations responsible for their actions.

Due to the prevalence of stories of sexual harassment from Harvey Weinstein and many other figures in media, many famous women and men spoke out about their experiences with sexual harassment, which helped to normalize it in public consciousness. According to Chloe Hart's article "The Penalties For Self-Reporting Sexual Harassment," "there is a cultural bias against women who self-report sexual harassment compared to other harassment targets." These biases could be motivating factors for women to stay quiet about this harassment, which normalizes it in office settings. Many women fear the negative connotations that follow sexual harassment, such as being "seen as less moral, warm, and socially skilled (but not as less competent), and .... less suitable for career advancement" (Hart 2019). After #MeToo, however, Hart acknowledges that these biases have shifted slightly, and while she cannot explicitly say whether that was the cause, she believes it may have contributed. The prevalence of celebrities, such as Ashley Judd, Gwyneth Paltrow, and Alyssa Milano, in this media attention may have also contributed to this shift in cultural thought. To many people who implicitly hold these biases, seeing famous women whom they respect and admire speaking about their own experience with sexual harassment may have forced them to reevaluate their ideas about women who report. Ultimately, the MeToo movement, especially the celebrities involved, helped increase cultural awareness about the truths of sexual harassment and shift cultural biases.

What Do We Do Now?

As the movement grew and focused more on awareness, which Time's Up has been aiding, it shifted the focus from creating spaces for women to share their stories, to acting against men who have been accused. In an editorial for the European Journal of Women's Health, Dubravka Zarkov, who is part of the International Institute of Social Studies, discusses her views on the harder questions to answer for the MeToo movement, such as the official end goal of the movement. She says:

First, we should not assume that what is happening among the political and cultural elites will automatically 'trickle down' to the streets. In other words, we should not expect that office workers, teachers, shop owners or policemen will be equally easily publicly 'blamed and shamed' or dismissed from their jobs because they have harassed and assaulted dozens of women (and men). (Zarkov, Davis 2018)

Many positive effects are associated with the recent ousting of several famous sexual harassers, such as a shifting of what is deemed normal as well as an improved working environment in those places. However, this massive push to change companies' cultures is not reaching all levels of society. It is not "trickling down" as Zarkov says. The movement's tools of supporting public accusations and pushing for consequences does not work in less privileged parts of society where not enough public support can be created. The people in these situations, often lower income women of color, have received some legal protections, depending on their residential state, but more needs to be done to create specific regulations. True change would come about by making it successful to hold someone accountable for sexual harassment through the criminal justice system and shifting our culture away from the normalization of sexual harassment.

In order to change culture and policies, the MeToo movement needs to stay true to Tarana Burke's original vision. Burke created a foundation for the movement that was built around intersectionality and a focus on the people most affected by sexual harassment. The loss of this original intention and shift to seeking justice against powerful men has resulted in a lack of true change in policies and cultures at all levels. Despite the media's focus on this aspect of the movement, Burke's original movement has utilized the overall media attention to expand into more issues while still holding true to her original ideas. On the organization's website, they promote specific programs which are working to change American culture by helping underrepresented communities, such as transgender people, women of color, migrant women, domestic abuse survivors, queer people, and more. These different programs are necessary to address the many different aspects of sexual harassment. The movement also focuses on improving gender equality in the workplace and in society, which may help reduce rates of sexual harassment. For those who seek to do their part to stop sexual harassment, taking action by being part of Burke's MeToo movement is the most effective way to do it. In addition to aiding programs they are currently working with, activists can also help by referring back to Burke's original idea of the importance of education.

The MeToo Movement has effectively brought awareness to the issue and prevalence of sexual harassment, but more steps can be taken to educate young people about it in order to make it less acceptable and normalized. One of the main issues with sexual harassment in the late twentieth century was that not many people talked openly or cared about it. Many women were being sexually harassed, but there was not even a term to capture the experience or laws banning it until well into the 1970s and 1980s. As sexual harassment becomes a more mainstream topic, many people are learning that behaviors they may do or experience from a coworker are unacceptable. In fact, younger workers and men are less likely to even realize that they are being sexually harassed (Acker 2006). To combat this, activists need to embrace Tarana Burke's original idea of education materials which schools, communities, and other organizations could use to help their members talk about sexual harassment and assault. By teaching people to recognize these situations from a younger age, it makes it less likely that they will believe these behaviors are acceptable, which could decrease sexual harassment and increase reporting. Unfortunately, many people fear reporting sexual harassment due to "reactions of coworkers, the mishandling of past reports, fear of job loss, and concerns about potential damage to one's reputation following reporting" (Acker 2006). With education, many more people would understand sexual harassment, so each of these

reactions would seem less likely, and the people not already involved may be more motivated to step in to help with education. In a 2007 study about bystander sexual violence prevention education, researchers found that, after this education, both men and women had more knowledge, changed attitudes, and improved bystander behavior for the length of the study: one year. Introducing education about sexual harassment and assault into schools and community groups would help shift the current culture of silence.

Conclusion

Starting in 2006, the MeToo movement aimed to bring more cultural awareness to the issue of sexual harassment and act as a way to support victims. Tarana Burke created the movement with the ideas of intersectionality and addressing the ways sexual harassment affects more vulnerable populations, such as lower income women or women of color. After the publication of articles exposing Harvey Weinstein in 2017, the movement grew on social media, leading to an outpouring of messages from survivors of sexual harassment which increased awareness. Some of those sharing their stories included famous women, who often accused famous men. Unfortunately, the celebrity involvement caused the media outlets to focus more on the involvement of these public figures than the core issues of the movement. In order to right these wrongs, those who want to support the movement need to return to Burke's original vision of intersectionality and using education to change culture. This paper highlighted the shift of the movement and the negative implications of that shift. It also sought to convince people who support the movement to take a step back from the justice-seeking side of the movement in order to focus on actions that can help a broader range of people, such as supporting education about sexual harassment and changing public policy. Throughout this paper, many new questions were raised about related topics, such as how public education could be used to affect a culture and empower underrepresented communities to speak up. The MeToo movement created a lot of positive change in American culture, but more can be done to combat sexual harassment and gender equality through Tarana Burke's ideas.

References

@Alyssa_Milano (Alyssa Milano) "If you've been sexually harassed or assaulted write 'me too' as a reply to this tweet" *Twitter*, 17 October 2017, twitter.com/alyssa_milano/status

Acker, J. "Inequality Regimes." *Gender & Society*, vol. 20, no. 4, 2006, pp. 441–464., doi:10.1177/0891243206289499.

Banyard, V. L., Plante, E. G., & Moynihan, M. M. (2007). Sexual violence prevention through bystander education: An experimental evaluation. *Journal of Community Psychology*, 35(4), 463–481.

Bever, L. "'You Took Away My Worth': A Sexual Assault Victim's Powerful Message to Her Stanford Attacker." *The Washington Post*, WP Company, 29 Apr. 2019, www.washingtonpost.com/news/early-lead/wp/2016/06/04/you-took-away-my-worth-a-rape-victim-delivers-powerful-message-to-a-former-stanford-swimmer/.

Burke, T. "The Movement." *MeToo Movement*, 2006, justbeinc.wixsite.com/justbeinc/the-me-too-movement-cmml.

Caputi TL, Nobles AL, Ayers JW. Internet Searches for Sexual Harassment and Assault, Reporting, and Training Since the #MeToo Movement. JAMA *Intern Med*. 2019;179(2):258–259. doi:10.1001/jamainternmed.2018.5094

Carlsen, A, et al. "#MeToo Brought Down 201 Powerful Men. Nearly Half of Their Replacements Are Women." *The New York Times*, The New York Times, 23 Oct. 2018, www.nytimes.com/interactive/2018/10/23/us/metoo-replacements.html.

CBS. "More than 12M 'Me TOO' Facebook Posts, COMMENTS, Reactions in 24 Hours." *CBS News*, CBS Interactive, 2017, www.cbsnews.com/news/metoo-more-than-12-million-facebook-posts-comments-reactions-24-hours/.

Cillizza, C. "Analysis: Andrew Cuomo Just Did the Least Andrew Cuomo Thing in the Most Andrew CUOMO WAY." CNN, Cable News Network, 10 Aug. 2021, www.cnn.com/2021/08/10/politics/cuomo-analysis-resign-new-york/index.html.

Gibson, C, et al. "Understanding the 2017 'Me Too' MOVEMENT'S TIMING." *Humanity & Society*, vol. 43, no. 2, 2019, pp. 217–224., doi:10.1177/0160597619832047.

Hart, C.G.. "The Penalties for Self-Reporting Sexual Harassment." *Gender & Society*, vol. 33, no. 4, 2019, pp. 534–559., doi:10.1177/0891243219842147.

Hernandez, T K. *Sexual Harassment and Racial Disparity: The Mutual Construction of Gender and Race*, 4 J. Gender, Race and Just. 183 (2000-2001) https://ir.lawnet.fordham.edu/faculty_scholarship/12

Johnson, A, et al. "Progress in Advancing Me Too Workplace Reforms in ..." *National Women's Law Center*, 2019, nwlc.org/wp-content/uploads/2019/07/final_2020States_Report-12.20.19-v2.pdf.

Kantor, J, & Twohey, M. "Harvey Weinstein Paid off Sexual HARASSMENT Accusers for Decades." *The New York Times*, The New York Times, 5 Oct. 2017, www.nytimes.com/2017/10/05/us/harvey-weinstein-harassment-allegations.html.

Kantor, J, & Twohey, M. *She Said*. Bloomsbury Publishing Ltd, 2020.

Kearl, H. (2018). *The facts behind the #metoo movement: A national study on sexual harassment and assault*. Stop Street Harassment.

Me Too. "Organizations in Action." *Me Too. Movement*, 16 July 2020, metoomvmt.org/take-action/organizations-in-action/.

Morgan, R, and Truman, J. "Criminal Victimization, 2019." Bureau of Justice Statistics, Sept. 2020, bjs.ojp.gov/library/publications/criminal-victimization-2019.

Ohlheiser, A. "The Woman BEHIND 'Me Too' Knew the Power of the Phrase When She Created It - 10 Years Ago." *The Washington Post*, WP Company, 28 Apr. 2019, www.washingtonpost.com/news/the-intersect/wp/2017/10/19/the-woman-behind-me-too-knew-the-power-of-the-phrase-when-she-created-it-10-years-ago/.

Pennebaker, J. W. "Telling Stories: The Health Benefits of Narrative." *Literature and Medicine*, vol. 19 no. 1, 2000, p. 3-18. Project MUSE, doi:10.1353/lm.2000.0011.

Race, 4 J. Gender, Race and Just. 183 (2000-2001)
Siegel, R. B. "Introduction: A Short History of Sexual Harassment." *Directions in Sexual Harassment Law*, 2003, pp. 1–28., doi:10.12987/yale/9780300098006.003.0001.

Sparks, G. "Anita Hill's Accusations Did Not Hurt Public Support for Clarence Thomas in '91 | *CNN* POLITICS." CNN, Cable News Network, 18 Sept. 2018, www.cnn.com/2018/09/17/politics/hill-thomas-kavanaugh-sexual-assault/index.html.

Time's Up. "Our Work." *TIME'S UP Now*, 3 Dec. 2019, timesupnow.org/work/.

Tippett, E. C., "The Legal Implications of the MeToo Movement" (2018). Minnesota Law Review. 57. https://scholarship.law.umn.edu/mlr/57

Totenberg, N. "A Timeline of Clarence Thomas-Anita Hill Controversy AS Kavanaugh to Face Accuser." *NPR*, NPR, 23 Sept. 2018, www.npr.org/2018/09/23/650138049/a-timeline-of-clarence-thomas-anita-hill-controversy-as-kavanaugh-to-face-accuse.

Tully-McManus, K. "Congress Passes Sexual HARASSMENT Bill by Unanimous Consent." *Roll Call*, Roll Call, 13 Dec. 2019, www.rollcall.com/2018/12/13/congress-passes-sexual-harassment-bill-by-unanimous-consent/.

Twohey, M, et al. "Weinstein's Complicity Machine." *The New York Times*, The New York Times, 6 Dec. 2017, www.nytimes.com/interactive/2017/12/05/us/harvey-weinstein-complicity.html.

Zarkov, D, & Davis, K. "Ambiguities and Dilemmas around #Metoo: #Forhow Long and #WhereTo?" *European Journal of Women's Studies*, vol. 25, no. 1, 2018, pp. 3–9., doi:10.1177/1350506817749436.

Assessing the Chinese Censorship System: The Case of LGBTQ Representations in Entertainment

By Summer Xu

Author Bio:

Born and raised in Xinjiang, China, Weijing (Summer) Xu is currently a high school senior student studying at Choate Rosemary Hall in Connecticut. In addition to being the Editor in Chief of her high-school Yearbook and a WeChat platform named NAnthology, Summer also exhibits her interest in media and journalism by promoting journalism in local public high schools in her hometown, Xinjiang. Summer also actively participates in school activities as the club president of the Choate Musician Association and active member of the Psychology Club. She is also a dedicated golfer who enjoys cooperating with her teammates in Choate's Varsity Golf team. In the future, Summer hopes to study media and communication, furthering her research in the rise of new media and cultural globalization.

Abstract

This research paper presents a case study of LGBTQ representations in entertainment in China to assess the Chinese censorship system, which is imperfect yet sophisticated. Censorship in China, which is translated in Chinese as "ShenCha," meaning to "check and refine" your work, has been causing heated discussion, primarily as the State Administration of Radio, Film, and Television (SARFT) exerts more power over the entertainment industry. In this paper, the writer compared the example of the film "Farewell My Concubine" in 1993 and the online TV show "Addicted" in 2016, which both contained LGBTQ contents, highlighting the complexity of the censorship system in China and addressing the content, target audience and censorship decisions of both works. This leads us to the preliminary conclusion, which provides a possible explanation of how the complicated relationship between SARFT, the officials, and feedback from audiences influences the censorship system.

Keywords: Media studies, censorship system, censorship policy, LGBTQ+ representation, movies and TV, entertainment

Introduction

Many observers question whether the sophisticated film censorship system in China, carried out by the State Administration of Radio, Film, and Television (known as SARFT), allows little to no flexibility to incorporate non-mainstream artistic elements or represent specific groups of people such as the Lesbian, Gay, Bi-sexual, Trans, and Queer (LGBTQ) people, or even to portray the lifestyles of people with tattoos and colored hair. For example, in 2017, there were two reality shows, *Idol Producer and Rap of China*, that consisted of young idols where many participants had dyed hair or tattoos, mostly small ones on their arms or fingers. After being aired and viewed by the government officials of the SARFT, the show's depictions were deemed inappropriate for the teenagers who were the primary audience of the shows. Future reality show participants with dyed hair were forced to dye their hair back to their natural colors and cover up any tattoos. This event changed future content and added another regulation change to content moderation despite not being listed in the law itself. This provides an example of how the SARFT can censor without listing it in the rules and regulations.

In addition, the movie rating system is still in its infancy in China, meaning that scenes that would make a film R-rated might be directly cut out to be released in Chinese cinemas. The complex system of censorship, which reaches beyond laws and regulations listed in Regulations on the Administration of Movies, is intertwined with elements such as perceived national honor and dignity, officials' preference in the SARFT, traditional Chinese values of the family system, and even socioeconomic factors. Under these circumstances, films and TV series may be banned in China for reasons not listed among the regulations carried out by the SARFT, for example, themes around the representation of the LGBTQ community.

Meanwhile, in recent years LGBTQ issues have been less of a taboo in Chinese society, yet the government maintains a hostile attitude towards its representation in art. LGBTQ activities, such as communities, parades, or singing groups, have been legal since 1997 (Wang, Belair-Gagnon, and Holton 2020). LGBTQ movements were recognized by the Chinese government in October 2017 through same-sex couples' legal rights, including medical and property management through the guardianship system, meaning that they became each other's guardians instead of a married couple so that they still can have legal obligations to each other. In this context, one question to ask is: how is censorship represented in Chinese movies? The following sections will describe the research proposal.

Censorship in China

The Chinese film censorship system, translated in Chinese as "ShenCha," which means to "check and refine" your work, has become a heated topic in recent years in Western and Chinese media, especially since the Western filming industry began to reach the Chinese audience.[44] With the rise of various video platforms on the internet, the ability of Chinese audiences to watch American television shows and movies, whether through Virtual Private Networks (VPNs) or traditional Chinese broadcasters or websites, has substantially increased. In online discussion communities, such as WeChat--a Chinese owned multimedia social platform, and DouBan--a Chinese owned social networking service website that allows registered users to record information and create content related to film, books, music, recent events, and activities in Chinese cities, people are commenting whether the censorship system was appropriate and tailored to a globalized era where information is more free-floating.[45]

Chinese censorship checks, analyzes, and recreates parts of media content for national unity purposes.[46] To achieve this purpose, the Chinese

[44] Wang, Yidong, Valerie Belair-Gagnon, and Avery E. Holton. 2020. "The Technologization of News Acts in Networked News Participation: LGBT Self-Media in China." *International Journal of Communication Systems* 14 (0): 19.

[45] W"Douban - Crunchbase Company Profile & Funding," Crunchbase, |PAGE|, accessed May 11, 2021, https://www.crunchbase.com/organization/douban)

[46] Gary King, Jennifer Pan, and Margaret E. Roberts, "How Censorship in China Allows Government Criticism but Silences Collective Expression," American Political Science Review, |PAGE|, accessed May 11, 2021, https://gking.harvard.edu/publications/how-censorship-china-allows-government-criticism-silences-collective-expression)

government organized the SARFT to control all forms of media directly by dividing the bureau into the following departments: the General Office (the Legal Department), the General Editorial Office, the Film Bureau, the Broadcasting Affairs Supervision Department, the Personnel & Education Department, and the Planning and Finance Department. The SARFT is in charge of China's radio, television and film industry and directly supervises China National Radio, China Radio International, and China Central television.[47] Despite negative impacts that will be discussed later, the SARFT filters out films that are not of high quality, exerts control over copyrights of movies and other media while preventing plagiarism and infringement of intellectual property. However, China's strict censorship system currently faces challenges as well. For example, the establishment of online streaming media such as Chinese Netflix and iQiyi, as well as the sheer quantity of content compared to the government's limited resources, impacts the extent to which government censors can check the content on such streaming media.

According to the Regulations on Administration of Films that was promulgated in 1996, films are forbidden from having the following content:

1. Those endangering the unity, sovereignty, and territorial integrity of the state;
2. Those harming the security, honor, and interests of the state;
3. Those inciting national splitism and disrupting the unity of nationalities;
4. Those divulging state secrets;
5. Those publicizing obscenity or superstitions or playing up violence;
6. Those libeling or insulting other people;
7. Other contents prohibited by the state.[48]

[47] "State Administration of Radio, Film and Television," State Administration of Radio, Film and Television, accessed May 11, 2021, https://in.china-embassy.org/eng/mt/jyjs/t61109.html.

[48] Jeremy Geltzer, "Censoring the Silk Screen: China's Precarious Balance Between State Regulation and a Global Film Market," |PAGE|, accessed May 11, 2021, https://www.swlaw.edu/sites/default/files/2017-04/JIMEL V6, N2 6-Censoring the Silk Screen-Geltzer.pdf)

Methods

The main goal of this project is to analyze the complexity of censorship. Thus, using two case studies that represent diversity in censorship applications is helpful. A multiple case study enables the researcher to explore differences within and between cases. The goal is to replicate findings across cases. Because comparisons will be drawn, it is imperative that the cases are chosen carefully so that the researcher can predict similar results across cases,or predict contrasting results based on a theory. In the project, a comparative case study will be conducted between the film "Farewell My Concubine" and the phenomenal online TV show, "Addicted", to reflect on the complexity of censorship as they received different treatments from the government by looking closely at their contents, social context, achievements, and the different ways they were censored by the SARFT. Meanwhile, a descriptive case study, which is a type of case study that is normally used to describe an intervention or phenomenon and the real-life context in which it occurred, was also incorporated into the research as both examples were analyzed thoroughly. In addition, the project plans to utilize empirical data to come up with themes around politics of the SARFT, public attention, and engagement.

Censorship in LGBTQ as a case study

The development cases below will allow for a consistent description of how films are censored in China.

Case #1: The film "Farewell My Concubine" in 1993

As Chinese society develops rapidly while being influenced by Western culture, there has been an increasing demand for representation of the LGBTQ community in films and other media. Homosexual behaviors are not criminalized in China, but they are not legally recognized.[49] In recent years, activists have

[49] Rebecca Davis, "China's Gay Rights Stance Can't Derail Demand for LGBT Films," Variety, June 06, 2020, |PAGE|, accessed May 11, 2021, https://variety.com/2020/film/asia/lgbt-movies-china-gay-rights-1234625634/)

focused on promoting a more tolerant environment for LGBTQ elements in Chinese media and foreign movies, such as the award-winning Italian coming-of-age movie with a gay theme, "Call Me By Your Name." This film has accumulated a sizable Chinese fan base without ever being screened in China and was removed from screening one day before appearing in various movie festivals. While there are still no explicit regulations that mention the censorship of LGBTQ elements, movies with homosexual themes are still banned from appearing on any streaming platforms. There are exceptions, which makes censorship in LGBTQ theme films and other media more complex. The cases analyzed below show how the context of each case is different, making it interesting to compare and showing the depth of the censorship system. Other types of content have different outcomes.

Historically, movies that revolve around LGBTQ themes were usually censored completely from screening in China, and the SARFT would remove any mention of homosexuality from screening as well. As mentioned before, there are exceptions. The film "Farewell My Concubine" by Kaige Chen in 1993 was one of the examples. Based on the parallel stories of two Beijing opera performers through twentieth-century China, the director weaved in hints of homosexuality, political influences, and even the Cultural Revolution theme that was particularly sensitive.[50] At the same time, elements of traditional Chinese culture, such as Beijing Opera, were also incorporated, which was strongly encouraged by the government at the time. This phenomenal film received significant international attention as it received La Palme d'Or from the Cannes Film Festival in 1993.

The mentioning of these themes led to the prohibition of screening this movie after its release in China. However, it was only temporarily banned for two months before it was screened again after being slightly modified to tone down its indication of homosexual behaviors and suicide.[51] This government reaction was influenced by the massive success of this movie internationally and due to public pressure as the film soon became popular across the country. Its box office performance was outstanding as it achieved a remarkable statistic of $5,985,074 worldwide, including $5,216,888 domestically (87.2%) and $768,186 (12.8%) internationally, a great success in the year 1993.[52] The audience reaction and the international feedback largely influenced the government's decision in terms of the film's screening. However, the film was eventually shortened by 14 minutes after numerous cuts made by the censors, who hold a prestigious position in Chinese society, and who are often graduate students and high-level administrators. Conclusively, the factors that led to this unique case of censorship include the Chinese opera culture and the international feedback that the film received, which was supported by the Chinese government, as well as its award-winning experience and audience reactions.

Case #2 The online TV show "Addicted" in 2016

The public's acceptance of diversity in China has gradually increased in recent years, with foreign and domestic films and TV shows containing brief mentions of similar content no longer being censored. However, Chinese media still undergoes a strict and complex censorship system despite the increasing popularity of the "Boys' Love" theme (known as BL culture) among Chinese fangirls and the significant influence of LGBTQ civil rights movements worldwide. The blocking of the popular gay drama, "Addicted," directed by Ding Wei and based on a novel written by Chaijidan, reflects how the explicit mention of homosexual themes remains taboo in the Chinese censorship system. Although it was highly successful nationwide (according to The Guardian, it garnered 10 million views the day after its initial release. It was the second most-watched show on iQiyi) and caused a fanatic fan culture, the censors still banned this show completely during a movement against "vulgar, immoral and unhealthy content."[53]

[50] "Farewell My Concubine," IMDb, January 01, 1993, |PAGE|, accessed May 11, 2021, https://www.imdb.com/title/tt0106332/)

[51] "Farewell My Concubine," The Censorship Files, April 27, 2018, |PAGE|, accessed May 11, 2021, https://thecensorshipfiles.wordpress.com/farewell-my-concubine/)

[52] "Farewell My Concubine," Box Office Mojo, |PAGE|, accessed May 11, 2021, https://www.boxofficemojo.com/release/rl3612378625/weekly/)

[53] "Farewell My Concubine," The Censorship Files, April 27, 2018, |PAGE|, accessed May 11, 2021, https://thecensorshipfiles.wordpress.com/farewell-my-concubine/)

With reference to the Regulations on Administration of Films promulgated in 1996, listed in the previous section, the 15-episode-show "Addicted," featuring a simple romance story between two teenage boys, neither formed a threat to the political system nor included obscene, violent scenes. Meanwhile, except for the homosexuality-based storyline, this show had no government forbidden content since its target audience was teenagers, especially teenage girls who were fascinated by the four young, good-looking rising stars featured as main characters. Yet, the SARFT still took it down from major Chinese video streaming sites along with other similarly themed online shows at the time.[54] Similar to the classic film "Farewell My Concubine" mentioned above, "Addicted" received national attention, and viewers were highly dissatisfied after the SARFT took the show off. "Why did they take away this drama? There are millions of reasons to cover their move, but the truth is that they are afraid of gay [issues]." one Weibo user commented, quoted by the *South China Morning Post*.[55]

However, there are reasons why this specific show was removed entirely from streaming platforms even in 2016. First of all, the drug-related nickname of this show, "Heroine," which was derived from the two main characters' names, could mislead the young audience it was targeting, especially as most of its viewers were teenagers in middle or high schools. Secondly, the simple, broadly criticized soap-opera plot reflected little artistic aesthetic compared with the award-winning film mentioned before. Last but not least, many argued that the handsome young actors in the show reflected a lack of diversity or misrepresentation of the gay community. It mainly served to please the heterosexual female audience attracted to the delicate appearance and bodies of the actors. Therefore, there were few excuses for the government officials not to demonstrate their discomfort towards homosexuality by taking down this show.

Preliminary observations and conclusion

Although the degree of complexity of the Chinese censorship system, patterns, and routines can be concluded from comparing the cases above, censorship can be retroactive, meaning that shows and films taken down initially might be re-screened under public demand or international achievements, such as the "Farewell my Concubine" case illustrated above. Despite the retroactive process, these shows and films had to make changes by making cuts or by refilming certain scenes. Many censorship decisions were spontaneous and primarily based on the preference of censors and officials in the SARFT and national government. As noted above, LGBTQ themes were not mentioned before in the seven articles of the regulation on the administration of films.

From the case study, a series of reasons might lead to the situation that LGBTQ scenes are banned or cut out. For example, older audiences would not approve of the kind of content, leading to a loss of audiences and fan base.[56] Also, the idea of maintaining Chinese traditional culture, which emphasizes marriage and reproduction, remains a critical element of filial piety in China despite the only recently modified one-child policy. Meanwhile, content and artistic factors also affect the censors' decisions as there are exceptions for domestic and international award-winning shows. It is also worth mentioning that popular TV series and films such as "Call Me by Your Name" still received significant attention in China due to the fanbase of the leading actor, Timothy Chalamet, even though it was taken down a day before its release at a film festival. However, officials working at the SARFT who generally belong to the older generations feel challenged in accepting LGBTQ themes since they contradict their traditional beliefs. Even if they have neutral opinions, they consider the majority of

[54] "China's Censors Take Another Gay-Themed Web Drama Offline," The Wall Street Journal, February 24, 2016, |PAGE|, accessed May 11, 2021, https://www.wsj.com/articles/BL-CJB-28737)

[55] Charlie Campbell, "China: Censors Pull Gay Drama 'Addiction', Sparking Outcry," Time, February 25, 2016, |PAGE|, accessed May 11, 2021, https://time.com/4236864/china-gay-drama-homosexuality/)

[56] Jean K. Chalaby, "The Format Age: Television's Entertainment Revolution," Amazon, 2015, |PAGE|, accessed May 11, 2021, https://www.amazon.com/Format-Age-Televisions-Entertainment-Communication/dp/1509502599)

the audience and how they might feel. Many people in China, particularly the older generations, are homophobic due to the traditional cultures they were raised in. At the same time, younger generations are more fully educated because life conditions have improved, and they have access to international news and Western media. Therefore, it is easier for younger people to accept the diversity of people and generating widespread support for LGBTQ issues in China still needs time.

To conclude, using comparative and descriptive case studies of the different situations that two LGBTQ-themed film and TV shows faced, this research potentially will show the complexity of the Chinese censorship system and provide a possible explanation as to how the relationship between SARFT, officials, and feedback from audiences influences the censorship system. However, artistic creations shall not be the victim of the automatic censorship system. The fact is that this censorship system is yet to be fully developed, and a more sophisticated, unbiased censorship system based on a strict rating system should be applied in the future to promote a more appropriate artistic creation environment that fits national conditions, which will be beneficial for both the artists and the audiences.

References

Beijing, Jessie Jiang /. "Beijing's Gay Community Fights Censorship." Time. July 01, 2011. Accessed May 11, 2021. http://content.time.com/time/world/article/0,8599,2080443,00.html.

Campbell, Charlie. "China: Censors Pull Gay Drama 'Addiction', Sparking Outcry." Time. February 25, 2016. Accessed May 11, 2021. https://time.com/4236864/china-gay-drama-homosexuality/.

Chalaby, Jean K. "The Format Age: Television's Entertainment Revolution." Amazon. 2015. Accessed May 11, 2021. https://www.amazon.com/Format-Age-Televisions-Entertainment-Communication/dp/1509502599.

"China Bans Depictions of Gay People on Television." The Guardian. March 04, 2016. Accessed May 11, 2021. https://www.theguardian.com/tv-and-radio/2016/mar/04/china-bans-gay-people-television-clampdown-xi-jinping-censorship.

"China's Censors Take Another Gay-Themed Web Drama Offline." The Wall Street Journal. February 24, 2016. Accessed May 11, 2021. https://www.wsj.com/articles/BL-CJB-28737.

Davis, Rebecca. "China's Gay Rights Stance Can't Derail Demand for LGBT Films." Variety. June 06, 2020. Accessed May 11, 2021. https://variety.com/2020/film/asia/lgbt-movies-china-gay-rights-1234625634/.

"Douban - Crunchbase Company Profile & Funding." Crunchbase. Accessed May 11, 2021. https://www.crunchbase.com/organization/douban.

"Farewell My Concubine." Box Office Mojo. Accessed May 11, 2021. https://www.boxofficemojo.com/release/rl3612378625/weekly/.

"Farewell My Concubine." IMDb. January 01, 1993. Accessed May 11, 2021. https://www.imdb.com/title/tt0106332/.

"Farewell My Concubine." The Censorship Files. April 27, 2018. Accessed May 11, 2021. https://thecensorshipfiles.wordpress.com/farewell-my-concubine/.

Geltzer, Jeremy. "Censoring the Silk Screen: China's Precarious Balance Between State Regulation and a Global Film Market." Accessed May 11, 2021. https://www.swlaw.edu/sites/default/files/2017-04/JIMEL V6, N2 6-Censoring the Silk Screen-Geltzer.pdf.

King, Gary, Jennifer Pan, and Margaret E. Roberts. "How Censorship in China Allows Government Criticism but Silences Collective Expression." American Political Science Review. Accessed May 11, 2021. https://gking.harvard.edu/publications/how-censorship-china-allows-government-criticism-silences-collective-expression.

"LGBT Couples in China File for Voluntary Guardianship." 404. Accessed May 11, 2021. https://news.cgtn.com/news/2019-08-09/LGBT-couples-in-China-file-for-voluntary-guardianship-J15eC8QcrC/index.html#:~:text=An increasing number of LGBT,new Civil Law in 2017.&text=. The couple set up a voluntary guardianship arrangement with each other.

State Administration of Radio, Film and Television. Accessed May 11, 2021. https://in.china-embassy.org/eng/mt/jyjs/t61109.html.

Wang, Yidong, Valerie Belair-Gagnon, and Avery E. Holton. 2020. "The Technologization of News Acts in Networked News Participation: LGBT Self-Media in China." *International Journal of Communication Systems* 14 (0): 19.

What makes the Colonna Altarpiece a Great Painting?

By Vincent Luo

Author Bio:

Vincent Luo is a student currently attending St. George's School in Vancouver, Canada. During his spare time, Vincent loves reading books, visiting museums, and doodling in his sketchbook. He hopes to study art and art history in greater depth in the future.

Abstract

In his painting the Colonna Altarpiece, Raphael took various approaches, like realism, linear perspective, balance and symmetry and merged them together to create this dynamic piece. Patient examination reveals some of the most unique aspects of this piece. Is the blue in the heavenly scene above the same as the blue of the earthly scene below? Can a painting have symmetry when the colors are different on the two sides? Aside from his use of perspective, how does Raphael use light and shade to make the figures appear three dimensional? What effect is achieved when the Christ Child coincides with the vanishing point of the perspectival construction? What is the significance of the fact that baby Jesus is holding a loop-like gathering of cloth in the center of the main panel? Asking basic questions such as these makes it possible to appreciate many of the fascinating features of this painting. The various features all appear to play different roles, yet they come together to create a powerful picture.

Keywords: Raphael; *Colonna Altarpiece*; realism; linear perspective; balance; symmetry.

Madonna and Child Enthroned with Saints
(Photo credit: The Metropolitan Museum)

Raphael painted *The Colonna Altarpiece* for a Franciscan convent in Perugia circa 1504. Since this altarpiece was created for the nuns of a local convent, Raphael worked painstakingly to meet their demands. Every detail of this painting was depicted with the utmost attention, including the direction of the light source, which was drawn to match the light coming into the church from its right window. In addition, it is also important to note that many features of *The Colonna Altarpiece*, such as the lunette, were based on the works of other artists, such as Perugino. Raphael likely imitated the composition of the lunette in Perguino's *Polyptych for the church of San Pietro in Perugia*, where Jesus sat in the center and was flanked by angels on both sides.

Soon after its completion, *The Colonna Altarpiece* inspired the artworks of many local artists, including Francesco da Citta di Castello's *Mystic Marriage of Saint Catherine with Saint Agnostion* and Sinibaldo Ibi's *Madonna and Child Enthroned with Saints* (1509). Although *The Colonna Altarpiece* was widely celebrated during the Renaissance, it eventually started to languish in the 18th century. Art critics like Roger Fry went as far as to say that "no one would buy it and no one wanted to look at it." Later on, however, this painting was purchased by J. Pierepont Morgan, whose son bequeathed it to the Metropolitan Museum in the 1920s.

Today, many historians claim that the importance of *The Colonna Altarpiece* cannot be compromised. It is significant because it represents an important period in Raphael's life and is currently recognized as one of the many masterpieces of the Renaissance. This painting is an essential artifact within the Metropolitan Museum and is Raphael's only altarpiece in the United States (Bayer, 2016).

What makes *The Colonna Altarpiece* so great is its complexity and nuance. The painting is characterized by an unusual combination of tranquility and liveliness. This combination creates a sense of harmony, if harmony is understood as the interweaving of different elements into an aesthetically balanced whole. In the resulting work, minor changes to any elements of the painting will disturb its congruence. For example, if one were to change the color of the characters' robes, the two sides of this painting would no longer be balanced. The main panel of Raphael's painting features seven figures, five standing in their distinct postures, as well as a woman and child sitting on a slightly elevated throne. The initial impression is one of great symmetry-- in the main panel, a male and a female on either side of an enthroned figure, and in the lunette above, a standing figure flanked by angels on either side. The basic symmetry of the figural arrangement is carried through by the floral patterns and ornaments of the throne. On closer examination, however, the painting is full of elements that escape symmetry. Identifying the symmetrical and asymmetrical aspects of the painting is one part of a larger slow observational process. Aside from enabling the audience to notice aspects of symmetry, this observational process will ultimately unveil to the audience many impressive features that may be overlooked at a glance. It is through this analytical process that we are able to pay close attention to its realism, structural balance, color modulation, linear and atmospheric perspective, all of which contribute toward the painting's excellence.

(Sketch of Baby Jesus. Photo Credit: The Metropolitan Museum of Art)

My own sketch of Raphael's work. This has helped me gain a better understanding of how he laid out his painting at its earlier stages.

Realism is achieved when artists depict objects as accurately as possible in the form they appear. For the sake of precision, Raphael focused on his preparatory drawings before painting his panels. While painting, he concentrated on different applications of the brush stroke, and in that process created color contrast. In the publication accompanying an exhibition of the Colonna Altarpiece, Linda Wolk Simon points out that a preparatory work, *The Head of a Child*, is a study made from life for the baby Jesus in the painting. This drawing of an actual baby forms the basis for the figure of the baby Jesus in the painting. Wolk-Simon also observes that in preparing the figure of the Madonna, Raphael had an assistant pose for him while he made multiple sketches. These studies from life helped him preserve the observations that he made in the life drawings and become familiar with the shape of the figures that he planned to draw onto his main panel. After sketching these figures, his next step was to lay out the landscape and any additional elements that he wanted to incorporate into it, such as the building with a tower at the corner (Wolk Simon, 2006). Then he could apply paint to the panels. While applying paint, Raphael modeled the draperies of the figures to create contour and volume, causing light and shadow to appear on different sides (Bayer, 2006). One example of Raphael's enhancement of the painting's volume can be shown in his depiction of Saint Cecilia's mantle, which shows a marked difference in tone between its golden front side, the spot where the light is shining, and the dark color of its lower left side. Moreover, vibrant color contrast is also notable on Madonna's mantle, which is not only drawn with a particularly delicate brushstroke, but also displays a wide contrast between the light source and the shadow of the robe. To indicate the presence of the light source, Raphael included a higher concentration of yellow dots in certain areas, such as her knees (Bayer, 2013). This separates it from areas without light, where the dots are submerged in the deep blue of her robe. The yellow dots allow the viewer to see the contour and shape of her legs, which otherwise would be difficult to make out given the darkness of her robe. The dots are effective in the modeling of the shape of her legs because of their natural tendency to catch light, which in turn increases the painting's three dimensional effect by making the lighter areas protrude from the rest of her dress. It is undoubtedly a unique element that makes Madonna's robe different from the draperies of the other figures. In the end, Raphael's initial planning,

the color contrast of his figures, as well as his varying use of the paint brush, are approaches that made his painting more realistic.

(This image shows how the yellow dots that Raphael painted on Madonna's dress can enhance the level of realism. Photo Credit: The Metropolitan Museum)

Other than realism, Raphael also carried out a variety of steps to achieve structural balance. Raphael's painting is emphatically symmetrical in its elaborate throne and in the disposition of two older male saints and two younger female saints to either side of it. Raphael drew diagonal lines from the corner of the painting to locate its center, which is where he drew Christ's hand (Bayer, 2013). But the symmetry is also explored in more subtle ways. Inspired by painters such as Fra Bartolomeo, the landscape in this altarpiece offers gentle rather than absolute symmetry: while not identical, the two hills at the back of the painting resemble one another, jutting into the water in the middle (Muntz, 2016). The additional layer of symmetry allows this artwork to achieve a sense of harmony and stability. Despite the blurriness of some of its features, the symmetrical landscape helps the figures stand out further in the foreground, which draws our attention immediately to their distinct appearances and gestures. Beyond managing forms and proportions, Raphael achieves a symmetry of attitudes: one male figure looks out, the other male is in (near) profile, and the same goes for the two female saints, though here profile and outward gaze are switched; even the angels above display this contrast. Thus, a symmetry of contrasts threads through the painting (Nagel, 2020). This painting shows that subtle symmetry is actually stronger than overt symmetry.

Umbrian landscape studies. Photo Credit: The Metropolitan Museum

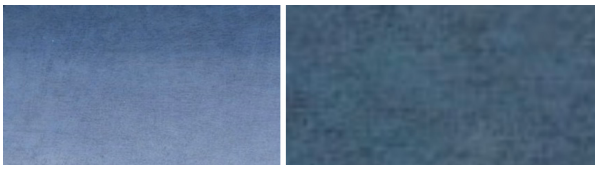
Fra Bartolomeo. Sketch of landscape. Photo credit: The Metropolitan Museum

Raphael's symmetrical landscape in the background. Photo Credit: The Metropolitan Museum

Likewise, Raphael's choice of color is also a notable aspect of this painting. Raphael used a technique called equalization, essentially the symmetry of color, to make the figures in his painting balanced in tone (Nagel, 2020). The mantles of Paul and Peter are different in color, yet this obvious difference is compensated by their similarity in tone. Though red is darker and more vibrant than yellow, Raphael made St. Paul's red mantle slightly closer to the light source in order to lighten its tone and thus match St. Peter's yellow mantle through equalization. By bringing their mantles into a proximate tonal range, this equalization technique gives the audience the impression that both Saints are equal in status with one another, and indeed

as the two principal Apostles they were. Raphael draws the audience's attention towards Madonna, whose red tunic reflects light more sharply than St. Paul's. Perhaps it is made out of a shinier, more reflective fabric. In respecting these differences in the reflective properties of different fabrics, Raphael is a realist. Yet, he uses these devices in order to shine a spotlight on the Virgin, the core of the painting, the miraculous vehicle through whom the divine was made flesh. This effect is also enhanced by contrasting the tone of her tunic with her dark mantle. Because Madonna is the central figure of this painting, it appears as if Raphael has established a symbolic hierarchy based on the color of their robes. (Nagel, 2020).

Additionally, Raphael used similar colors to connect related characters. For instance, he used red and purple to paint Christ and Madonna's clothes to match God's robe in the lunette- a red thread that encapsulates the theology of the incarnation: the divine becomes flesh through the body of a virgin (Nagel, 2020). All the figures are wearing a combination of purple, red, yellow, and green, which implies that they are connected to one another. Since God in the lunette is dressed in a similar fashion to the Saints, it is likely that Raphael is trying to demonstrate God's connection to humans (Nagel, 2020). Nevertheless, it is important to take into account that God has been angry with humans several times, as indicated by the Old Testament: Adam and Eve's Expulsion from heaven, The Great Flood, The plague of Egypt. To take it one step further, it is likely that Raphael's depiction of similar colored robes suggests people's desire for reconciliation. Because people see God as the almighty being, it makes sense why they want to be on his good side. Aside from the color of their robes, there is also a modulation of light in the background; Raphael added a lighter hue to the blue-sky layer by layer as he worked his way down the central panel. By contrast, Raphael painted the sky of heaven with one unmodulated blue to highlight its purity, whereas the sky at the bottom is painted with more depth and feels more spacious, which gives it an earthly atmosphere. This establishes the difference between the sphere of the living and the sphere of the immortals in heaven (Nagel, 2020), which in effect makes this painting particularly interesting to examine in depth.



Left: The blue of the sky in the central panel. Right: The blue of heaven in the lunette. Photo Credit: The Metropolitan Museum

On top of that, Raphael effectively uses technical elements such as linear perspective to address details of a larger story. In the painting, it is notable that the perspectival lines of the throne extend diagonally into the center of the painting, giving the impression that they converge at Jesus's left hand. This directs our attention to baby Jesus's unique posture, where his right hand seems to be in a blessing posture, while his left characterizes his childlike innocence (Wolk-Simon, 2006). Like all the techniques mentioned above, this serves to depict the duality of baby Jesus's character as both a human and a living God. In addition, it is also important to note that both Jesus and God the Father in the lunette above are holding circular objects. Jesus's left hand seems to be holding a loop, which parallels God the Father holding a sphere above. Since a circle, having no beginning nor end, suggests the concept of infinity, we can then infer that Raphael is trying to highlight their immortality. Even though scrutinizing technical aspects like linear perspective might seem minor to our understanding of the painting, it can sometimes reveal contextual information that is crucial to the implications behind the altarpiece.

The discussion of Raphael's Colonna Altarpiece is highly logical as it comes in a sequence. This essay began by explaining the steps that he took to create this painting, followed by a discussion of various features that made this painting remarkable. It included a discussion of features that enhance realism, different ways of achieving structural harmony, incorporation of linear perspectives, as well as approaches that provide contextual elements about Christianity. However, the Colonna Altarpiece as a whole does not have a sequence in which it is being looked at. All the features are where they are supposed to be and they all serve unique purposes. Nevertheless, none of these features, such as the equalization of Paul and Peter's robes, the variation in the different shades of the sky, or the Christ Child holding a loop would make sense on their own. Raphael deliberately made St. Paul's red robe the same tone as St. Peter's

yellow robe to indicate their equality in status, which is impossible to notice without observing carefully both of the Saints; Likewise, it would be difficult to understand the implication behind the changing shades of the blue sky if merely the central panel is being looked at; it would also be puzzling to assess why the Christ child is holding a loop without taking God the Father into consideration. Given these examples, it is reasonable to conclude that many features of the painting work together with one another to form a powerful picture. Along with Raphael's brilliant technique, the cohesion of these subtle elements is what makes the Colonna Altarpiece brilliant as a whole.

Christ holding a loop in front of Madonna's womb. Photo Credit: The Metropolitan Museum

References

Bayer, A., & The Metropolitan Museum of Art. (2013). Madonna and Child Enthroned with Saints ca. 1504. The Metropolitan Museum of Art. Retrieved August 9, 2020, from https://www.metmuseum.org/art/collection/

Muntz, E. (2016). Rapheal. Parkstone Press International.

Nagel, A. (Presenter). (2020, August 7). The Colonna Altarpiece. Speech presented in New York City, NY, United States.

Wolk-Simon, L., & Metropolitan Museum of Art (New York, N.Y.). (2006). Raphael at the Metropolitan: The Colonna Altarpiece. New York: Metropolitan Museum of Art.

contact@scholarlyreview.org