

AJUR

American Journal of
Undergraduate Research

Volume 16 | Issue 4 | March 2020

www.ajuronline.org

Print Edition ISSN 1536-4585
Online Edition ISSN 2375-8732

AJUR

American Journal of
Undergraduate Research

Volume 16 | Issue 4 | March 2020

www.ajuronline.org

- 2 **AJUR History and Editorial Board**
- 3 **Special Thanks to AJUR's Sponsors**
- 5 **Coating Polyurethane with Palmitoleic Acid and Bovine Serum Albumin to Prevent the Host Response to Foreign Materials**
Sheherbano Hussain, Zoha Babar, Jimmy Hadid, & Jacqueline McLaughlin
- 15 **Music Intervention in Undergraduates: the Relationship between Heart Rate Variability and State Anxiety**
Annalise J. Tolley & Robert S. Vick
- 23 **Can Expense Ratios Signal Performance? An Analysis of Equity ETFs & Mutual Funds**
James Monroe Gamble IV
- 41 **Acoustic Identification of Wild Gray Wolves, *Canis lupus*, Using Low Quality Recordings**
Cara B Hull, Caitlin M McCombe, & Angela M Dassow
- 51 **The Allele Frequency of the HFE gene mutation H63D (rs1799945) and Its Relationship to a Hereditary Hemochromatosis Diagnosis in Metabolic Nutrition Students at Virginia Tech**
Tyler R. Ferqueron, Angela S. Anderson, & Deborah J. Good
- 59 **Meningococcal Meningitis in College Students at United States Universities**
Mikafui Dzotsi
- 71 **Urban Wind Harvesting Using Flow-Induced Vibrations**
Levon Ghabuzyan, Christopher Luengas, & Jim Kuo
- 81 **Exploring the Relationship between Dystopian Literature and the Activism of Generation Z Young Adults**
Aysba Jerald
-

American Journal of Undergraduate Research (AJUR) is a national, independent, peer-reviewed, open-source, quarterly, multidisciplinary student research journal. Each manuscript of AJUR receives a DOI number. AJUR is archived by the US Library of Congress. AJUR was established in 2002, incorporated as an NFP in 2018. AJUR is indexed internationally by EBSCO and Crossref with ISSNs of 1536-4585 (print) and 2375-8732 (web).

EDITORIAL TEAM Volume 16 | Issue 4 | March 2020

Dr. Kestutis G. Bendinskas, Executive Editor, editor@ajuronline.org

Dr. Anthony Contento, Copy Editor, Treasurer

Peter Newell, Editor, Secretary

Daniel Laird, Web Master

EDITORIAL BOARD by subject area

ACCOUNTING

Dr. Dean Crawford,
dean.crawford@oswego.edu

ARCHEOLOGY

Dr. Richard Redding,
rredding@umich.edu

ART HISTORY

Dr. Lisa Seppi,
lisa.seppi@oswego.edu

ASTROPHYSICS

Dr. Shashi Kanbur,
shashi.kanbur@oswego.edu

BEHAVIORAL NEUROSCIENCE

Dr. Aileen M. Bailey,
ambailey@smcm.edu

BIOCHEMISTRY

Dr. Pamela K. Kerrigan,
pamela.kerrigan@montsaintvincent.edu

Dr. Nin Dingra,
ndingra@alaska.edu

BIOENGINEERING

Dr. Jorge I. Rodriguez,
jorger@clmson.edu

Dr. Jessica Amber Jennings,
jjennings@memphis.edu

BIOINFORMATICS

Dr. Kevin Daimi,
daimikj@udmercy.edu

Dr. John R. Jungck,
jungck@udel.edu

Dr. Isabelle Bichindaritz,
ibichind@oswego.edu

BIOLOGY, PHYSIOLOGY

Dr. David Dunn,
david.dunn@oswego.edu

BIOLOGY, DEVELOPMENTAL

Dr. Poongodi Geetha-Loganathan,
p.geethaloganathan@oswego.edu

BIOLOGY, MICROBIOLOGY

Dr. Peter Newell,
peter.newell@oswego.edu

BOTANY

Dr. William R. Bromer,
wbromer@stfrancis.edu

Dr. Julien Bachelier,
julien.bachelier@fu-berlin.de

CHEMISTRY

Dr. Alfredo Castro,
castroa@felician.edu

Dr. Charles Kriley,
ckriley@gcc.edu

Dr. Douglas Mulford,
douglas.mulford@emory.edu

Dr. Vadoud Niri,
vadoud.niri@oswego.edu

COMMUNICATION DISORDERS AND SCIENCES

Dr. Kim Tillery,
Kim.Tillery@fredonia.edu

COMPUTER SCIENCES

Dr. Dele Oluwade,
deleoluwade@yahoo.com

Dr. Kevin Daimi,
daimikj@udmercy.edu

Dr. Levent Ertaul,
levent.ertaul@csueastbay.edu

Dr. Mais W Nijim,
Mais.Nijim@tamuk.edu

COMPUTATIONAL CHEMISTRY

Dr. Alexander Soudackov,
alexander.soudackov@yale.edu

ECOLOGY

Dr. William R. Bromer,
wbromer@stfrancis.edu

ECONOMICS

Dr. Elizabeth Schmitt,
elizabeth.schmitt@oswego.edu

EDUCATION

Dr. Marcia Burrell,
marcia.burrell@oswego.edu

EDUCATION, PHYSICS

Dr. Andrew D. Gavrin,
agavrin@inpui.edu

ENGINEERING, ELECTRICAL

Dr. Michael Omidiora,
michael.omidiora@nyu.edu

ENGINEERING, ENVIRONMENTAL

Dr. Eileen M. Cashman,
eileen.cashman@humboldt.edu

ENGINEERING, SOFTWARE

Dr. Kevin Daimi,
daimikj@udmercy.edu

ENVIRONMENTAL SCIENCES

Dr. Eileen M. Cashman,
eileen.cashman@humboldt.edu

FILM AND MEDIA STUDIES

Dr. Lauren Steimer,
lsteimer@mailbox.sc.edu

GEOLOGY

Dr. Rachel Lee,
rachel.lee@oswego.edu

HISTORY

Dr. Richard Weyhing,
richard.weyhing@oswego.edu

Dr. Murat Yasar,
murat.yasar@oswego.edu

HONORARY EDITORIAL BOARD MEMBER

Dr. Lorrie Clemo,
lorrie.a.clemo@gmail.com

JURISPRUDENCE

Bill Wickard, Esq.,
William.Wickard@KL.Gates.com

KINESIOLOGY

Dr. David Senchina,
david.senchina@drake.edu

LITERARY STUDIES

Dr. Melissa Ames,
mames@ein.edu

Dr. Douglas Guerra,
douglas.guerra@oswego.edu

MATHEMATICS

Dr. John Emert,
emert@bsu.edu

Dr. Jeffrey J. Boats,
boatsjjf@udmercy.edu

Dr. J.D. Phillips,
jophilli@nmu.edu

Dr. Dele Oluwade,
deleoluwade@yahoo.com

Dr. Christopher Baltus,
christopher.baltus@oswego.edu

Dr. Mark Baker,
mark.baker@oswego.edu

MEDICAL SCIENCES

Dr. Thomas Mahl,
Thomas.Mahl@va.gov

Dr. Jessica Amber Jennings,
jjennings@memphis.edu

METEOROLOGY

Dr. Steven Skubis,
steven.skubis@oswego.edu

MUSIC

Dr. Juliet Forshaw,
juliet.forshaw@oswego.edu

NANOSCIENCE AND CHEMISTRY

Dr. Gary Baker,
bakergar@missouri.edu

NEUROSCIENCE

Dr. Pamela E. Scott-Johnson,
pscottj@calstatela.edu

PHYSICS

Dr. Mohammad Islam,
mohammad.islam@oswego.edu

POLITICAL SCIENCE

Dr. Katia Levintova,
levintoe@uwgb.edu

PSYCHOLOGY

Dr. Joseph DW Stephens,
jdstephe@ncat.edu

Dr. Melanie Dyan Hetzel-Riggins,
mdb33@psu.edu

Dr. Pamela E. Scott-Johnson,
pscottj@calstatela.edu

SOCIAL SCIENCES

Dr. Rena Zito,
rzito@elon.edu

STATISTICS

Dr. Mark Ecker,
mark.ecker@uni.edu

Dr. Mark Baker,
mark.baker@oswego.edu

TECHNOLOGY, ENGINEERING

Dr. Reg Pecan,
regpecan@sbsu.edu

ZOOLOGY

Dr. Maria Sagot,
maria.sagot@oswego.edu

SPECIAL THANKS

AJUR is made possible through the volunteer efforts of our editorial team.

We sincerely thank our financial sponsors:



WILEY

Please support undergraduate research
Request sponsorship information from editor@ajuronline.org

Coating Polyurethane with Palmitoleic Acid and Bovine Serum Albumin to Prevent the Host Response to Foreign Materials

Sheherbano Hussain, Zoha Babar, Jimmy Hadid, & Jacqueline McLaughlin*

Department of Biology, Penn State Lehigh Valley, Center Valley, PA

<https://doi.org/10.33697/ajur.2020.002>

Students: skb5406@psu.edu*, zpb5044@psu.edu, jjb5916@psu.edu,

Mentor: jxm57@psu.edu

ABSTRACT

Macrophages are cells of the immune system that play a pivotal role in the host inflammatory response by attacking and engulfing any foreign molecule not seen as 'self.' They also help regulate the host response by releasing a variety of cytokines and growth factors that act as signals to other cells to amplify the host response. However, the host response causes degradation of implanted medical devices composed of polyurethane as well as other synthetic materials which it does not identify as self. Research was undertaken to investigate the potential of coating polyurethane with the self-like molecules palmitoleic acid and albumin to reduce or prevent the body's host response from damaging implanted medical devices. Using an *in vitro* THP-1 bioassay, polyurethane films coated with palmitoleic acid and bovine serum albumin showed a reduction in macrophage adherence. The individually coated palmitoleic acid and bovine serum albumin films significantly reduced the number of cells attached to the films with increasing concentration while the films coated with the conjugate of both showed no statistical difference. This suggests the potential role of self-like molecules in reducing the inflammatory response to foreign materials.

KEYWORDS

Macrophages; host inflammatory response; THP-1; prosthetics; palmitoleic acid; bovine serum albumin; cytokines; immune system

INTRODUCTION

Macrophages are large white blood cells that engulf and digest cellular debris, foreign substances, microbes, cancer cells, and anything else seen as foreign. Thus, they play a key role in the human body's 'host response,' which is the defense mechanism and reaction against the presence of foreign molecules in the body. It constitutes a variety of mechanisms that contribute to how the body responds to anything that is foreign or non-self. Macrophages are formed from monocytes, which are white blood cells that differentiate from stem cells produced in the bone marrow and reside in the extracellular matrices between cells in the body's tissues. To differentiate between foreign molecules and self, a number of phagocytic membrane receptors have evolved on the surface of macrophages that aid the immune system in the ability to recognize and respond to foreign materials and tolerate self-components of the body. These receptors can be identified using flow cytometry or immunohistochemical staining.¹

Recognition of a foreign molecule by a macrophage through receptor binding involves phagocytosis. A pocket called a phagosome is formed around the unwanted material. Enzymes digest and degrade the unknown materials.² In addition, macrophages also work to help regulate the inflammatory response by producing and releasing a variety of cytokines and growth factors. For activation, cytokines and growth factors send signals to other cells to amplify the host response. On the other hand, anti-inflammatory cytokines like interleukin-10 can mediate the host response.³

One of the most important functions of macrophages is that they can activate the adaptive immune system. After a macrophage has phagocytized and digested foreign material, it displays some of the broken-down proteins (antigens) on its cell surface. These antigens act as identification signals for Helper T cells. Helper T cells can "read" these signals and tell what kind of particle the macrophage has ingested. If the T cell determines the macrophage has eaten something harmful (a pathogen), it can trigger a powerful reaction towards the specific pathogen.⁴

THP-1 cells are a human monocyte-derived macrophage cell line derived from the peripheral blood of a one-year old male infant with acute monocytic leukemia.⁵ These cells are considered an ideal cell line for growing suspension cell cultures and monolayers on experimental surfaces. THP-1 cells have been used to study the host response to biomaterials and biocompatibility because they are an excellent model of monocytes to macrophage differentiation.^{6,7} This knowledge is essential in medicine, bioengineering, and biomaterial science, since the host response detects foreign materials through activated macrophage attack via

phagocytosis, release of proinflammatory cytokines, and neutrophil activation – ultimately degrading the material.⁸ Device failures due to the host response can lead to harmful health effects. Moreover, the pathophysiology in accordance with this biological process following medical device implementation exhibits a considerable healthcare burden.⁹

Specific fatty acids have been shown to affect control processes in inflammatory responses and have been researched in this respect. Due to their hydrophobic nature, fatty acids are incorporated into cell membranes. Research shows that they can bind to membrane receptors or change intracellular protein activation, thereby altering cellular function.¹⁰

Palmitoleic acid (PA) is an n-7 monounsaturated fatty acid secreted by adipose tissue. Studies have shown that monounsaturated fatty acids, like palmitoleic acid, modulate the secretion of important transcription factors involved with inflammatory pathways by inhibiting NF- κ B, which is essential for cytokine production.¹¹ This occurred specifically by a decrease in the Th1, Th17, and CD28 responses, and an increase in the inhibitory receptor CD95.

In cultured macrophages, evidence has also been shown that PA decreases the level of proinflammatory cytokine expression. Macrophages were isolated from the intraperitoneal cavity of mice and were exposed to PA conjugated with albumin for 24 hours in culture. RNA was isolated and gene expression was quantified to determine the relation to metabolic and inflammatory pathways. PA was observed to stimulate anti-inflammatory effects by inhibition of the inflammasome pathway, which constitutes the complex of proteins associated with activating the inflammatory response. PA decreased the expression of inflammatory markers, specifically NF κ B and IL-1 β .¹²

In the following experiment, to further research potential deterrents of the human host response to implanted biomaterials, PA was investigated for its ability to decrease the number of adhered THP-1 cells to polyurethane *in vitro*, a commonly used polymer in implantable devices. It was reasoned that there would be a significant decrease in THP-1 cell adherence to polyurethane films coated with PA through the THP-1 Attachment Assay.¹³ Since using PA in a cell-based assay may prove to be challenging due to its hydrophobic nature, PA conjugated with albumin was also investigated to assist in the coating process. Human albumin is a blood protein that can serve as an adhesion location for macrophages and other cells of the immune response once an implant has interacted with the tissue. However, since bovine serum albumin (BSA) is commonly used in THP-1 cell culture protocols, it is a molecule both recognizable by THP-1 cells and unable to increase the host response.¹⁴ Due to its negative charge, it can bind water, salts, fatty acids, vitamins and hormones and carries these bound components between tissues and cells.¹⁵ With fatty acids in particular, BSA creates an aqueous-soluble reagent that can be absorbed and employed by cells as exemplified through its known conjugation with palmitate.¹⁶ With this composite knowledge, the following hypothesis was developed and tested: there will be a significant decrease in activated THP-1 cell (macrophage) adherence to polyurethane films coated with the self-like molecules PA, PA conjugated to BSA, and BSA through the THP-1 Attachment Assay.¹³

METHODS AND PROCEDURES

A. Experimental Design for Cell Culture

The experimental design for this experiment is depicted in the schematic 24-well plate presented in Figure 1. Polyurethane control films coated with sodium chloride only, and experimental films coated with different concentrations of PA, PA conjugated with BSA, and BSA were prepared and utilized. For the first set of wells, the concentration of PA and BSA was 100 μ M. As for the second set, the concentration of PA and BSA was 500 μ M. Lastly, the third set of wells had films coated with 1000 μ M of PA and BSA. The ratio for PA conjugated with BSA in each experimental concentration equaled 1 mM of PA to 0.17 mM of BSA (see below section, Coating of Polyurethane Films).

B. Coating of Polyurethane Films

Sheets of polyurethane films were obtained from Children's Hospital of Philadelphia (CHOP, PA), and cut into 1 cm x 0.5 cm rectangles, then placed into the wells of a 24-well culture plate - all in a vented hood using sterile scissors and forceps. A final concentration of 100 μ M, 500 μ M, and 1000 μ M palmitoleic acid was diluted into 95% ethanol and dispensed onto the polyurethane films in their respective wells (**Figure 1**).

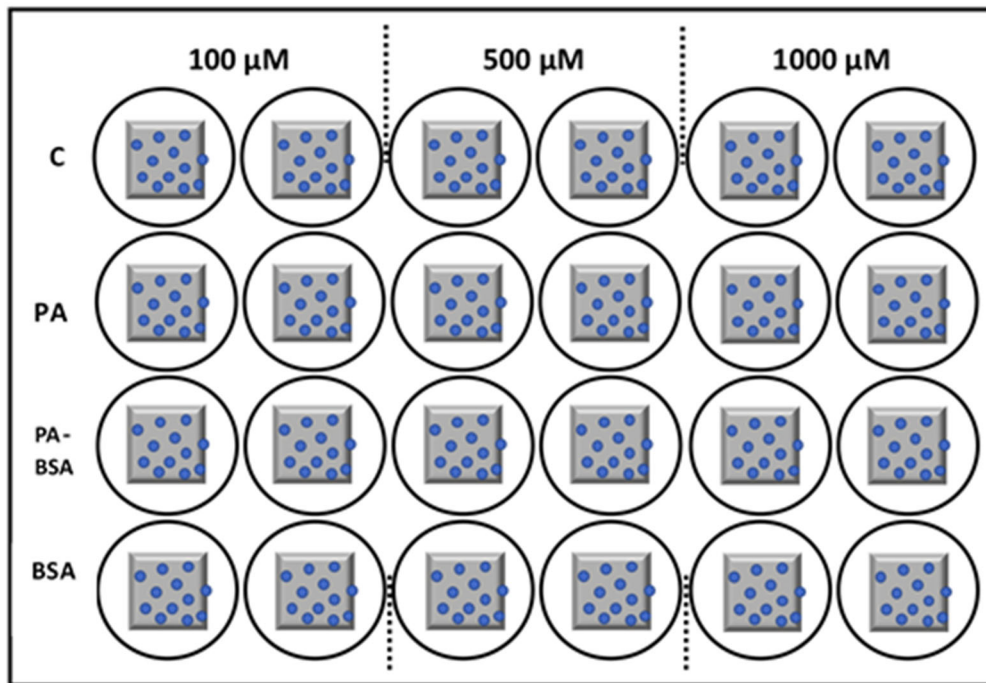


Figure 1. The experimental design consisted of polyurethane films in a 24-well dish. There were two wells per specific treatment and concentration. For each concentration designation (100 μ M, 500 μ M, and 1000 μ M), there were control samples and films coated with palmitoleic acid (PA), palmitoleic acid conjugated with bovine serum albumin (PA-BSA), and bovine serum albumin (BSA), respectively. Ten samples were analyzed for each polyurethane film, resulting in twenty samples per specific concentration and treatment condition.

The control films (C) were coated with NaCl only. All polyurethane films were dried at room temperature overnight to ensure adhesion. For each of the respective BSA wells, the same concentrations of 100 μ M, 500 μ M, and 1000 μ M were added into NaCl instead of ethanol. A ratio of 1 mM of PA to 0.17 mM of BSA was utilized for the wells reserved for the PA-BSA conjugate.¹⁶ For the six PA-BSA conjugate films, the first two films were coated with 100 μ M of PA conjugated with BSA. The second pair was coated with 500 μ M of PA conjugated with BSA, and the third pair was coated with 1000 μ M of PA conjugated with BSA. The conjugate was prepared by stirring the PA and BSA solutions under heat.¹⁶ All polyurethane films were coated with 50 μ L of designated solution, then dried at room temperature overnight (in a sterile hood) to ensure adhesion. Any excess coating mixture was removed by pipetting.

These specific PA concentrations were chosen to determine the effects of a low, medium, and high concentration of PA on macrophage attachment to the polyurethane films. Previous research¹² utilized the concentration of 600 μ M of PA in cell culture; thus, concentrations around this value were chosen to see the effects of activated THP-1 cell (macrophage) adhesion on the polyurethane films. NaCl was used to ensure macrophage adherence to the control polyurethane films.

C. The THP-1 Attachment Assay

The THP-1 Attachment Assay¹³ was utilized in this experiment. All work for the assay was conducted in sterile conditions under a hood. 70% ethanol was used to disinfect any cell culture materials. THP-1 cells purchased from ATCC (Manassas, VA) were utilized. Cells were grown in RPMI-1640 media (Thermo, Denver, Colorado) with 5% FBS (Gibco, Grand Island, New York) and +50 μ M 2-ME (2-Mercaptoethanol; Gibco, Grand Island, New York) at 37 $^{\circ}$ C and 5% CO₂ in a Thermo Forma incubator. A minimum of 3 million active growing THP-1 cells were collected and centrifuged at 3,000 rpm. They were counted with the aid of trypan blue and an automated cell counter (Olympus). The cells were deemed acceptable with percent viabilities that were calculated to be above 90%. The cell concentration and the volume of THP-1 cells required to seed 300,000 cells per well were determined. Essentially, the THP-1 cells were activated with Phorbol 12-myristate 13-acetate (PMA) before cell addition into the wells. PMA assists in the differentiation process from monocytes to macrophage by activating protein kinases. Adhesion of differentiated macrophages onto polyurethane films was allowed by a three-day incubation at 37 $^{\circ}$ C and 5% CO₂. Following this allotted incubation time, the polyurethane films were washed with phosphate buffered saline (PBS) three times, and then the

attached cells were fixed with 1% formaldehyde fixation buffer. After being washed by PBS for three times again, they were stained with Vectashield plus DAPI (Vector Laboratories, Inc) to label the cell nuclei that had attached to the film. The films were analyzed by fluorescence microscopy.

D. Microscopy and Cell Counting

Light micrographs were taken with an inverted phase contrast microscope (Olympus, BX41 and Camera DP73) to view the THP-1 cells after three days of incubation. To further determine the number of adhered macrophages, fluorescence micrographs with the DAPI stain were taken at 100x magnification and analyzed. Ten fields of view were randomly selected and an area between 151,000 and 151,500 μm^2 was utilized for cell counting for each control and experimental film. The selected fields and associated areas were not repeated or duplicated. Twenty sampling areas were analyzed per experimental condition. The adhered number of macrophages were counted by the number of nuclei present and totaled for each field. CellSens software was used to demarcate the specific areas for counting in each field and to assist in the cell counting process. The software allowed markers to be placed on each cell nucleus in each micrograph and total numbers were recorded. Cell count values for each set of two control and experimental films were averaged and compared to determine which group of films experienced the most macrophage adhesion. P values which compared each experimental group to its designated control group were obtained using a GraphPad¹⁷ unpaired t-test with a significance value of < 0.01 . Analysis of the p-values influenced conclusions regarding the significance of results.

RESULTS

As displayed in **Figures 2-4**, there was an apparent decrease in macrophage (THP-1 activated cells) attachment to the PA and BSA coated polyurethane films for all concentrations tested (100 μM , 500 μM , and 1000 μM). The PA-BSA coated films presented more adhered macrophages, similar to the control films. Ten samples per well (twenty per experimental condition) were analyzed for macrophage adherence per random sampling area (151,000 - 151,500 μm^2).

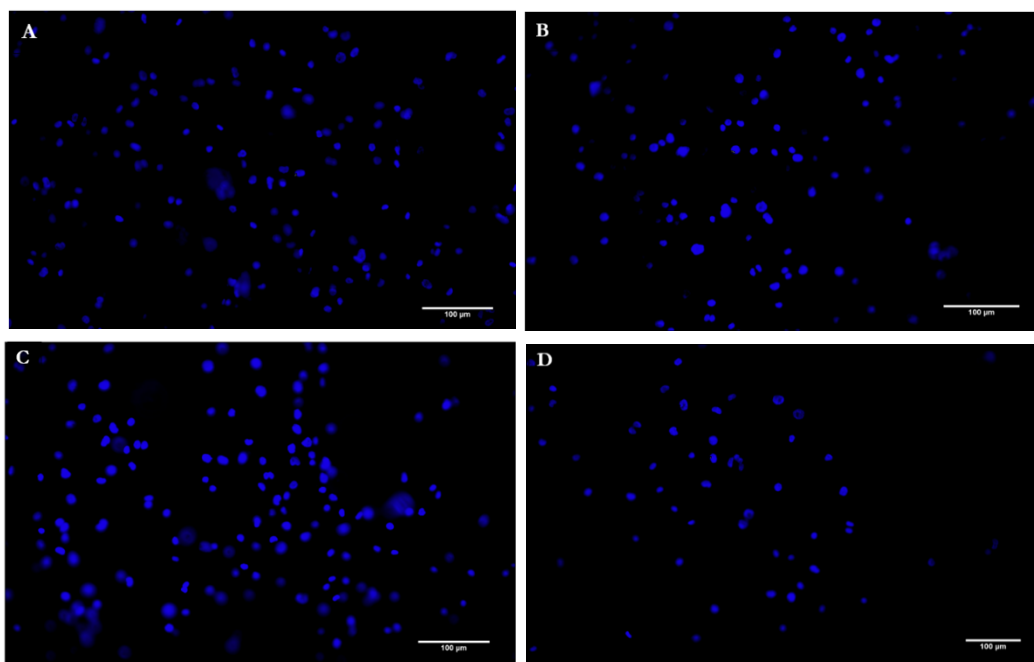


Figure 2. Fluorescence micrographs taken at a magnification of 100X of adhered macrophages (THP-1 activated cells) to polyurethane films coated with 100 μM : NaCl control (A), PA (B), conjugated PA-BSA (C), BSA (D).

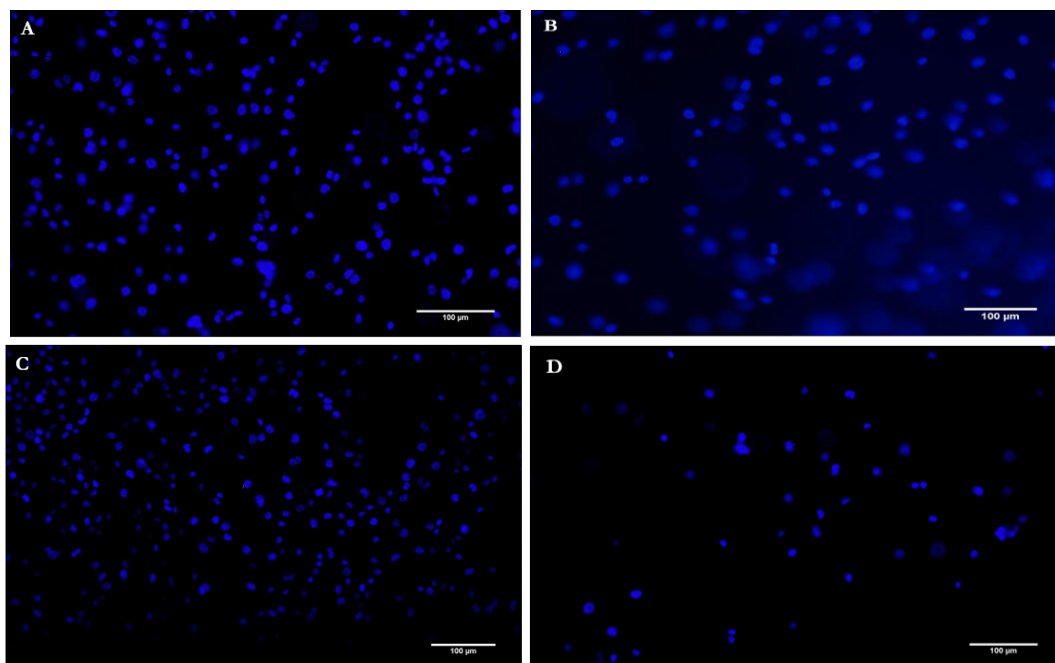


Figure 3. Fluorescence micrographs taken at a magnification of 100X of adhered macrophages (THP-1 activated cells) to films coated with 500 μ M: NaCl control (A), PA (B), conjugated PA-BSA (C), BSA (D).

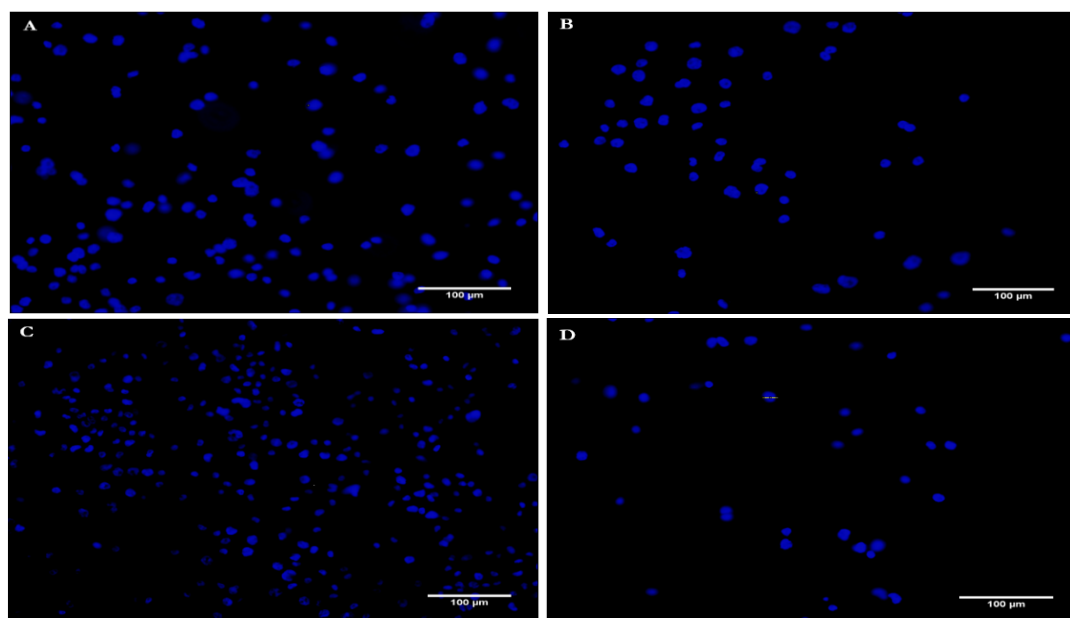


Figure 4. Fluorescence micrographs taken at a magnification of 100X of adhered macrophages (THP-1 activated cells) to polyurethane films coated with 1,000 μ M: NaCl control (A), PA (B), conjugated PA-BSA (C), BSA (D).

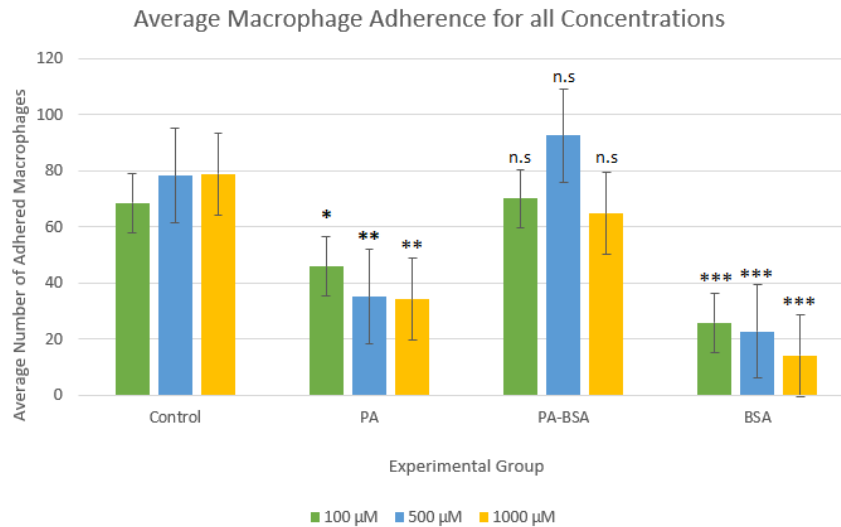


Figure 5. The average number of adhered macrophages per random sampling area (151,000 - 151,500 μm^2) in each experimental group showing all concentrations tested: control, PA, PA-BSA, and BSA. Twenty samples were analyzed per experimental condition of concentration and treatment. Error bars are the standard deviations of the mean. P-values were calculated for significance: * $p < 0.01$, ** $p < 0.001$, *** $p < 0.0001$.

The average number of adhered macrophages to polyurethane films with NaCl, 100 μM of PA, PA-BSA conjugate, and BSA are displayed in green bars in **Figure 5**. There was an average of 68.5 macrophages adhered to the polyurethane films with no coating; an average of 45.7 macrophages adhered to the polyurethane coated with PA; an average of 70 macrophages adhered to the polyurethane coated with the PA-BSA conjugate; and, an average of 25.7 macrophages adhered to the polyurethane coated with BSA. P-values were calculated for the PA, PA-BSA conjugate, and BSA groups using a GraphPad¹⁷ unpaired t-test with a significance value of less than 0.01. Based on the results, polyurethane films coated with 100 μM of PA and BSA exhibited a significant decrease in adhered macrophages.

The average number of adhered macrophages to polyurethane films with NaCl, 500 μM of PA, BSA, and PA-BSA conjugate is presented by the blue bars in the figure. The average number of adhered macrophages per random sampling area (151,000 - 151,500 μm^2) for each experimental group was 78.25, 35.05, 92.45, and 22.75, respectively. There was a significant decrease in the number of adhered macrophages on the polyurethane films coated with 500 μM of PA and BSA individually.

The yellow bars in **Figure 5** display the average number of adhered macrophages to polyurethane films with NaCl, 1,000 μM of PA, the PA-BSA conjugate, and BSA. The average number of adhered macrophages per random sampling area (151,000 - 151,500 μm^2) on each experimental group was 78.75, 34.15, 64.65, and 14.05, respectively. When 1000 μM of PA and BSA was applied to the polyurethane films, the number of adhered macrophages decreased significantly when compared to the results seen using the 100 μM and 500 μM concentrations.

From these data, it can be concluded that the polyurethane films coated with BSA had the most significant decrease in the number of adhered macrophages compared to that of the control, PA, and PA-BSA experimental groups. PA also caused a significant decrease in macrophage adherence, as apparent by the significant p-values. Both PA and BSA exhibited a reduced number of adhered macrophages to the polyurethane films as compared to the control for all concentrations of 100 μM , 500 μM , and 1000 μM . The average number of macrophages attached to PA coated films were 45.7, 35.05, and 34.15, respectively. The average number of macrophages attached to BSA coated films were 25.7, 22.75, and 14.05. The p-values were less than the significance value of 0.01. On the other hand, the polyurethane films coated with the PA-BSA conjugate presented no significant differences from saline control surfaces as the p-values were much greater than the significance value of 0.01.

DISCUSSION

Studies have shown the damaging effects of the body's host inflammatory response instigated by the presence of implanted medical devices. The presence of implants and the degradation of such implants attract cells of the immune system, prominently differentiated macrophages, that recruit other cells to the affected area by the release of a variety of cytokines and chemokines.¹⁸ This research was undertaken to investigate how the human host response can be reduced by coating such devices with molecules that are not seen as foreign in the body (self-like). The specific hypothesis under investigation was PA, PA conjugated to BSA, and BSA will decrease the number of adhered activated THP-1 cells (macrophages) to polyurethane films *in vitro*. Polyurethane films were chosen since polyurethane is a commonly used polymer in implantable devices and biomaterials. Polyurethanes are produced by reactions between alcohols and isocyanates, the latter of which is considered a 'non-self' molecule in the body. PA, a hydrophobic monounsaturated fatty acid commonly found in adipose tissue, was reasoned because of previous research indicating the anti-inflammatory nature of fatty acids by inhibiting macrophage cytokine production¹¹; and, a PA-BSA conjugate was chosen because of BSA's ability to bind hydrophobic fatty acids and efficiency as an aqueous-soluble reagent in cell-based assays.¹⁴ Albumin is a negatively charged, hydrophilic large molecule found in blood plasma, and BSA is a common supplement used in THP-1 suspension cell culture systems. It was assumed that coating polyurethane films with the high concentration of PA, the PA-BSA conjugate, or BSA would exhibit the least macrophage attachment while the control films would exhibit the most macrophage attachment. On the other hand, coating the polyurethane films with the low concentrations of PA, PA conjugated with BSA, or BSA would exhibit intermediate macrophage attachment. The experiment led to contradictory results. Although the hypothesis was supported by the PA and BSA adherence values, the adherence values on films coated with the PA-BSA conjugate across all concentrations were higher than expected.

Data across polyurethane films of all three concentrations supported the hypothesis that coating with PA reduced macrophage adherence *in vitro* as compared to the control films with no coating. However, the BSA films exhibited a significant reduction in macrophage adhesion as well. This indicates that both PA and BSA possess the capability to reduce macrophage attachment. In addition, there was an apparent decrease in macrophage attachment for both substrates as the coating concentration increased. Therefore, the films coated with the higher concentration (1,000 μM) of PA and BSA displayed the lowest adherence values. In addition, data gathered from films of all three concentrations showed that BSA had the least number of adhered macrophages, followed by PA.

Although the results of the PA and BSA individually coated films supported the hypothesis in reducing macrophage attachment to polyurethane films, the results of the PA-BSA conjugate did not. There was no apparent decrease of macrophage adherence on the PA-BSA coated films for all three concentrations. In fact, the data obtained from the cell counts showed that the number of attached macrophages to the PA-BSA films were not statistically different than that of the control films with no coating. In summary, due to these results, the hypothesis under investigation in this study was only partially supported as the PA-BSA conjugate had no repressive effect on macrophage attachment, while the individually coated PA and BSA films did.

This discrepancy regarding the PA-BSA conjugate can be related to the ability of macrophages to recognize "self-like" molecules from "non-self-molecules". It is assumed that by forming the conjugate through constant heating and stirring, a conformational change(s) took place in the PA and BSA conjugate as the fatty acids covalently bonded to the protein.¹⁹ Thus, the activated macrophages proceeded as they are charged to do; they treated the conjugate as a foreign molecule by attaching to it and most likely releasing cytokines and transcription factors that amplified the response. On the other hand, since PA is commonly found in the human body, and BSA in tissue culture for THP-1 cells, the macrophages recognized them as such. This is supported by the increase in number of adhered macrophages to the polyurethane films coated with the conjugate as compared to the reduced number of adhered macrophages on the PA and BSA coated films. With the rise of temperature at pH 7.0, the proportion of alpha-helix decreased above 30 °C and those of beta-structure and disordered structure increased in the same temperature range.¹⁹

CONCLUSIONS

In conclusion, we found that PA and BSA coated polyurethane films reduce macrophage adherence *in vitro*, while films coated with conjugated PA-BSA had no significant effect. We believe that this research can aid other researchers in finding a viable solution to help slow down the degradation of implanted medical devices composed of polyurethane via activation of the host

response. By exploring the ability of 'self-like' molecules like palmitoleic acid to significantly reduce macrophage adherence, revised manufacturing measures may be investigated to reduce the host response.

For future research, specific cytokine expression released by macrophages in our assays will be evaluated using ELISA and gene expression. We can also test a human peritoneal macrophage cell line instead of THP-1 using a similar experimental design. In vivo experimentation can also be conducted by using animal models. Although this is not the most favorable method, it is vital to the next step.

ACKNOWLEDGEMENTS

We thank Ms. Yu Mi Jeong and Dr. Joshua B. Slee for their expertise and technical support throughout this experiment and Mr. Jeff Noseworthy from Olympus for allowing us to use a high-quality fluorescence microscope and CellSens™ software. We are also grateful to Heather Amatore and Mit Patel for their assistance.

REFERENCES

1. Khazen W., M'bika J.P., Tomkiewicz C., Benelli C., Chany C., Achour A., Forest C. (October 2005). "Expression of macrophage-selective markers in human and rodent adipocytes". *FEBS Letters*. 579 (25): 5631–4. doi:10.1016/j.febslet.2005.09.032
2. McDougal, W. (n.d.). Macrophages: definition, function & types. Retrieved October 29, 2017, from <http://study.com/academy/lesson/macrophages-definition-function-types.html>
3. Fujiwara, N., & Kobayashi, K. (2005). Macrophages in inflammation. *Current Drug Targets Inflammation & Allergy*, 4(3), 281-6. doi:10.2174/1568010054022024
4. Chaplin, D. D. (2010, February). Overview of the immune response. *The Journal of Allergy and Clinical Immunology*, 125(2), S3-S23. doi:10.1016/j.jaci.2009.12.980
5. Tsuchiya, S., Yamabe, M., Yamaguchi, Y., Kobayashi, Y., Konno, T., & Tada, K. (1980). Establishment and characterization of a human acute monocytic leukemia cell line (THP-1). *International Journal of Cancer*, 26(2), 171-6. doi: 10.1002/ijc.2910260208
6. Bosshart, H., & Heinzemann, M. (2016). THP-1 cells as a model for human monocytes. *Annals of Translational Medicine*, 4(21), 438. doi:10.21037/atm.2016.08.53
7. Qin, Z. (2012). The use of THP-1 cells as a model for mimicking the function and regulation of monocytes and macrophages in the vasculature. *Atherosclerosis*, 221 1, 2-11. doi:10.1016/j.atherosclerosis.2011.09.003
8. Anderson, J.M., Rodriguez, A., Chang, D.T. Foreign body reaction to biomaterials. *Seminars in Immunology*. 2008, 20, 86–100. doi:10.1016/j.smim.2007.11.004
9. Paparella, D.; Yau, T.M.; Young, E. Cardiopulmonary bypass induced inflammation: Pathophysiology and treatment. An update. *European Journal of Cardio-Thoracic Surgery*. 2002, 21, 232–244. doi:10.1016/s1010-7940(01)01099-5
10. Furuhashi, M., & Hotamisligil, G. S. (2008). Fatty acid-binding proteins: role in metabolic diseases and potential as drug targets. *Nature Reviews. Drug Discovery*, 7(6), 489. doi:10.1038/nrd2589
11. Passos, M. E. P., Alves H. H. O., Momesso, C. M., Faria, F. G., Murata, G., Oury-Boaventura, M. F., Hatanaka, E., Massao-Hirabara, S., & Gorjao, R. (2016). Differential effects of palmitoleic acid on human lymphocyte proliferation and function. *Lipids in Health and Disease*, 15(217), 1-11. doi:10.1186/s12944-016-0385-2.
12. Souza, Camila O., Teixeira, Alexandre A.S., Biondo, Luana A., Silveira, Loreana S., Calder, Philip C., Neto, & Jose C. Rosa. (2017). Palmitoleic acid reduces the inflammation in LPS-stimulated macrophages by inhibition of NF- κ B, independently of PPARs. *Clinical and Experimental Pharmacology and Physiology*, 44, 566-575. doi:10.1111/1440-1681.12736
13. Slee, J. B., Alferiev, I. S., Nagaswami, C., Weisel, J. W., Levy, R. J., Fishbein, I., & Stachelek, S. J. (2016, May). Enhanced Compatibility of CD47-functionalized vascular stents. *Biomaterials*, 87, 82-92. doi:10.1016/j.biomaterials.2016.02.008
14. Vroman, L., Adams, A.L., Fischer, G.C., Munoz, P.C. (1980). Interaction of high molecular weight kininogen, factor XII, and fibrinogen in plasma at interfaces. *Blood*. 55, 156–9. doi:10.1182/blood.v55.1.156.bloodjournal551156
15. Bovine Serum Albumin. (n.d.). Retrieved from <https://www.biowest.net/products/serum/bovine-serum-albumin-bsa/>
16. Alsabeeh, N., Chausse, B., Kakimoto, P. A., Kowaltowski, A. J., & Shiriha, O. (2018). Cell culture models of fatty acid overload: Problems and solutions. *Biochimica et Biophysica Acta. Molecular and cell biology of lipids*, 1863(2), 143–151. doi:10.1016/j.bbalip.2017.11.006
17. P-values of the results were calculated using the GraphPad QuickCalcs Web site: <https://www.graphpad.com/quickcalcs/ttest1.cfm> (accessed April 2018).

18. Cobelli, N., et. al. (2011). Mediators of the inflammatory response to joint replacement devices. *Nature Reviews Rheumatology*, 7, 601-608. doi:10.1038/nrrheum.2011.128
19. Takeda, K., et al., (1989). Conformational change of bovine serum albumin by heat treatment. *Journal of Protein Chemistry*, 5, 653-659. doi:10.1007/bf01025605

ABOUT STUDENT AUTHORS

Sheherbano Hussain is an undergraduate student at Penn State University Park majoring in Biochemistry with a Cell and Molecular Biology option. After graduating in May of 2020 with a B.S. degree, she hopes to apply her skills from undergraduate research for future work in the medical field.

Zoha Babar is an undergraduate student at Penn State University Park. She will graduate in May, 2020 with a B.S. degree in Biology. She plans to pursue a career in the medical field.

Jimmy Hadid is an undergraduate student at Penn State University Park. He will graduate in May, 2020 with a B.S. degree in Biology with a neuroscience option. In the future, he is planning on working in the medical field.

PRESS SUMMARY

The immune system works to defend the body by enacting the host inflammatory response towards foreign materials. Major players in this response are macrophages, which are white blood cells that attack and engulf molecules that are considered 'non-self.' While the host inflammatory response is crucial for protection, it poses a significant issue in healthcare. Since the immune system does not recognize medical implants and prosthetics, macrophages attack the foreign substance, causing deterioration and infection. To suppress the effects of the host response, this research was undertaken to investigate the ability of 'self-like' molecules to reduce activated THP-1 cell (macrophage) attachment to polyurethane, a commonly used polymer in prosthetics. Previous research indicates the anti-inflammatory character of palmitoleic acid in cell culture; and, bovine serum albumin is a common supplement used in tissue culture for THP-1 cells. The reduction of macrophage adherence to polyurethane films coated with palmitoleic acid, palmitoleic acid conjugated to bovine serum albumin, and bovine serum albumin using a tested *in vitro* THP-1 bioassay is presented.

Music Intervention in Undergraduates: the Relationship between Heart Rate Variability and State Anxiety

Annalise J. Tolley* & Robert S. Vick

Department of Biology, Elon University, Elon, NC

<https://doi.org/10.33697/ajur.2020.003>

Student: atolley2@elon.edu*

Mentor: vickers@elon.edu

ABSTRACT

A low heart rate variability (HRV) is indicative of autonomic inflexibility, which has important implications for physical and psychological health. This study investigates autonomic functioning and its relationship to state anxiety in the context of music intervention. A within-subjects, quasi-experimental design was used with undergraduates, a population frequently impacted by state anxiety. Participants pre-selected music that they identified as the most “relaxing” before being administered the State-Trait Anxiety Inventory (STAI). Baseline physiological indices were then collected using an EKG, followed by administration of the selected music and an additional collection of physiological indices. After the full duration of music intervention was administered, participants took the STAI once more. Results indicated that even short durations of music listening may serve to reduce anxiety in the undergraduate student, as was evidenced by the significant increase in physiological indices and decrease in self-reported anxiety levels. However, while changes in HRV and anxiety may trend together, results indicated no statistical association between these parameters.

KEYWORDS

Heart Rate Variability; Anxiety; Music Intervention; Autonomic Tone; Vagal Control; Autonomic Flexibility; College Students; State Anxiety

INTRODUCTION

State Anxiety in the Undergraduate

Undergraduates face a variety of stressors in their lives. Their competitive academics, leadership roles, volunteer positions, and extracurriculars are just a few of the many commitments that contribute to a hectic schedule. While all of these expectations may seem benign, statistics reveal a more disconcerting picture. Anxiety in undergraduates is on the rise, with almost 60% of students in 2016 reporting overwhelming anxiety in the past 12 months and 85% reporting being overwhelmed by all that had to be done.¹ Close to 50% of students also reported that their academics felt “traumatic or very difficult to handle.”¹ While only 17% of the students surveyed were diagnosed with anxiety by a medical professional,¹ it is clear that these feelings of sub-clinical anxiety are distressing to the individual and deserve attention. As such, diagnosis or not, the mental health of undergraduates is increasingly in jeopardy.

Anxiety can be classified as either state anxiety or trait anxiety.² Trait anxiety refers to personality-driven anxiety. This is the general tendency for someone to perceive the environment as threatening, and it is relatively constant. State anxiety, on the other hand, is subjective and fleeting. It is a psychological and physiological response to situational feelings of tension.³ Given the large proportion of undergraduates who report feelings of anxiety and tension without qualifying for diagnosis, these individuals are presumably feeling the effects of situational, state anxiety. From this point forward, unless otherwise specified, the use of the word anxiety refers to anxiety at large, rather than a clinical diagnosis.

Measuring Anxiety Through Autonomic Flexibility

Various self-reported measures exist to determine anxiety, but research has also provided evidence that anxiety can be measured physiologically using heart rate variability (HRV).⁴⁻⁷ HRV is a physiological indicator of autonomic/ vagal tone, which is the interplay between the sympathetic and parasympathetic branches of the autonomic nervous system (ANS). Vagal tone has long been accepted as an indicator of cardiac and physical health,^{8,9} with low vagal tone correlating to diabetes, ischemic heart failure, and early mortality.^{8,9} Low vagal tone is an indicator of poor health because it is a measurement of poor autonomic flexibility. High autonomic flexibility is an adaptive trait allowing for organismic stability in the face of a complex, dynamic environment.^{5,10}

Measuring autonomic flexibility through vagal tone provides important ties to other forms of flexibility, such as psychological flexibility.^{11, 12} Vagal control is unique because it integrates both the ANS and the CNS via the central autonomic network (CAN), which is mediated by the vagal nerve.^{5, 13} The CAN provides a pathway for autonomic tone to reflect and impact behavioral processes and vice versa. It is the mechanistic explanation for the impact of vagal tone on mental health and is the basis of the autonomic flexibility-neurovisceral integration model of anxiety. This model states that low vagal tone is a predictor of all forms of anxiety due to a rigid inability to dis-inhibit negative feedback mechanisms both physiologically and psychologically.⁵ If this model is accepted as true, HRV (as a measurement of vagal tone) measures not only autonomic flexibility but psychological flexibility as well. Relevant to this study, lack of psychological flexibility can present as anxiety, which is the inability to dis-inhibit an inappropriate response following a benign stimulus.¹⁴

The Present Study

To further explore this psychophysiological model of anxiety, investigators of this study included music intervention. Research indicates that music can alter both psychological and autonomic states,¹⁴⁻¹⁸ although the exact mechanism of music's effect on the ANS is not entirely known.^{14, 19, 20} Music intervention served a primary purpose in this study: to function as a situational stimulus that could decrease state anxiety as measured by self-report and physiological indices. Any resultant changes in parameters would then be compared. If psychological and physiological anxiety are integrated to the degree that the autonomic flexibility-neurovisceral integration model theorizes, then comparison should yield a significant relationship. This finding would support that HRV is, in fact, a psychophysiological measurement rather than just a physiological one.

Other studies investigating music intervention and its effects on state anxiety have shown promising results that provide reason to believe it may be effective for undergraduates.^{15, 16, 18} While not the primary goal of the study, if the hypothesis of this study were confirmed, results may imply that music intervention is a brief and efficacious method for reducing state anxiety in the undergraduate.

In summation, the present study seeks to determine the extent of the possible relationship between anxiety and HRV by using music intervention. Based on the literature, the study is conducted with the assumption that STAI score is a measurement of psychological anxiety, and HRV is a measurement of physiological anxiety. It is hypothesized that (1) music intervention will lead to decreased psychological and physiological anxiety, as measured through STAI score and HRV indices, respectively; and (2) the changes in psychological and physiological anxiety will be related to one another.

METHODOLOGY

Participants

Healthy undergraduate students at Elon University (n=32) were recruited to participate (mean age 19.2 yrs., range 17-22 yrs.). Participants were not screened for anxiety before participating; there were no exclusion criteria. Previous sample sizes in the field and prior sample size calculations provide basis to believe that 32 participants may allow for adequate assessment ($\beta=0.8$) of a medium to large effect size.¹⁸ However, across studies, sample size calculations are inconsistent. This study was approved by the Elon University's Institutional Review Board (#17-083). The purpose of the study, methods, potential risks, and benefits were outlined to the participants by an informed consent form. The participants' rights and privacy were protected throughout the experiment, and participants had the right to withdraw from the study at any given time.

Study Design

All participants were part of both the pre- and post-experimental conditions to control for individual differences. To control for HRV variation as a result of circadian rhythm, which differs throughout the day,²¹ data collection occurred within the span of 5pm to 6pm. After written consent was obtained, the State-Trait Anxiety Inventory (STAI) was administered through pencil and paper to the subjects. After the survey was completed, the participants were played a sample from four preselected songs and asked to choose the most relaxing song option, which would be played in the latter part of the experiment. The subjects were then attached to an EKG, and seven minutes of baseline HRV data was collected. The pre-selected music was then played for an additional seven minutes as HRV data continued to be collected. After the fourteen minutes of EKG data collection, participants were asked to complete the STAI Y-1 once again. The total experiment duration was 30 minutes.

Psychological Measure

To measure anxiety, the State-Trait Anxiety Inventory (Y-1 and Y-2) was used. A leading limitation in anxiety research thus far has been the surveys used for evaluation. Very few tests have the ability to differentiate between depression and anxiety,²² and due to this methodological flaw, much of the research that explores anxiety and HRV may very well be an extension of preexisting HRV and depression research.²³ The State-Trait Anxiety Inventory (STAI) have been shown to better distinguish anxiety with high reliability and validity across different languages, cultures, and populations.^{22, 24-26} The STAI includes two 20-questions surveys, in which the participant is asked to rate their response on a Likert 1-4 scale. The Y-1 measures state anxiety (anxiety

attributed to external stimulus), and the Y-2 measures trait anxiety (anxiety attributed to personality). The average alpha reliability coefficients for college students were 0.92 and 0.91 on the Y-1 and Y-2, respectively.² A higher score indicates more anxiety, and it is thought that scores below 40 are not considered clinically significant symptoms.²²

Physiological Measure

HRV was monitored and measured with lead I of a three-lead EKG. Electrodes were connected to a computer-based data acquisition system (ADInstruments, Boulder, CO) and analyzed using Lab Chart Pro and its HRV module (ADInstruments).

Acoustic Stimuli

This study offered participants four options for music intervention, and participants were asked to choose which song they found most “relaxing.” In preexisting studies, participants have been allowed to self-select music used for their intervention,²⁰ which leaves room for a myriad of confounding variables. To reduce these confounding effects without completely removing the autonomy of the participant, subjects were given four musical genre options (jazz, orchestral, elemental healing sounds, and instrumental pop) controlled for tempo (60 bpm),²⁰ lyrics (no lyrics),²⁷ and relative volume.¹⁷ Because the aim of the study was to measure whether or not music could lower psychological and physiological measures of state anxiety in subjects, it was important to avoid accidental arousal of the individual via PNS withdrawal or subjective dislike due to the song.^{17, 20} Of the four available options, subjects were able to choose the most subjectively relaxing choice. Research has shown that maintaining options in music intervention studies is important for both the subjects and the experimental design.¹⁶

Statistical Analysis

IBM SPSS Statistics 24 was used for data analysis. Descriptive statistics were run, and data was normally distributed. A paired t-test was used to evaluate if there were differences between the STAI scores and HRV indices pre- and post- music administration. Principal components analysis was used to derive a composite HRV index, which was then compared using repeated measures ANOVA to STAI score. Significance was set at $p < 0.05$ for all analyses.

RESULTS

Psychological Measure

For STAI score, the Y-1 score (state anxiety) was used in analysis and the Y-2 (trait anxiety) was not. Because the Y-2 is representative of general, personality-related anxiety, it would not be expected to change based on external stimuli. It therefore was only administered once. A one-way, unpaired t-test between individuals with clinically significant STAI scores (above 40 on Y-2) and non-clinically significant anxiety indicated no significant difference between any of the HRV indices. As a result, Y-2 was not used at this time for data analysis.

The STAI Y-1 score decreased for 31 out of 32 of the participants. To test if the overall reduction in Y-1 score was significant, a one-way, paired t-test was run. **Figure 1** illustrates the results: a significant reduction in Y-1 score ($p < 0.0001$) from an average STAI score of 38 ± 8.87 before music administration to 29 ± 5.98 after music administration.

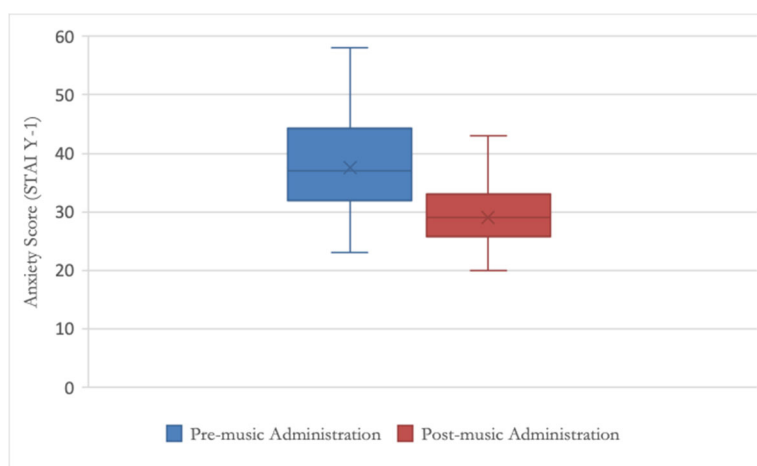


Figure 1. Average STAI score before and after music administration. Error bars are SEM (n=32).

Physiological Measure

In addition to the STAI score, average RR, SDRR, pRR50, and HR were collected both pre- and post-music administration. RR is the time interval measurement from one R peak to the next in milliseconds (ms). The SDRR is standard deviation of all RR measurements (ms). The pRR50 can be thought of as the percentage of RR interval differences greater than 50 ms from the previous RR. HR is average heart rate (bpm). The data analysis in this experiment used time domain analysis. There was high variation in HRV range based on the individual. To both account for this and exclude ectopic beats, the data was plotted and the range was adjusted to include the majority of heart beats. HRV intervals for each participant were kept within 400 ms and ranged from 400–800 ms to 800–1200 ms.

Paired t-tests were used to analyze physiological indices before and after music intervention. Results indicated that average RR increased from 743.08 ± 122.71 ms ($p < 0.001$) to 759.5 ± 123.78 ms, SDRR increased from 51.84 ± 25.29 ms to 57.27 ± 24.62 ms ($p < 0.01$), pRR50 increased from $17.29 \pm 17.9\%$ to $20.61 \pm 20.3\%$ ($p < 0.01$), and average HR decreased from 83.19 ± 12.78 bpm to 81.46 ± 12.70 bpm ($p < 0.001$). Participants showed a statistically significant difference in HRV indices after music administration. This difference reflected physiological relaxation after music administration, as determined by an increased RR, SDRR, and pRR50 and a decrease in HR, is reflected in **Table 1**.

	Pre-music	Post-music	p-value
Average RR (ms)	743.1	759.5	<0.005
SDRR (ms)	51.84	57.27	<0.005
pRR50 (%)	17.29	20.61	<0.005
Average HR (bpm)	83.19	81.46	<0.005

Table 1. Average HRV indices pre- and post-music administration. The p-values were determined with one-tailed, paired t-tests.

Psychological and Physiological Measure

Although results indicate that both STAI scores and HRV indices are significantly different after music administration, to determine the extent to which these variables may be related, a Pearson's correlation was run between STAI and each HRV index both pre- and post-music administration. It was determined that there was no significant correlation between anxiety and average RR ([pre] $r = -0.11$, $p = 0.95$; [post] $r = 0.03$, $p = 0.87$), anxiety and SDRR ($r = 0.01$, $p = 0.98$; $r = -0.18$, $p = 0.34$), anxiety and pRR50 ($r = 0.14$, $p = 0.45$; $r = -0.04$, $p = 0.82$), or anxiety and HR ($r = 0.02$, $p = 0.90$; $r = -0.46$, $p = 0.80$), which is summarized in **Table 2**.

Pearson's Correlation			
Variables	Treatment	R-value	P-value
RR v. STAI	Pre	-0.11	0.95
	Post	0.03	0.87
SDRR v. STAI	Pre	0.01	0.98
	Post	-0.18	0.34
pRR50 v. STAI	Pre	0.14	0.45
	Post	-0.46	0.82
HR v. STAI	Pre	0.02	0.90
	Post	-0.46	0.80

Table 2. Pearson's correlation table between HRV indices and anxiety measure.

Because Pearson's correlation is most often used to investigate the relationship between either an independent variable and a dependent variable or two factorials, principal component analysis (PCA) was used to explore the relationship between the five dependents. The first run of PCA yielded a correlation matrix that showed no significant correlation between STAI score and RR, SDRR, pRR50, and HR with r-values respectively as follows: $r = -0.28$, $r = -0.110$, $r = 0.008$, $r = 0.030$. After STAI was removed, PCA was again run. **Table 3** shows the produced correlation matrix and the scree plot, which were used to identify relevant component PC1.

The PC1 was used as the composite variable of HRV indices and a univariate ANOVA was run to confirm a statistically significant physiological difference after music intervention. The average PC1 pre-music was -0.092 ± 1.01 and the average PC1 post-music was 0.092 ± 1.00 . A statistically significant difference was confirmed using a paired sample t-test ($p < 0.005$). A negative PC1 value, as is the value pre-music, corresponds to an anxious physiological state, as measured by a high HR, low average RR, low SDRR, and low pRR50. Conversely, a positive PC1 value, as in the post-music score, would correspond to a

relaxed physiological state, as measured by a lower HR, higher average RR, higher SDRR, and higher pRR50. Results indicate that after music administration, participants were more physiologically relaxed than before music administration.

A Pearson's correlation was once more run but this time between STAI score and PC1 pre- and post-music administration. No statistically significant correlation was found between STAI score and PC1 ($r = -0.043$, $p = 0.737$).

		RR	SDRR	pRR50	HR
Correlation	RR	1.000	0.592*	0.741*	-0.981*
	SDRR	0.592	1.000	0.784*	-0.555*
	pRR50	0.741	0.784	1.000	-0.711*
	HR	-0.981	-0.555	-0.711	1.000

Table 3. PCA correlation matrix of HRV indices with excluded STAI score. Relevant correlations are starred.

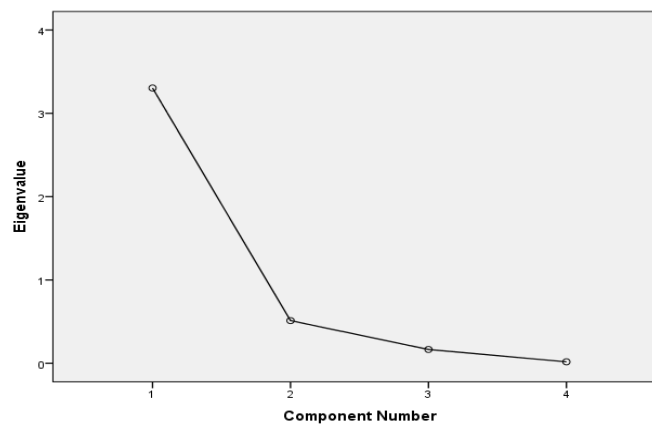


Figure 2. Scree plot of HRV components. Following the broken stick model, the relevant component was PC1.

DISCUSSION

It was hypothesized that (1) music intervention would lead to decreased psychological and physiological anxiety, as measured through STAI score and HRV indices, respectively; and (2) the changes in psychological and physiological anxiety would be related to one another, as determined by statistical significance.

Results supported hypothesis one and indicated that relaxing music intervention resulted in a statistically significant decrease in self-reported state anxiety, as measured by the STAI Y-1. These findings support previous research on the positive effects of music intervention on psychological and physiological measures.^{19, 28} However, many of these studies administer music for 20-30 minutes. This study shortened the duration of administered music to seven minutes, which is more manageable for situational intervention given the busy schedule of the undergraduate. Results indicated a significant benefit to undergraduates from just seven minutes of music, lowering state anxiety by 23% on average. Thus, the use of music intervention to lower state anxiety, even for a few minutes a day, should be encouraged to the undergraduate population who greatly suffers from daily anxiety.¹ Because this study investigated the lowered state anxiety at baseline, it would be interesting to further investigate the persistence of this effect. Could a lowered state anxiety mitigate the anxiety-inducing effects of a situational stressor? If so, this could have practical implications for the undergraduate, who often knows when situational stressors will occur, such as a test or presentation.

In addition to a significant decrease in self-reported anxiety, results further support hypothesis one: music intervention would result in increased HRV indices, a reflection of reduced physiological anxiety. Although there is little research on the effect of music specifically on HRV, there is a breadth of literature that examines the effect of music on other physiological indices related to autonomic functioning.^{17, 29, 30} Research has been inconclusive, but this study hypothesized that with the administration of controlled, “relaxing” music, there would be a significant effect on autonomic modulation as measured by HRV, which was supported by the results. PC1, the component representative of all the HRV indices, significantly shifted in the positive direction after music intervention. A positive shift reflected a decrease in physiological anxiety through an increase in HRV, as measured by higher RR, SDRR, pRR50, and lower HR.

Hypothesis two postured that the significant changes in HRV and STAI score would be statistically correlated to one another based on the autonomic flexibility-neurovisceral integration model of anxiety. This was unsupported. While results indicated that both dependents changed significantly in the predicted direction after music intervention, further analysis revealed that the changes in psychological and physiological anxiety were not statistically correlated to each other, as measured by Pearson's correlation. Since both Pearson's correlation and correlation matrices assume a linear relationship between variables, it is possible that the changes in HRV and anxiety were not linearly related. However, graphs between STAI and PC1 as well as STAI and HRV individual indices indicated no clear trend to fit the data. Thus, while it is likely that HRV and anxiety trend together, this study found no empirical base to support a significant relationship. As such, HRV may only be a reliable measurement of physiological anxiety, rather than a psychophysiological measurement of anxiety. If this were the case, future research should take caution using HRV as a measurement of anxiety independent of self-report.

Although state anxiety and trait anxiety are somewhat related, correlations between the state and trait anxiety for college students has varied from 0.11 to 0.53 for females and 0.37 to 0.67 for males.² Because of this variability, when conclusions are drawn, they should not be extrapolated to anxiety in general but rather to the specific anxiety being measured. This study's design has implications only for state anxiety, which is anxiety produced from situation, not disposition. It is therefore possible that HRV is more indicative of trait anxiety, but not state anxiety.

Several limitations exist, and results should be interpreted modestly. These limitations include the generalizability of the data, given the minimal diversity of the subjects collected and small sample size. Because subjects were recruited from Elon University, Caucasian females were overrepresented, while males and minority groups were underrepresented. Previous sample sizes in the field and prior sample size calculations provide basis to believe that 32 participants may allow for adequate assessment ($B=0.8$) of a medium to large effect size.¹⁸ However, across studies, sample size calculations are inconsistent. This poses the risk that many studies in the field, including this one, do not qualify for adequate statistical power.

To reduce the effects of individual differences and account for the high degree of individual variability associated with HRV, a within-subjects, pretest/posttest design was employed. It is worth noting that it is possible that the musical pieces chosen may have had unknown effects on autonomic tone for different individuals. This study tried to maintain a sense of choice for the participants while holding constant variables known to affect autonomic tone and perceptual experience (i.e. lyrics, tempo, volume, genre),^{17, 20, 27} but it is possible that the subjective perception of music could alter an individual's ANS unknowingly. Likewise, it is possible that not all participants found the music subjectively relaxing.

The aim of this study was to develop a deeper understanding of the possible relationship between psychological state anxiety and physiological anxiety, as measured by STAI and HRV respectively. The hypothesis that music intervention would both lower self-reported anxiety scores and increase HRV was supported, which provides evidence that brief music intervention is a successful tool to lower state anxiety in undergraduate students and to promote autonomic relaxation. The prediction that the change in psychological anxiety (STAI score) and that the change in physiological anxiety (HRV indices) would be statistically correlated was unsupported. Therefore, despite the significant effect that music had on both autonomic functioning and self-reported anxiety, these effects seem only to trend together, as they are not statistically related. While these findings strengthen the potential of music to reduce state anxiety and promote autonomic relaxation in the undergraduate, they also suggest that interpreting HRV as a measurement of something other than physiological anxiety should be approached with caution. Physiological anxiety may trend, even interface, with psychological state anxiety, but not to a significant degree. This may qualify the broader implications of HRV as a measurement of state anxiety.

ACKNOWLEDGEMENTS

The author thanks Dr. Vick for his critical support and contribution to this study as a whole and to this manuscript. Additionally, the author thanks the staff of the Elon Biology Department for their mentorship and guidance. This research would not have been possible without funding from the Undergraduate Research Program and the Elon College Fellow's program.

DISCLOSURE STATEMENT

There are no conflicts of interest to disclose.

REFERENCES

1. American College Health Association: National College Health Assessment. (2016) *Fall 2016 Reference Group Executive Summary* (Reference Group Executive Summary) (p. 14).
2. Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., and Jacobs, G. A. (1983) *Manual for the State-Trait Anxiety Inventory*, Consulting Psychologists Press, Palo Alto, CA.

3. Horikawa, M., and Akihiro, Y. (2012) The relationships among trait anxiety, state anxiety and the goal performance of penalty shoot-out by university soccer players, *PLoS One*, 7(4). <https://doi.org/10.1371/journal.pone.0035727>
4. Chalmers, J. A., Quintana, D. S., Abbott, M. J.-A., and Kemp, A. H. (2014) Anxiety disorders are associated with reduced heart rate variability: A meta-analysis, *Front Psychiatry*, 5(80). <https://doi.org/10.3389/fpsy.2014.00080>
5. Friedman, B. H. (2007) An autonomic flexibility-neurovisceral integration model of anxiety and cardiac vagal tone, *Biol Psychol* 74(2), 185–199. <https://doi.org/10.1016/j.biopsycho.2005.08.009>
6. Sharma, R.K., Yatan, P. S. B., Sagar, R., Deepak, K. K., and Mehta, M. (2011) Heart rate variability study of childhood anxiety disorders, *J Cardiovasc. Dis. Res.* 2(2), 115–122. <https://doi.org/http://dx.doi.org/10.4103/0975-3583.83040>
7. Watkins, L. L., Grossman, P., Krishnan, R., and Sherwood, A. (1998) Anxiety and vagal control of heart rate, *Psychosom Med* 60(4), 498–502. <https://doi.org/10.1097/00006842-199807000-00018>
8. Stein P. K., Kleiger R. E. (1999) Insights from the study of heart rate variability. *Annu Rev Med* 50, 249–261. <https://doi.org/10.1146/annurev.med.50.1.249>
9. Shigetoh, Y., Adachi, H., Yamagishi, S., Enomoto, M., Fukami, A., Otsuka, M., Kumagai, S., Furuki, K., Nanjo, Y., Imaizumi, T. (2009) Higher heart rate may predispose to obesity and diabetes mellitus: 20-year prospective study in a general population, *Am J Hypertens* 22(2), 151–155. <https://doi.org/10.1038/ajh.2008.331>
10. Lloyd, D., Miguel, A. A., and Cortassa, S. (2001) Why homeodynamics, not homeostasis?, *Sci World J* 1, 133–145. <https://doi.org/10.1100/tsv.2001.20>
11. Mori, K., and Iwanaga, M. (2014) Resting physiological arousal is associated with the experience of music-induced chills, *Int J Psychophysiol* 93(2), 220–226. <https://doi.org/10.1016/j.ijpsycho.2014.05.001>
12. Friedman, B. H., and Thayer, J. F. (1998) Anxiety and autonomic flexibility: A cardiovascular approach, *Biol Psychol* 49(3), 303–323. [https://doi.org/10.1016/S0301-0511\(97\)00027-6](https://doi.org/10.1016/S0301-0511(97)00027-6)
13. Porges, S. W. (2009) The polyvagal theory: New insights into adaptive reactions of the autonomic nervous system, *Cleve Clin J Med* 76(Suppl 2), S86–90. <https://doi.org/10.3949/ccjm.76.s2.17>
14. Ellis, R. J., and Thayer, J. F. (2010) Music and autonomic nervous system (dys)function, *Music Percept* 27(4), 317–326. <https://doi.org/10.1525/mp.2010.27.4.317>
15. Bradt, J., and Dileo, C. (2009) Music for stress and anxiety reduction in coronary heart disease patients, *Cochrane Database Syst Rev* (2). <https://doi.org/10.1002/14651858.CD006577.pub2>
16. Cooke, M., Chaboyer, W., and Hiratos, M. A. (2005) Music and its effect on anxiety in short waiting periods: a critical appraisal, *J Clin Nurs* 14(2), 145–155. <https://doi.org/10.1111/j.1365-2702.2004.01033.x>
17. do Amaral, J. A. T., Guida, H. L., Vanderlei, F. M., Garner, D. M., Osório, E., de Abreu, L. C., and Valenti, V. E. (2015) The effects of musical auditory stimulation of different intensities on geometric indices of heart rate variability, *Altern Ther Health Med* 21(5), 16–23.
18. Wang, S.-M., Kulkarni, L., Dolev, J., and Kain, Z. N. (2002) Music and preoperative anxiety: A randomized, controlled study, *Anesth Analg* 94(6), 1489–1494. <https://doi.org/10.1213/0000539-200206000-00021>
19. Wu, P.-Y., Huang, M.-L., Lee, W.-P., Wang, C., and Shih, W.-M. (2017) Effects of music listening on anxiety and physiological responses in patients undergoing awake craniotomy, *Complement Ther Med* 32, 56–60. <https://doi.org/10.1016/j.ctim.2017.03.007>
20. Nilsson, U. (2008) The anxiety- and pain-reducing effects of music interventions: A systematic review, *AORN J* 87(4), 780–807. <https://doi.org/http://dx.doi.org/10.1016/j.aorn.2007.09.013>
21. Guo, Y.-F., and Stein, P. K. (2003) Circadian rhythm in the cardiovascular system: Chronocardiography, *Am Heart J* 145(5), 779–786. [https://doi.org/10.1016/S0002-8703\(02\)94797-6](https://doi.org/10.1016/S0002-8703(02)94797-6)
22. Julian, L. J. (2011) Measures of Anxiety, *Arthritis Care Res*, 63(0 11). <https://doi.org/10.1002/acr.20561>
23. Buckby, J. A., Yung, A. R., Cosgrave, E. M., and Killackey, E. J. (2007) Clinical utility of the Mood and Anxiety Symptom Questionnaire (MASQ) in a sample of young help-seekers, *BMC Psychiatry* 7(50). <https://doi.org/10.1186/1471-244X-7-50>
24. Hishinuma, E. S., Miyamoto, R. H., Nishimura, S. T., Goebert, D. A., Yuen, N. Y. C., Makini, G. K., Andrade N. N., Johnson, R. C., and Carlton, B. S. (2001) Prediction of anxiety disorders using the State-Trait Anxiety Inventory for multiethnic adolescents, *J Anxiety Disord* 15(6), 511–533. [https://doi.org/10.1016/S0887-6185\(01\)00079-2](https://doi.org/10.1016/S0887-6185(01)00079-2)
25. Vitasari, P., Wahab, M. N. A., Herawan, T., Othman, A., and Sinnadurai, S. K. (2011) Re-test of State Trait Anxiety Inventory (STAI) among engineering students in Malaysia: Reliability and validity tests, *Procedia Soc Behav Sci* 15, 3843–3848. <https://doi.org/10.1016/j.sbspro.2011.04.383>
26. Zeleníková, R., Homzová, P., Homza, M., and Bužgová, R. (2017) Validity and reliability of the Czech version of the Amsterdam Preoperative Anxiety and Information Scale (APAIS), *J Perianesth Nurs* 32(5), 429–437. <https://doi.org/10.1016/j.jopan.2016.03.007>
27. Ribeiro, M. K. A., Alcântara-Silva, T. R. M., Oliveira, J. C. M., Paula, T. C., Dutra, J. B. R., Pedrino, G. R., Simões, K., Sousa, R. B., and Rebelo, A. C. S. (2018) Music therapy intervention in cardiac autonomic modulation, anxiety, and depression in mothers of preterms: Randomized controlled trial, *BMC Psychol* 6(57). <https://doi.org/10.1186/s40359-018-0271-y>

28. Sendelbach, S. E., Halm, M. A., Doran, K. A., Miller, E. H., and Gaillard, P. (2006) Effects of music therapy on physiological and psychological outcomes for patients undergoing cardiac surgery, *J Cardiovasc Nurs* 21(3), 194–200. <https://doi.org/10.1097/00005082-200605000-00007>
29. Okada, K., Kurita, A., Takase, B., Otsuka, T., Kodani, E., Kusama, Y., Atarashi, H., and Mizuno, K. (2009) Effects of music therapy on autonomic nervous system activity, incidence of heart failure events, and plasma cytokine and catecholamine levels in elderly patients with cerebrovascular disease and dementia, *Int Heart J* 50(1), 95–110. <https://doi.org/10.1536/ihj.50.95>
30. Phipps, M. A., Carroll, D. L., and Tsiantoulas, A. (2010) Music as a therapeutic intervention on an inpatient neuroscience unit, *Complement Ther Clin Pract* 16(3), 138–142. <https://doi.org/10.1016/j.ctcp.2009.12.001>

ABOUT STUDENT AUTHOR

Annalise J. Tolley graduated from Elon University in May of 2018 with a Bachelor of Science in biology, a concentration in Foundations for Medical Sciences, and a minor in psychology. She has recently completed a one-year fellowship in India working for a public health NGO that works with communities to provide health services, promote empowerment, and enable sustainable development among the poor and marginalized. After refining her research interests over the next year, she intends to pursue a PhD in Social Psychology.

PRESS SUMMARY

Research suggests that the number of heart beats per minute might be less significant as a determinant for overall health than the interval spacing of those beats. Heart rate variability (HRV) measures the variation in time interval from one beat to the next. A lower HRV indicates less flexibility in a dynamic environment, which is believed to indicate poor physical and mental health. Current theories of anxiety suggest that HRV may be a psychophysiological measure of anxiety, although the degree to which the psychological and the physiological manifestations of anxiety interact need further study. This study explored this possible relationship by measuring changes in psychological anxiety (as measured by State-Trait Anxiety Inventory) and physiological anxiety (as measured by HRV) in response to music intervention. These changes were then compared. Undergraduates were chosen as the population to investigate this trend, as they are often subject to changes in state anxiety and could benefit from brief and accessible intervention. Results indicated that even short durations of music listening reduce anxiety in undergraduates, as was evidenced by the significant decrease in self-reported state anxiety and increase in HRV indices. However, results indicated no statistical correlation between these two parameters. These results may qualify the application of HRV as a psychophysiological indicator of anxiety.

Can Expense Ratios Signal Performance? An Analysis of Equity ETFs & Mutual Funds

James Monroe Gamble IV*

Department of Economics, University of Missouri, Columbia, MO

<https://doi.org/10.33697/ajur.2020.004>

Student: jmg2g2@mail.missouri.edu*

Mentor: sbastrisandy@gmail.com

ABSTRACT

This study examines the impact of the emergence of exchange-traded funds (ETFs) as an alternative investment vehicle to mutual funds. As the number of ETFs continues to rise, we investigate potential risks and disadvantages posed by ETFs in comparison to traditional mutual funds. We compare the returns, performance, and expense ratios of ETFs to those of mutual funds. We find that expense ratios are positively correlated with actively managed mutual fund returns and that passive funds have outperformed active funds since their inception. There is downward pressure on mutual fund fees over time, suggesting increased competition between mutual funds and ETFs. We also find, up to a certain threshold, actively managed funds are worth their costs.

KEYWORDS

Exchange-Traded Fund (ETF); Mutual Fund; Investing; Fee Structure; Expense Ratio; Passive (Active) Investing; Portfolio Management; Indexing

INTRODUCTION

Innovations in financial products have shaped and reshaped how effective we are at investing. Exchange-traded funds (ETFs) are no exception. Their proliferation has brought diversification, top tier management, and sophisticated investment strategies to the retail investor.¹ No longer do mutual funds have a stronghold on the market. From 2009 to 2018, index-based equity ETFs attracted roughly one and a half times the net flows of domestic equity index mutual funds (**Figure 1**).² Investment decisions that once required a team of dedicated money managers can now be made with a few simple clicks of a mouse. Diversified asset classes in financials, consumer staples, or broad market indices are accessible to ordinary people without the need of a financial manager. This is significant as defined contribution retirement plans have begun to supplant defined benefit pension plans, placing more responsibility on individuals to manage retirement savings.

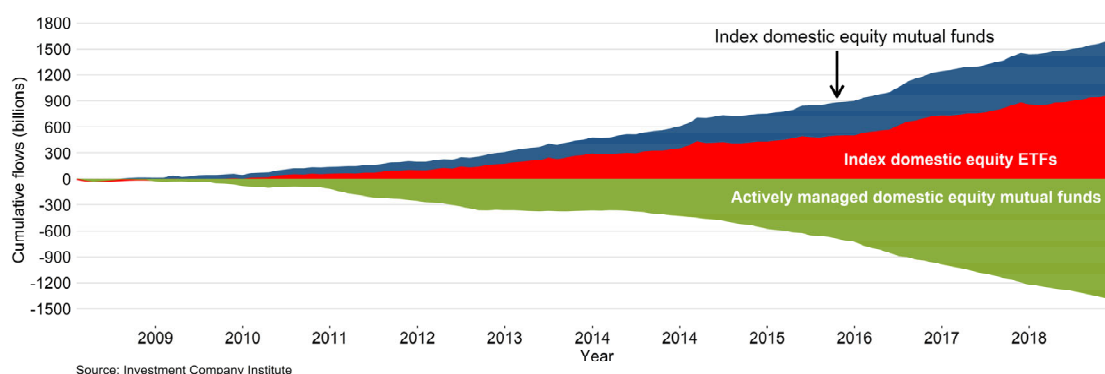


Figure 1. Mutual funds outflows have gone to ETFs.²

Cumulative flows to domestic equity mutual funds and net share issuance of domestic equity index ETFs; billions of dollars, monthly. Mutual fund data include net new cash flow and reinvested dividends; ETF data for net share issuance include reinvested dividends.

Before October 2009, domestic equity index ETF data include a small number of actively managed domestic equity ETFs.

ETFs compete with mutual funds for investor's money and are generally perceived to have sweeping benefits, such as liquidity and tax efficiencies. Innovations in financial technology (fintech) have democratized investing by making financial markets more accessible and lowering transaction costs. Increasingly, brokerage firms offer commission-free trades for ETFs allowing investors to avoid 100% of the trading costs and maximize returns. Fintech startup Robinhood disrupted the brokerage industry when it

began offering commission-free trades. Robinhood's popularity has pushed other firms to adapt by lowering or removing commissions on select ETFs. As a result, most ETFs can be bought and sold by retail investors without transaction costs.

Background

Research examining the efficiencies of ETFs exists, but there is little study into the confluence of fee structures underpinning these investments and fund performance.^{3, 4} With investment complexity comes the need for careful analysis to uncover not only the benefits but also the hidden costs to investors. Are ETFs better than mutual funds along characteristics such as fee and tax structures or performance as measured by long term returns on investment? Our study addresses this question.

For mutual funds, a wealth of research exists examining how management and fees affect value. Financial Planners, retail, and institutional investors can use our results to increase portfolio profitability for both retirement planning and capital growth. Investors, given an easy-to-implement investment technique with significant returns, can effectively manage their savings. Due to an overload of information and fear of making incorrect retirement investments, Americans frequently turn to professionals for advice. Commissions, kickbacks, and fees drawn from investor accounts create conflicts of interest.⁵ It is common practice for mutual fund providers to compensate brokers for attracting investors. Incentivized brokers may steer clients' savings into funds with higher fees and lower returns or suggest inappropriate rollovers out of low-cost retirement plans to generate commissions. The White House Council of Economic Advisers (CEA) estimates \$17 billion in losses every year to working and middle-class families due to poor investment advice.⁶

To combat the principal-agent problems that exist between brokers and their clients, the SEC adopted Regulation Best Interest (Regulation BI) in June of 2019. Regulation BI requires that broker-dealers put their client's interests first when recommending securities transactions or investment strategies. The SEC simultaneously adopted the Customer Relationship Summary (CRS) form, which requires brokers and registered investment advisors (RIAs) to state their roles clearly.⁷ Our research complements these reforms. We explore whether ETFs provide effective investment alternatives to mutual fund shares purchased through intermediaries. Investors can purchase ETFs directly, thereby avoiding fees charged by traditional mutual funds. These fees ultimately lower returns.

Factors such as tax and management expenses affect investors' per-share returns. For consistency, we focus on long-term capital gains, which are classified as investments of one year or more. In the United States, capital gains are taxed at different rates depending on income and term length. Capital gains on short-term investments, held for a year or less, are taxed as ordinary income. Long-term capital gains are taxed at maximum rates of 0%, 15%, or 20%, depending on the investor's total income. At all income levels, the maximum tax rate on long-term capital gains is less than the tax rate on ordinary income. Furthermore, qualified retirement plans can cause the income tax cost to be avoided altogether or drastically reduced. Individual Retirement Accounts (IRAs) and defined contribution 401(k) plans are common tax-deferred investment vehicles. In a tax-deferred IRA or 401(k), federal and state income taxes are deferred until funds are withdrawn in retirement. Then, withdrawals are taxed as ordinary income. Public and private employers offer these plans to employees, matching employee contributions anywhere from 0% to 100% after an initial vesting period. Other retirement accounts, such as Roth IRAs, allow after-tax income to be invested in accounts that allow tax-free withdrawals in retirement.

Understanding the effects of tax structures and transaction costs is important because investing is an exponential process. Small improvements in earnings can lead to substantial gains, while hidden costs hinder the growth of nest eggs and lengthen maturity dates. Over 35 years, one percentage point of lower returns could reduce savings by more than 25%. With a four percent effective annual return, a \$10,000 retirement investment will grow to more than \$39,000. With a three percent return, it would be just over \$28,000, a reduction of 28.7%. Our research will prove beneficial to those investing in both deferred and non-tax deferred accounts. ETFs offer the advantage of providing the diversification of a mutual fund with the liquidity of a stock, in addition to tax savings on reinvested dividends (**Figure 2**). A mutual fund must pay a capital gains tax on earnings, which is then compounded by a second tax on the investor. ETFs reinvest dividend payments or distribute them directly to shareholders, thus sparing this double taxation expense.

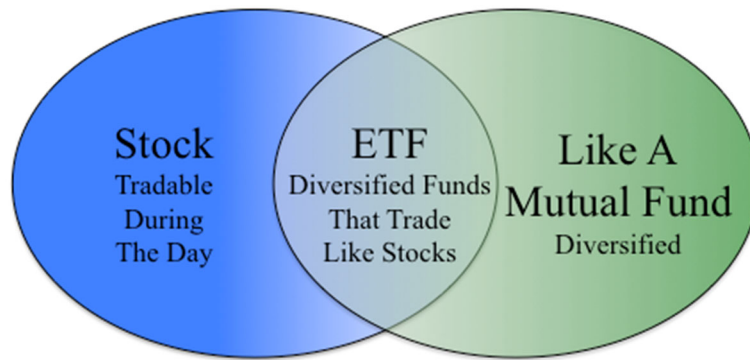


Figure 2. ETFs offer the liquidity of a stock and the diversification of a mutual fund.

Our findings are helpful to regulators making decisions on how to protect consumers best. The Securities and Exchange Commission (SEC) and the Financial Industry Regulatory Authority (FINRA) govern the types of fees that mutual funds and ETFs can charge. Fees such as 12b-1 fees, sales loads, and redemptions are regulated with caps on the amount funds can charge.¹⁰ However, many other hidden fees can be costly to investors. ETFs are one solution to simplifying fee structures as their fees can be represented by one number, the expense ratio.

Funds can be classified by their investment strategy. Actively managed funds try to beat the market while passive funds aim to mimic the performance of a benchmark index. In this study, we classify index-based strategies as passive. There has been an extensive debate on the merits of active versus passive management, but few studies have focused on the relationship between active and passive ETFs in addition to mutual funds. This is in part due to the infancy of ETFs. The first ETF listed on a US stock exchange was the Standard & Poor's Depository Receipts (SPDR) introduced in 1993 by State Street Global Advisors. The fund tracks the S&P 500 and trades under the ticker SPY. As of 2019, the SPDR S&P 500 ETF Trust is the world's largest ETF with more assets under management (AUM) than the second and third largest ETFs combined. ETF growth was slow until the mid-2000s. In 2006, ETF growth began to accelerate. Between 2006 and 2019, growth in the number of listed ETFs surpassed 450% (Figure 3). As of year-end 2018, there were 1,988 actively and passively managed ETFs in the US. In combination, US-listed ETFs account for \$3.4 trillion in total net assets and make-up 71% of the \$4.7 trillion held by ETFs worldwide.²

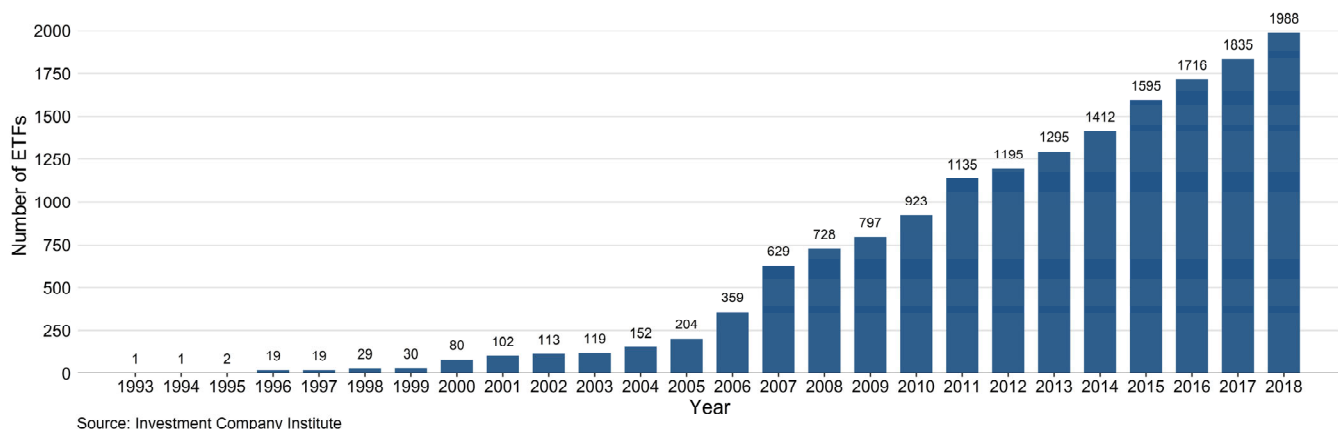


Figure 3. ETF growth in the United States between 1993 and 2019.

With growth comes complexity. Though the first passive ETF was introduced in 1993, it took fifteen years for active ETFs to emerge. In 2008, the Bear Stearns Current Yield Fund (YYY) became the first active US-listed ETF. The fund was doomed from the start as Bear Stearns was collapsing. Also, the fund's strategy was limited to trading short-term debt securities curbing managerial ability to generate high returns.¹¹ Soon after, PowerShares introduced the Active Alpha family of funds, which debuted the first active equity ETFs. Each fund uses complex quantitative trading methodologies to select stocks. Since their introduction, the number of active ETFs has grown by over 200%. As of year-end 2018, there are 289 actively managed ETFs.²

Our study is novel in that we categorize ETFs into active and passive investment strategies. The overwhelming majority of ETFs are index-based; however, actively managed ETFs are becoming increasingly common. The two approaches are fundamentally different and should be treated as such. There has been an extensive debate about the merits of active and passive management

styles. We add to this literature by introducing ETFs into the debate. We compare active funds and passive ETF performance to the historical performance of mutual funds during the same period spanning 1993 to 2019. We break funds into one of four types: mutual fund active, mutual fund passive, ETF active, and ETF passive. Past studies that cast a wide net over all fund strategies and types obscure the costs and benefits of each.

We focus on singling out the relationship between high fees and performance, high fees being an indicator of fund success measured using the dollar percentage gain per share. High fees indicate success because an underperforming fund that charges high fees will have negative cash flows. To attract investment, funds must appeal to investors charging a price that the market will bear. Underperforming funds must lower fees or become insolvent, while outperforming funds will continue to operate compensated by the market accordingly. By breaking down the relationship between performance and management expenses, we can discover a pattern of successful funds and demonstrate the independence of a fund portfolio's underlying assets. Analysts typically spend a lot of time and resources investigating these properties. We streamline this information search by basing evaluations solely on a fund's expense ratio and the performance it predicates.

Literature Review

Our paper contributes to the literature on ETFs and mutual funds as investment vehicles. Guedj and Haung investigate whether an ETF is a more efficient indexing vehicle than an open-ended mutual fund (OEF). Examining organizational structure via an equilibrium model, Guedj and Haung show that OEFs are beneficial to risk-averse investors due to partial insurance against liquidity shocks embedded in the fund's structure. They also find that OEF structures create moral hazard problems inducing excessive trading. Overall, they find that transaction costs to all investors are the same in OEFs and ETFs, though the allocations of these costs differ. Guedj and Haung conclude that investors with similar liquidity needs should be indifferent between the two fund types.³

Poterba and Shoven examine the perception of ETFs as tax-efficient alternatives comparing the pre-tax and post-tax returns of the largest ETF, the SPDR S&P 500 Trust (SPY), with the returns of the largest equity index fund, the Vanguard 500 Index Fund. They find that between 1994 and 2000, the after-tax performance of the two funds are very similar.⁴ Poterba and Shoven do not examine other ETF or mutual fund alternatives. This is of concern because assets held by SPY account for less than 7.8% of the passive ETF marketplace.^A The timing of Poterba and Shoven's study precludes a considerable increase in ETF offerings. We include in our study a representative sampling of funds offering diverse strategies, objectives, and goals.

Poterba and Shoven have found that ETFs may be more attractive for equity investments outside of tax-deferred accounts than IRAs or 401(k)s due to their low rate of taxable distributions and liquidity. Conversely, retirement account investors may find traditional equity mutual funds more attractive due to their unique tax attributes.⁴ Poterba and Shoven do not go beyond SPY in their analysis of ETFs, raising the issue of sample size. Our study builds on their examination of after-tax returns encompassing a larger body of ETFs and traditional funds. Poterba and Shoven account for two passive funds to the exclusion of the broader ETF and conventional fund market, which includes both active and passive strategies. Due to the timing, their study excludes innovations in the ETF and mutual fund market. In 2001, there were 102 ETFs, far fewer than the number available in today's marketplace.

For regulators addressing retirement savings, ETF investment strategies can counteract future uncertainties of social programs, such as Social Security, which are increasingly under the pressure of ballooning deficits. State pension fund managers will also find this work beneficial as they may spare many unnecessary search costs and expenses charged by brokers and hedge funds. Society's capitalized cost of price discovery is at least 10% of the current market cap. It has been estimated that individual returns would be increased by 67 basis points if every investor switched to a passive market portfolio.¹²

Chung et al. study the relationship between hedge fund manager compensation and ability, determining managers receive compensation in relation to their ability.¹³ An explicit pay structure is coupled with an implicit measurement to determine management compensation. A manager's ability to generate future cash flows is indicative of the manager's current ability. Therefore, managers who can generate more cash flow will receive higher levels of compensation because these managers earn fees on larger funds. Because there is a link between performance and compensation, we predict higher mutual fund fees will correlate to increasing levels of fund performance as fees are paid to managers by willing investors eager to compensate high caliber managers for higher returns. Work linking hedge fund and mutual fund performance with compensation has been performed by Holmstrom, Berk and Green, Berk and Van Binsbergen, and Lim, Sensoy, and Weisbach.¹⁴⁻¹⁷

^A As of September 05, 2019. Assets under management change dynamically.

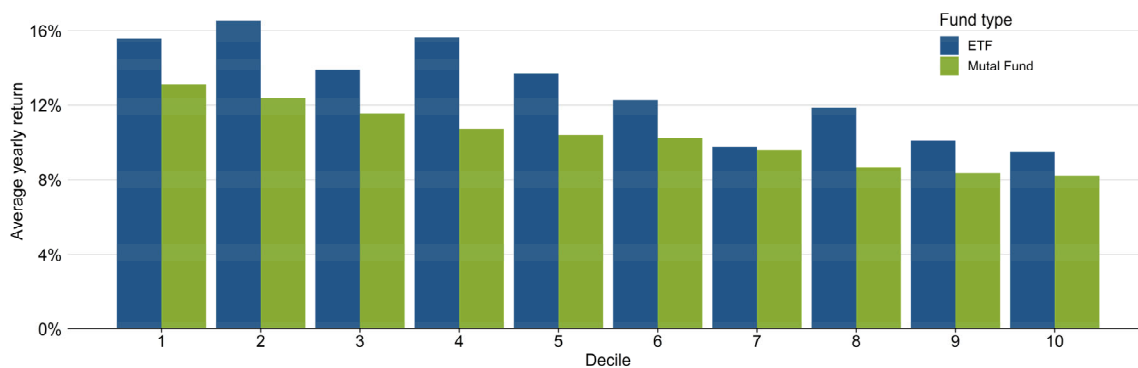


Figure 4. Displays the relationship between expense ratios and returns using the average yearly change in net asset value (NAV).

Mutual fund and ETF expense ratios represent the percentage of a fund's total assets used to operate the fund. These fees are automatically deducted from a fund's assets by managers. We use changes in a fund's net asset value (NAV) to measure an investor's returns. NAVs are computed net of all expenses taken out by fund managers. **Figure 4** shows that both mutual fund and ETF performance decrease as their expense ratios increase. The best performing funds have low-cost structures, while the high-cost funds exhibit the worst performance. This is likely attributable to the fact that low-cost funds generally use index-based strategies that require less decision making and turnover (percentage of holdings that change over a year) lowering costs.

Figure 5 paints a more detailed picture. Again, we display the relationship between fund expense ratios and performance. However, this time we categorize by fund type and investment strategy. Though these results are far from conclusive, **Figure 5** offers three important insights that lead us to believe that funds should be categorized by strategy instead of lumping them together by fund type as past studies have done.

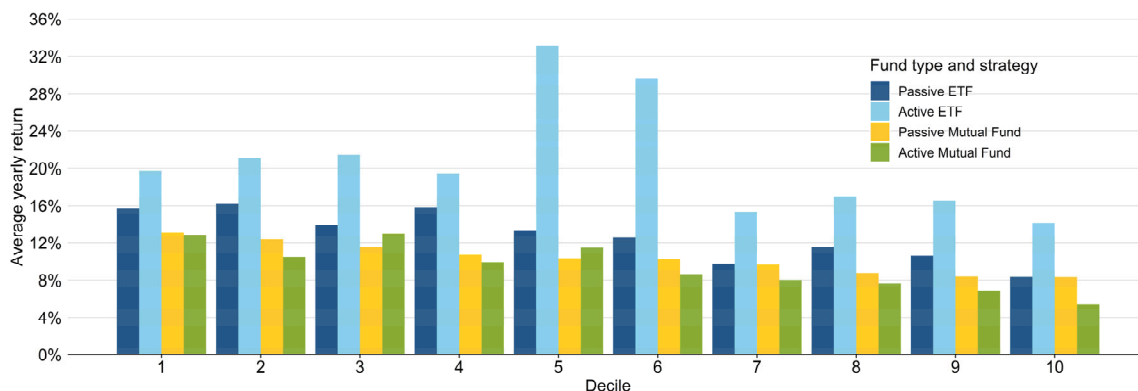


Figure 5. Expense ratio on fund performance by fund type and strategy.

First, active ETFs' change in NAV increases as the fund's expense ratio increases. Performance peaks at the fifth decile, but performance is still high at the sixth decile before dropping by nearly 50%. Passive ETFs do not appear to penalize for higher expense ratios until the fifth decile. From this, we gather that a positive relationship will only hold up to a certain threshold that differs by fund type and strategy. For funds with higher expense ratios, managers are unable to overcome the higher costs with improved performance.

Second, passive funds provide consistent returns. Passive fund performance varies less when compared to active funds. For mutual funds, excluding deciles three and five, passive strategies outperform active strategies. Passive ETFs outperform both mutual fund strategies across all decile groups. This, in conjunction with the performance of active ETFs, suggests ETFs may have an advantage over mutual funds. Although in the seventh and tenth decile, passive ETFs only slightly outperform passive mutual funds.

Third, eventually, all funds succumb to the behavior displayed in **Figure 4**. Higher expense ratios lead to lower returns. This is particularly true for index-based funds whose performance does not change as expense ratios increase. Because passive funds mimic major-indices, the composition of funds held by tenth decile funds is the same as those held by lesser decile funds. It

follows that index-based funds with higher fees will always perform worse than their lower fee equivalent. Investment in these funds contradicts rational expectations theory. However, behavioral economists have sought to explain why investors purchase high-fee index funds in lieu of a low-cost equivalent, finding that susceptible investors overweight funds' past performance and that this behavior decreases with financial literacy.¹⁸

Overall, we find that ETFs outperform mutual funds across investment strategies and that passive investment strategies perform well relative to active strategies. These findings are far from conclusive. Pástor, Stambaugh, and Taylor have found decreasing returns to scale in the active mutual fund industry. Over time they find that fund performance deteriorates at both the fund and industry level. They also find that young funds tend to outperform older funds. It is possible that we may see the same thing happen to active ETFs. As the active ETF market becomes more crowded, average performance will likely decrease over time.¹⁹ It is also possible that outliers may skew performance though we find similar results after winsorizing the top and bottom 10% of funds by expense ratio and NAV. Active ETFs are particularly susceptible due to the relatively small number of active ETFs relative to other funds.

We observe fierce competition in the industry, with mutual fund fees steadily declining since 2003 (**Figure 6**).

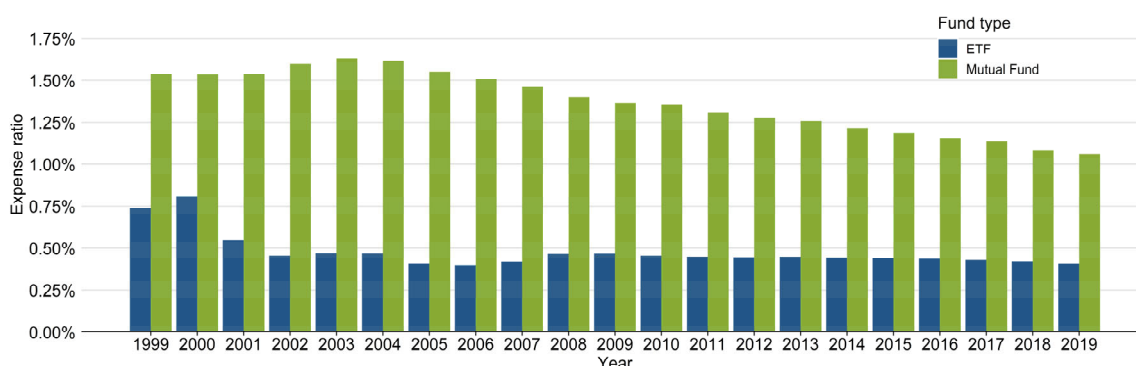


Figure 6. Expense Ratio on Fund Performance by Fund Type.

As Robinhood has transformed the brokerage industry, ETFs are having a similar effect on mutual funds. Some of the largest mutual fund companies offer ETF versions of their most popular funds allowing investors to avoid front-end loads and initial minimum investments. For example, Vanguard offers its most popular fund, the Vanguard 500 Index Fund, as the Vanguard S&P 500 ETF (VOO).

Mutual funds emulating indices have experienced inflows, while outflows from active mutual funds have accelerated (**Figure 1**). We predict, due to competition, increased transparency, and lower barriers to entry, that mutual fund and ETF management fees will be driven toward convergence. Hortaçsu and Syverson discuss the dispersion of fund fees amongst S&P 500 index funds finding that search costs largely drive differences in fees, product differentiation via marketing, and competition.²⁰ All three mechanisms have changed since the release of their paper in 2004. As seen in **Figure 6**, we find that mutual fund fees peaked in 2003 a year before the release of their findings.

METHODS AND PROCEDURES

Data and Variable Definitions

We use the Center for Research in Security Prices (CRSP) Survivor Bias-Free Mutual Fund Database for our analysis.²¹ The data tracks returns at a daily frequency and spans 20 years, 1998 to 2019.^B CRSP began tracking daily fund returns on September 2, 1998. The first observation in our dataset is from October 1, 1998. For periods prior, fund data is monthly, so we choose to focus on the later period. We measure fund performance using the change in net asset value. NAVs include reinvested dividends and subtracts management expense and 12b-1 fees.¹⁰

$$\text{Net Asset Value} = \frac{\text{Assets} - \text{Liabilities}}{\text{Shares Outstanding}}$$

Fee tables can be found in a fund's prospectus, which is an important document detailing a fund's investment objectives, risk, fees, expenses, and other pertinent information. The fees we focus on in our analysis are "Annual Fund Operating Expense." As defined by the SEC, this section lists management fees, distribution and service (12b-1) fees, other expenses, and total annual

^B October 1, 1998 – July 28, 2019

fund operating expenses. Expenses under the heading “Shareholder Fees” are excluded from a fund’s expense ratio. These are fees such as sales loads, redemption, purchase, exchange, and account fees. These costs are excluded from the analysis because they depend more on investors’ actions than on managers’ choices. Also, because ETFs are traded on exchanges, to acquire these assets, investors must go through a stockbroker. Brokerage fees are classified as transaction costs and are excluded from our analysis. Long term, these fees are marginal and have no relationship to ETF performance. Shareholder fees vary by fund. The SEC does not impose specific limits on fees that a mutual fund may charge. The SEC does limit redemption fees to 2% in most situations; however, FINRA is mostly in charge of imposing limits on mutual fund fees.¹⁰

Turnover is a proxy for management expenses located within the “Annual Fund Operating Expense.” Turnover tracks the number of trades made by a fund. The more trades, the higher the expense to the investor, as brokerage fees and commissions must be paid to intermediaries. Portfolio turnover is calculated per period by deducting from newly purchased securities the number of securities sold and dividing this number by the fund’s NAV.

$$\text{Turnover} = \frac{\text{Newly Purchased Securities} - \text{Securities Sold}}{\text{Net Asset Value}}$$

ETFs reflect management costs in the expense ratio and are not subject to front or back end loads. Actively managed funds tend to incur more of these expenses as managers trade in and out of positions as the market fluctuates. Passive funds are designed to do the opposite, minimizing trading costs through strategies such as indexing.¹⁵ The minimization of turnover helps lead to lower expense ratios and other expenses. Fees and expenses vary by fund. A fund with high costs may perform better than a low-cost fund and still generate a smaller return. Small differences in fees translate into significant differences in returns over time.

Poor performing funds are more likely to exit the market creating an upward bias in performance as only the survivors remain.²²⁻²⁵ By including active and inactive funds, our dataset controls for this. A weakness of the CRSP dataset is that it does not control for incubation bias. Incubation bias occurs when firms utilize a series of privately held funds from which they handpick the best performers to offer publicly. During the incubation phase, funds accrue performance histories that are used in the fund’s prospectus to increase the fund’s marketability. CRSP reports performance from the incubation period, so the selection of incubation survivors leads to an upward bias. Only the past histories of strong funds are included in the dataset. Richard Evans has shown that this bias can be eliminated by excluding fund data from the incubation period.²⁶ We do this by excluding fund data from before the public start date.

We limit our study to domestic equity mutual funds with an average of 90% or more of their assets in common stocks over their lifespan. Using the CRSP Style Codes, we exclude fixed-income funds and funds classified as “other” from our analysis. CRSP Style Codes consists of up to four characters that, from left to right, represent an increasing level of granularity. We restrict the first character to ‘E’ for equity funds. Fixed income and money market funds are excluded from our analysis because we want to measure a fund manager’s ability to pick stocks. We create flag variables for fund characteristics. We differentiate ourselves from other studies by focusing on fund type and investment strategy. Investment strategy consists of active and passive management styles and fund type differentiates ETFs from mutual funds.

Passively managed funds can be identified using CRSP’s index fund flag; however, this data was not tracked until 2003. Other variables such as fund ticker, objective code, or ETF flag are not uniformly reported across each fund’s identification number. We find that CRSP’s summary data for a variety of reasons is missing or elects not to fill in certain fund characteristics. To correct for inconsistencies, we use a backward matching algorithm allowing us to fill in missing observations that go unreported for a portion of a fund’s history. Using data from First Bridge, we can match Active ETFs by ticker.²⁷ For example, by applying our matching algorithm to the index fund flag, we can identify index funds prior to 2003 if they existed after. To catch funds that existed before 2003, but not after, we use regular expressions to perform textual analysis on fund names to identify passive investment styles. For example, the terms “S&P,” “Dow Jones,” “Russell,” “Wilshire,” or “MSCI” identify funds whose strategies are index-based. Similarly, ETFs can be flagged using backward matching on CRSP’s ETF flags and then double-checked using textual analysis. We try a variety of identification schemes to cross-validate our results. Our most reliable indicator of fund type is the NASDAQ ticker. Tickers are widely issued across the entire sample and mutual funds can be identified using the fifth-letter, which is an X indicating that the fund is a Mutual Fund Quotation Service Instrument (MFQS).²⁸

Funds with negative or zero expense ratios are removed from our sample. We consider all three of CRSP’s index fund classifications as passive: pure index, index-based, and index enhanced. This creative approach allows us to widen the scope of our study to include four groups: active ETFs, passive ETFs, active mutual funds, and passive mutual funds. Expense ratios are reported as of the most recently completed fiscal year. We then allocate the costs over trading days to compute the average daily expense ratio for each fund. A fund’s total expense ratio (TER) is calculated by multiplying its expense ratio by the fund’s total assets. Expense ratios account for size differences amongst funds allowing us to make comparisons across firms. In absolute

terms, a fund's total cost is attributable to its size. We further classify our data by computing fund age and creating an indicator for survivorship.

Analysis

The dependent variable in our model is returns. Daily and monthly returns are calculated as the change in NAV, including reinvested dividends. NAVs are net of all management expenses and 12b-1 fees. Front and rear load fees are excluded. Each day a total adjustment factor is calculated by CRSP. The adjustment factor begins with a value of one for each day and is then modified depending on the types of dividends found for the fund that day. Modifications include distributions, reinvestments, and stock splits. We use the daily adjustment factor calculation as specified by the *CRSP Survivor-Bias-Free US Mutual Fund Guide*.²⁹

The independent variable of interest is management expense. For this variable, we use the net expense ratio provided by the *CRSP Survivor-Bias-Free US Mutual Fund Database*. The annual net expense ratio, often referred to as the "audited expense ratio," is drawn from a fund's audited annual report and is calculated as the percentage of a fund's assets paid during a fiscal year for operating expenses and management fees. Fees include accounting, administrator, advisor, auditor, board of directors, custodial, distribution (12b-1), legal, organizational, professional, registration, shareholder reporting, subadvisor, and transfer agency expenses. Waivers and reimbursements are included in the net expense ratio, while brokerage costs and investor sales charges are excluded. Annual expense ratios may differ from the prospectus expense ratios, as the annual ratio reflects a fund's actual expense, while the prospectus ratio is an estimate of expenses for the current period.³⁰

We also include in our analysis control variables to capture the influence of external factors that can affect returns independent of managerial skill. If the economy is in an expansion, a manager may appear to possess more skill though increases in asset prices are attributable to growth in the broader economy. Interactions among control variables are permissible if all the interacting terms are included in the model.³¹ Coming up with a comprehensive list of outside factors influencing asset prices is not realistic. However, there is a method for calculating risk-adjusted returns that spares us from introducing thousands of control variables.

In 1993, Eugene Fama and Kenneth French identified three factors that explain up to 95% of a diversified portfolio's returns.³² These factors are firm size, book-to-market value, and excess returns on the market. These factors are commonly referred to as "Fama-French Factors" or the "Three-Factor Model".³³ In 1997, Mark Carhart extended the model by adding momentum as a fourth factor.¹⁹ Thus, we use risk-adjusted returns to evaluate a manager's performance. We follow standard conventions small minus big (SMB), high minus low (HML), risk-free rate using the one-month Treasury Bill rate (RF), excess return on the market (MKTRF), and momentum (UMD).²⁴ The first three factors are constructed using six value-weight portfolios formed on size and book to market equity ratio. The portfolios are classified as small value, small neutral, small growth, big value, big neutral, and big growth. This data can be found on Kenneth French's website.³³ We use a compilation of these values provided by Wharton Research Data Service (WRDS).³⁴ Below, we provide a high-level overview of our model. For more detail on the construction of the six portfolios, see Fama and French, "*Common Risk Factors in the Returns of Stocks and Bonds*."³³

$$\text{Excess Portfolio Returns}_{i,t} = \alpha + \beta_1 \text{MKTRF}_{i,t} + \beta_2 \text{SMB}_{i,t} + \beta_3 \text{HML}_{i,t} + \beta_4 \text{UMD}_{i,t} + \varepsilon_{i,t}$$

Excess portfolio returns or "risk-adjusted returns" are calculated as a fund's return minus RF. MKTRF is based on the value-weighted return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ that have a CRSP share code of 10 or 11 at the beginning of month t , good shares and price data at the beginning of t , and good return data for the entire month. From this value, we subtract the return on riskless assets represented using the one-month Treasury bill (i.e. the risk-free rate, RF) from Ibbotson Associates.³⁵ We refer to MKTRF and the market interchangeably as MKTRF. MKTRF is intended to be representative of the overall market.

$$\begin{aligned} \text{Excess Portfolio Return}_{i,t} &= \text{Fund Return}_{i,t} - \text{RF}_{i,t} \\ \text{MKTRF}_{i,t} &= \text{Market Return}_{i,t} - \text{RF}_{i,t} \end{aligned}$$

Each factor has a coefficient or loading. The market portfolio loading, β_1 , captures systematic risk by measuring the extent to which a fund's returns are attributable to changes in the broader market. A beta greater than one, e.g. 1.5, means that for every 1 unit increase in the MKTRF the fund's returns will increase by 1.5 units. Conversely, a one-unit decrease would equate to a 1.5 unit decrease in the fund's returns. Thus, the fund is more sensitive to changes than the broader market and carries more systemic risk. The average firm has a beta of 1.0 meaning it moves with the market. A beta less than one means that a fund carries less systematic risk than average and a beta of zero equates to no systematic risk, i.e. the fund's performance is unaffected by the market. A negative beta means that the expected return of a fund moves in a direction opposite the expected return of the market portfolio. Most equity-only funds have betas close to 1.0.³⁶ Intuitively, higher betas equate to higher than expected returns as

investors are taking on more risk. This intuition underlies the capital asset pricing model (CAPM) which measures a fund's risk relative to the market. In our model, we subtract RF from both sides giving us excess portfolio return on the left-hand side.

$$CAPM: Fund\ Return_{i,t} = RF_{i,t} + \beta_1 MKTRF_{i,t} + \varepsilon_{i,t}$$

$$Fund\ Return_{i,t} - RF_{i,t} = \beta_1 MKTRF_{i,t} + \varepsilon_{i,t}$$

We estimate a single-factor market model as:

$$Excess\ Portfolio\ Return_{i,t} = \alpha + \varphi_1 Expense\ Ratio_{i,t} + \beta_1 MKTRF_{i,t} + \varepsilon_{i,t} \quad \text{Equation 1.}$$

In **Equation 1**, the intercept alpha (α) is interpreted as the value added by a fund manager. After accounting for the expense ratio, alpha tells us how well the manager is performing relative to the market. An alpha of zero suggests a manager is capturing market returns perfectly, while a positive alpha suggests that the manager is delivering returns higher than the market's expected return. To evaluate a fund manager's ability, we subtract the portfolio's actual return from the benchmark return. If alpha is negative, the manager is underperforming the market.

$$\alpha_{i,t} = Excess\ Portfolio\ Return_{i,t} - MKTRF_{i,t}$$

We estimate Fama-French 3-Factor and Carhart 4-Factor models:

$$Fund\ Return_{i,t} - RF_{i,t} = \alpha + \varphi_1 Expense\ Ratio_{i,t} + \beta_1 MKTRF_{i,t} + \beta_2 SMB_{i,t} + \beta_3 HML_{i,t} + \varepsilon_{i,t} \quad \text{Equation 2.}$$

$$Fund\ Return_{i,t} - RF_{i,t} = \alpha + \varphi_1 Expense\ Ratio_{i,t} + \beta_1 MKTRF_{i,t} + \beta_2 SMB_{i,t} + \beta_3 HML_{i,t} + \beta_4 UMD_{i,t} + \varepsilon_{i,t} \quad \text{Equation 3.}$$

SMB is the average return on the three small portfolios minus the average return on the three big portfolios. It accounts for the outperformance of small companies relative to large ones. Its loading, β_2 , indicates the extent to which excess returns are attributable to its size. An SMB loading of zero signifies that a fund is primarily composed of large-cap stocks, while a value greater than 0.5 indicates a fund is tilted toward small-cap stocks. A statistically significant coefficient with a high R-value indicates that the excess return of a fund may be attributable to its large size.³⁶

HML is the average return on the two value portfolios minus the returns on the two growth portfolios. It encapsulates the historical excess returns to value stocks (high book-to-price ratio) over growth stocks (low book-to-price ratio). Its loading, β_3 , represents the portion of excess returns that are attributable to its book-to-market equity value. A loading of zero implies a growth portfolio, while a loading of more than 0.3 implies the fund is tilted toward value stocks.

The fourth factor, UMD, was added by Mark Carhart and is calculated as the difference of the equal-weighted average of the lowest-performing funds and the equal-weighted average of the highest performing funds. Carhart found that returns for the previous 12-months can have predictive value on the future performance of that stock. A positive and statistically significant coefficient (β_4) equates to positive momentum.²⁴ Epsilon represents firm-specific risk. This is the part of the return that is unexplained by the factors.

We then add in fund controls:

$$Excess\ Portfolio\ Return_{i,t} = \alpha + \varphi_1 Expense\ Ratio_{i,t} + \gamma_2 4\ Factors_{i,t} + \gamma_3 Fund\ Controls + \varepsilon_{i,t} \quad \text{Equation 4.}$$

We then compare these equations to evaluate the relationship between performance and fees. In **Equation 1**, we use a single factor model upon which we expand to three and four-factor models, **Equations 2 and 3**. Eugene Fama and Kenneth French showed that the average explanatory value (r-squared) went from explaining 70% to 95% when moving from CAPM to three factors.³² Carhart later demonstrated that past performance could be used to explain future performance introducing momentum as a fourth factor.

Since Carhart's findings, other factors have been used; however, size, value, and momentum are universally accepted in the finance literature. In **Equation 4**, we add fund controls for daily net asset value (DNAV), total net asset value (TNAV), and turnover. **Equations 2, 3, and 4** include our explanatory variable, the expense ratio. This allows us to compare alphas between equation parings. If a fund's performance is attributable solely to factors, then the manager has not demonstrated an ability to add value. High performing managers should enhance performance through the selection of stocks that outperform expected returns. We proxy a manager's value added by their ability to charge a fee for their stock-picking ability. The more a manager charges, the

more value they should be able to demonstrate through superior returns. We expect high performing managers to have positive and statistically significant ϕ_1 values.

If a fund's alpha improves positively with the manager's compensation, we conclude that the improvement is attributable to the manager's ability to select stocks. Ordinary least of squares regressions are used to assess the relationship between independent and dependent variables. We use robust standard errors and include year and time fixed effects to account for differences across funds and years.

RESULTS

In **Table 2**, we present descriptive statistics of our data.

Group	Observations	Median Year	Median Expense Ratio	Average Expense Ratio	Median Annualized Return	Average Annualized Return	Median NAV (millions)	Median Turnover Ratio	Average Fund Age (years)
By Category:									
ETF	1,844,683	2014	0.46%	0.44%	19.97%	12.88%	39.37	0.22	12.36
Mutual Fund	39,442,121	2010	1.29%	1.36%	20.21%	10.32%	14.92	0.61	15.99
By Fund Type:									
Active	1,179,273	2011	1.24%	1.32%	18.52%	9.67%	13.79	0.58	13.31
Passive	40,107,531	2010	1.25%	1.22%	20.24%	10.46%	15.39	0.59	15.90
By Current Status:									
Live Funds	24,753,127	2013	1.12%	1.16%	21.60%	13.00%	17.42	0.51	18.04
Dead Funds	16,533,677	2006	1.50%	1.56%	17.61%	6.62%	13.05	0.72	12.51
All funds	41,286,804	2010	1.25%	1.32%	20.19%	10.44%	15.33	0.59	15.82

Table 2. Time series of cross-sectional average and median annual attributes from September 1999 to June 2019.

Our dataset consists of 41,286,804 daily fund observations, spanning 1998 to 2019. In 1998, CRSP began reporting daily returns data. CRSP's adaptation coincides with the earliest observations retained in our data set. We perform transformations on 21 variables and have approximately 20,000 unique CRSP fund numbers that map to approximately 46,000 unique funds in our dataset.^{37, 38} Mutual funds make up the majority of our observations as ETFs are still in their nascent stage. Median returns for ETFs and mutual funds are similar at approximately 20%; however, average yearly returns are higher for ETFs at approximately 12.88% while mutual funds' average yearly returns are 10.32%. The difference, 2.56%, is markedly large, especially for long-term investors. Over 10 years, this equates to approximately a 20% difference in cumulative return. Expense ratios are much lower for ETFs than they are for mutual funds, further widening the gap between ETF and mutual fund returns. Both the average and median expense ratios for mutual funds are nearly three times larger than the expense ratios for ETFs. (**Figure 6**). Mutual fund turnover is approximately three times that of ETFs. Turnover may be a key factor driving the difference in expense ratios between fund types. The average age of mutual funds in our sample exceeds ETFs by approximately four years.

Passively managed funds outnumber actively managed funds in our sample. We find that fund characteristics across groups are similar though passive ETFs have higher average and median returns. Passive fund turnover is nearly identical to that of active funds. On average passive funds have lower expense ratios, which is in line with expectations. Approximately 60% of the funds in our sample are actively managed, while approximately 40% are no longer active. Dead funds, on average, have lower returns and higher expense ratios than live funds. Live funds are almost double the average yearly return of dead funds. The average lifespan of a dead fund is almost six years shorter than the lifespan of a live fund. This suggests that underperforming funds exit the market quickly. High turnover is associated with these funds, which may contribute to their higher costs.

In **Table 3**, we provide summary statistics for active and passive ETFs as well as active and passive mutual funds.

Fund Type by Strategy	Distinct Funds	Median Year	Median Expense Ratio	Average Expense Ratio	Median Annualized Return	Average Annualized Return	Median NAV (millions)	Median Turnover Ratio	Average Fund Age (years)	Live / Dead Funds
ETF										
Active	51	2017	0.75%	0.71%	22.69%	20.76%	31.20	0.58	4.07	43 / 8
Passive	1,165	2014	0.45%	0.44%	19.92%	12.78%	39.60	0.22	12.47	890 / 275
Mutual Fund										
Active	907	2011	1.25%	1.33%	18.43%	9.43%	13.63	0.58	13.50	392 / 515
Passive	17,932	2010	1.29%	1.36%	20.26%	10.35%	14.98	0.61	16.06	8,106 / 9,826

Table 3. Time series of cross-sectional average and median annual attributes by fund type and strategy September 1999 to June 2019.

Most funds in our sample are passive. We have 24,223 active ETFs and 1,820,460 daily passive ETF observations. For mutual funds, there are 1,155,050 active daily fund observations and 38,287,071 daily passive fund observations. These observations map to 51 active and 1,165 passive ETFs in addition to 907 active and 17,932 passive mutual funds that meet our criteria as described in *Data and Variables Definitions*.

Revisiting **Figure 5**, we must be careful in our assessment of active ETFs as its strong performance is drawn from a small sample of funds. Even so, these results should not be dismissed. Active ETFs are in their nascent phase. The first active equity ETFs, the PowerShares Active Alpha funds, commenced operations in 2008 and did not have a full calendar year of performance history until 2009. Whereas, State Street introduced the SPDR S&P 500 Trust fund in 1993. This is evidenced by the average age of active ETFs; the average age of active ETFs is slightly over four years, while passive ETFs average age is 12.5 years. The first active ETF was introduced in 2008 by Bear Stearns; however, we exclude it from our analysis as it was not an equity-based fund. **Table 3** supports our earlier findings that active ETFs have the highest returns. Passive ETFs, passive mutual funds, and active mutual funds follow in the stated order. This aligns with **Figure 5**, as we see that both ETFs categories outperform mutual funds on average. This is also shown in **Table 2**'s comparison of ETFs and mutual funds.

In **Figure 7**, we plot yearly returns by fund type. Again we find that, on average, ETFs outperform mutual funds. We also note that both ETFs and mutual funds experience negative market shocks during the early 2000's bursting of the dot-com bubble and the 2008 Great Recession. Mutual fund returns were slightly more negative during these periods.

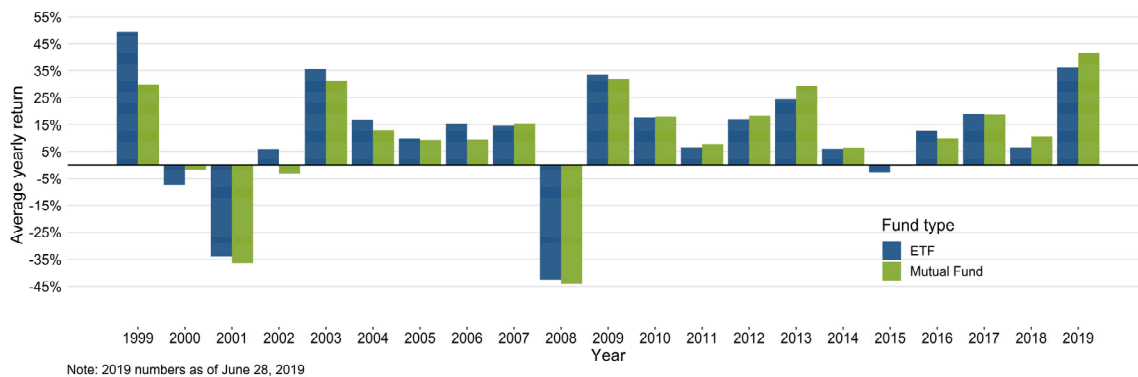


Figure 7. Yearly Return by Fund Type.

Figure 8 provides a more detailed picture. Here, we separate funds by type and strategy. We add the SPDR S&P 500 ETF (SPY) as a benchmark for performance. Active ETF observations begin with their introduction in 2009.

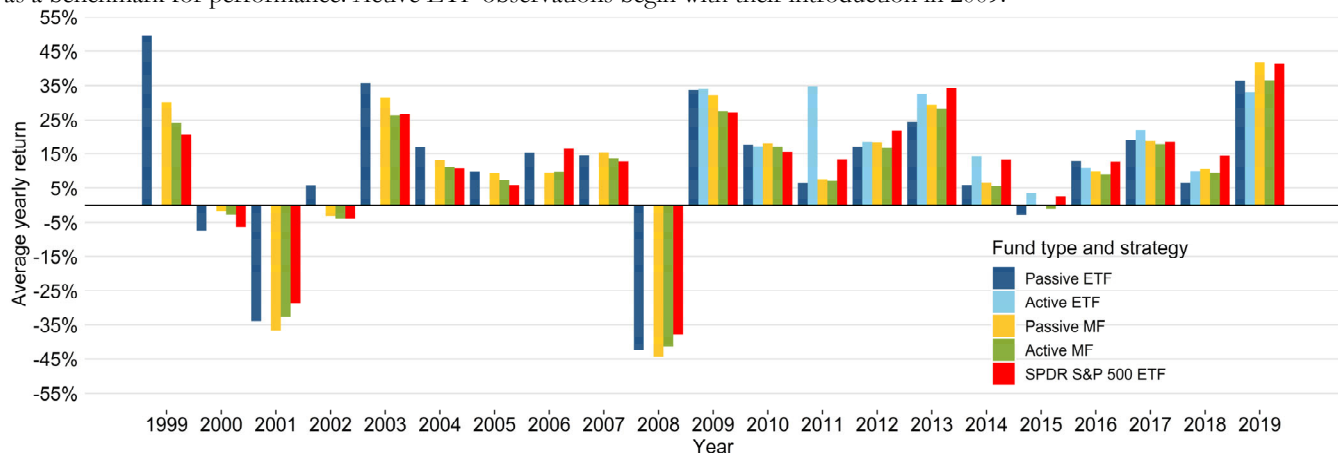


Figure 8. Yearly return by fund type & strategy.

In **Figure 8**, we see that ETFs outperform mutual funds on average. Our benchmark fund is SPY, the first ETF introduced in 1993. From January of 1999 to June 2019, SPY had an average return of 10.74%, exceeding the returns of active and passive mutual funds. However, SPY's median return of 16.88% is less than the median returns of all funds, meaning that over 50% of funds in our sample outperformed SPY.

In **Table 4**, we look at performance during recessionary periods. Using the National Bureau of Economic Research's (NBER) business cycle expansion and contractions dates we restrict our sample to dates between March 2001 and November 2001 and between December 2007 and June 2009.³⁹

Fund Type by Strategy	Distinct Funds	Average Annualized Return
ETF		
Active	2	24.68%
Passive	392	-19.01%
Mutual Fund		
Active	416	-23.33%
Passive	11,662	-24.30%
All Funds		
Active	418	-23.22%
Passive	12,050	-24.14%

Table 4. Fund performance during recessionary periods.

During recessionary periods ETFs outperform mutual funds, while active funds come out ahead of passively managed funds though not by much. Having a predetermined strategy, passive ETFs do not adjust in bear markets, whereas active fund managers may adjust their holdings to limit losses. During these periods, we expect actively managed funds to perform well relative to passively managed funds. **Table 4** suggests that this explanation may be plausible though active funds sustain losses that are only slightly less than passive funds. It may be advantageous to invest in an actively managed firm during economic downturns. In 2009, the newly introduced PowerShares Active Alpha Q fund and the PowerShares Active Mega-Cap Fund performed well during the recessionary period; however, we cannot draw strong conclusions from such a limited sample. These funds enter our sample in January of 2009.

In **Tables 5** and **6**, we show our regression estimates for **Equations 1** through **4** across fund types and strategies. **Table 5** shows active and passive ETFs, while **Table 6** presents active and passive mutual funds.

For passive ETFs, ϕ_1 is positive and statistically significant for **Equations 1, 2, and 3**. **Equations 1** and **2** equate to our market and Fama French regressions in **Table 5**. For both, we estimate expense ratio coefficients that are statistically significant at the 1% confidence level. From **Equation 3**, we estimate Carhart's four-factor model finding an expense ratio coefficient that is significant at the 5% level. Adding control variables to the Carhart model as in **Equation 4** we find no significance for ϕ_1 . For active ETFs, ϕ_1 is negative, but only statistically significant for the market regression (5% confidence level).

Multiplying this by the median expense ratio for passive ETFs found in **Table 3**, 0.44%, we estimate that the median passive ETFs' expense ratio contributes a daily return gain of 0.015%. The three-factor model coefficient rounds to the same gain while our four-factor model produces a daily gain of 0.013%. Using the number of trading days in a year, we find that passive ETFs gain between 3.29% and 3.80% annually.

The expense ratio coefficients for active ETFs are negative; however, the market model is the only regression with a statistically significant coefficient. The market model expense ratio coefficient, -0.292, is significant at 5%. Using the median expense ratio for active ETFs, 0.71%, we find that the median actively managed ETF has a daily reduction in NAV of 0.207% or 52.45% annually.

For active mutual funds, ϕ_1 is positive and statistically significant across all regressions. The coefficient is consistently 0.0259 for all regressions except for when we add controls to the four-factor model. With controls, the coefficient drops to 0.0247. Using the median expense ratio for active funds, we estimate the yearly gains for investing in active mutual funds to be between 7.83% and 8.19% annually. For passive funds, ϕ_1 is not statistically significant in any of our models. Even so, the coefficient near zero for all the regressions, indicating that even if significant, the expense ratio would have no effect.

	Exchange-Traded Funds							
	Market		Fama French		Carhart		Carhart Controls	
	(1) Active	(2) Passive	(3) Active	(4) Passive	(5) Active	(6) Passive	(7) Active	(8) Passive
MKTRF	0.878*** (25.40)	0.906*** (67.24)	0.865*** (28.69)	0.882*** (70.57)	0.876*** (29.60)	0.871*** (70.24)	0.866*** (29.69)	0.869*** (69.87)
HML			0.0507 (1.15)	0.101*** (6.93)	0.0622 (1.23)	0.0567*** (4.04)	0.0623 (1.23)	0.0564*** (4.03)
SMB			0.0932 (1.75)	0.106*** (8.08)	0.0981 (1.89)	0.102*** (7.76)	0.0985 (1.90)	0.100*** (7.67)
UMD					0.0179 (0.88)	- (-13.08)	0.0174 (0.84)	-0.0666*** (-13.19)
EXP_RATIO	-0.292* (-2.18)	0.0332** (2.75)	-0.273 (-1.86)	0.0319** (2.67)	-0.268 (-183)	0.0303* (2.55)	-0.169 (-1.00)	0.0109 (0.74)
DNAV							0.0000552 (0.81)	0.0000870*** (5.24)
NAV_LATEST							-0.0000199 (-0.39)	- (-4.92)
TNA_LATEST							0.000000283 (-1.11)	3.28E-10 (0.43)
TURN_RATIO							-0.00000159 (-0.13)	0.0000587 (1.76)
Constant	0.000572 (0.35)	-0.000479 (-1.55)	0.000409 (0.24)	-0.000284 (-0.94)	0.000420 (0.25)	-0.000165 (-0.54)	-0.000858 (-0.53)	-0.000224 (-0.76)
Observations	24223	1820460	24223	1820460	24223	1820460	24223	1820460
R-squared	0.486	0.527	0.488	0.530	0.488	0.532	0.488	0.532
Fund Fixed Effects	yes	yes	yes	yes	yes	yes	yes	yes
Year Fixed Effect	yes	yes	yes	yes	yes	yes	yes	yes
Robust standard errors	yes	yes	yes	yes	yes	yes	yes	yes

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5. ETF expense ratio on fund performance.

DISCUSSION

We show that returns for passive mutual funds and ETFs decrease as their expense ratios enter the upper decile range for their peer group (**Figure 4**). Breaking this finding down into management strategies, active and passive, we find that this result is not evenly distributed across subcategories. Active funds exhibit the opposite behavior, with returns increasing into the fifth decile. Passive ETF returns decrease in the first four decile groups (**Figure 5**). This behavior suggests that studies of ETF and mutual fund performance should account for the investment strategies of sampled funds. Failing to account for fund strategy may skew a study's findings based on the composition of the funds in the study. For example, a study of randomly selected ETFs that are tilted toward actively managed funds will result in higher returns than a randomly selected sample of ETFs that happens to be tilted toward passively managed funds, holding everything else constant.

	Mutual Funds							
	Market		Fama French		Carhart		Carhart Controls	
	(9) Active	(10) Passive	(11) Active	(12) Passive	(13) Active	(14) Passive	(15) Active	(16) Passive
MKTRF	0.960*** (147.06)	0.975*** (662.69)	0.949*** (152.45)	0.958*** (643.29)	0.950*** (153.60)	0.960*** (665.26)	0.950*** (153.28)	0.958*** (656.76)
HML			0.0349*** (3.70)	0.0558*** (19.40)	0.0377*** (3.97)	0.0591*** (20.49)	0.0375*** (3.95)	0.0583*** (20.27)
SMB			0.109*** (7.57)	0.170*** (59.28)	0.108*** (7.55)	0.170*** (-59.1)	0.108*** (7.54)	0.168*** (58.79)
UMD					0.00531 (1.25)	0.00597*** (6.88)	0.00521 (1.23)	0.00543*** (6.21)
EXP_RATIO	0.0259*** (4.22)	-0.000445 (-0.04)	0.0259*** (4.26)	-0.000303 (-0.03)	0.0259*** (4.26)	-0.000312 (-0.03)	0.0247*** (4.12)	0.000823 (0.06)
DNAV							0.0000428 (1.19)	0.000161*** (7.31)
NAV_LATEST							-0.0000418 (-1.19)	-0.000155*** (-7.30)
TNA_LATEST							-3.94E-08 (-1.83)	-7.29e-09*** (-4.08)
TURN_RATIO							0.00000563 (0.60)	0.0000134*** (3.80)
Constant	0.00219*** (15.79)	-0.0000723 (-0.12)	0.00211*** (15.92)	-5.03E-05 (-0.08)	0.00211*** (15.95)	-0.0000583 (-0.09)	0.00202*** (13.18)	-0.000392 (-0.65)
Observations	1155050	38287071	1155050	38287071	1155050	38287071	1155050	38287071
R-squared	0.838	0.724	0.841	0.729	0.841	0.729	0.841	0.730
Fund Fixed Effects	yes	yes	yes	yes	yes	yes	yes	yes
Year Fixed Effect	yes	yes	yes	yes	yes	yes	yes	yes
Robust standard errors	yes	yes	yes	yes	yes	yes	yes	yes

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6. Mutual fund expense ratio on fund performance.

In **Figure 5**, we show that active ETFs have higher returns across all decile groups which is confirmed in **Table 3**. We find that both average and median yearly returns are highest for active ETFs. Though we find that the performance of active ETFs has been stellar since their inception, the sample size is too small for us to conclude that their relative outperformance will continue. Since 2013, we have seen some signs of decline in their relative returns. The SPDR S&P 500 ETF has produced similar returns during this period. In addition, Pástor, Stambaugh, and Taylor have shown that upward trends in skill in the mutual fund industry coincide with industry growth. As AUM increases, a fund manager's ability to outperform passive benchmarks declines.¹⁹ Young firms tend to outperform older funds, and as more firms enter the space competition will increase eroding performance. It is possible that the trajectory of the active ETF industry may follow a similar path.

We use the median expense ratio for each fund type found in **Table 3** and multiply it by the coefficient φ_1 to estimate the median increase or decrease in daily returns dependent on fund strategy. We annualize this value to get the median change in returns for active and passive strategies. For passive ETFs, we estimate an increase between 3.29% and 3.80% annually. For active ETFs, the coefficient is large and negative suggesting a decrease in fund returns as the expense ratio increases. The market model is the only regression with a statistically significant expense ratio coefficient (5% confidence level). The magnitude (52.45%) of the coefficient is abnormally large relative to our other regressions. This may be attributable to our limited number of active ETF observations. Even so, the coefficient suggests that there is a negative relationship between active ETF fees and returns.

So, which is better, active or passive management? In aggregate, passive funds outperform active funds. **Table 2** shows that passive funds have higher average and median returns. Their median expense ratios are nearly identical; however, passive funds

have lower average expense ratios. **Figure 1** shows inflows into index-based ETFs and mutual funds while active mutual funds are experiencing negative outflows. This is likely attributable to active fund's underperformance of benchmark indices. **Figure 8** shows that the SPDR S&P 500 ETF has consistently outperformed active mutual funds for much of the sample.

We also sought to answer whether ETFs are better investment vehicles than mutual funds. In our analysis, we find that ETFs have the advantage of lower median expense ratios 0.75% to 1.25% and higher median returns from 22.69% to 18.43%. Thus, broadly, ETFs are better investment vehicles. Sorting by fund type and strategy, we find that active ETF expense ratios are 35% lower than active mutual funds expense while passive ETF expense ratios are 60% lower than passive mutual funds. In **Figure 6**, we show a downward trend in mutual fund expense ratios and speculate that this trend will continue as mutual funds are forced to compete with ETFs on price. The mutual fund industry has traditionally charged more for its services, but this has begun to change (**Table 3**). Mutual funds may come with sales loads and account minimums that make it difficult for small investors to get started. ETFs offer transaction cost and liquidity advantages over mutual funds.

If passive funds outperform active funds and ETFs outperform mutual funds, does that mean passive ETFs are the best investment choice? This is a tricky question. For both ETFs and mutual funds, passive funds with high expense ratios perform worse than their lower-cost peers. If two funds are benchmarked to the same index, the fund with the higher expense ratio will provide lower returns if both funds' tracking error – the deviation from their underlying index – is the same. Rational investors should choose the index fund with the lowest expense ratio; however, Choi, Laibson, and Madrian have shown that investors do not always behave rationally when selecting between index funds. Thus, the market for high expense index funds will continue to exist irrespective of their underperformance.¹⁸

Innovation in business models, such as commission-free trades, has lowered the barrier of entry for ETFs and has increased their popularity. Past research looking at the dispersion of S&P 500 returns attributes the dispersions of expense ratios across similar indices to search costs and financial literacy.¹⁸ Since this work, search costs have diminished and financial markets are easily accessible. Investigation needs to be done as to how commission-free trades affect the individual investor's long-term decisions. Additionally, prior studies that take brokerage costs into account when measuring performance need to be updated. By including transaction costs, previous authors did not anticipate financial innovations that allow for frictionless transactions.

We suspect that there is a relationship between the underperformance of active mutual funds and their net outflows. At their expense, passive funds have made tremendous gains in terms of AUM. As active mutual funds exit the market, only the strongest funds will survive.²² Our regression results show that actively managed mutual fund returns are positively correlated with their expense ratios, suggesting performance increases with management fees (**Table 6**). This positive correlation may be partially attributable to the fact that, for active investors, skill plays a major role in whether a fund outperforms index-funds. Funds with strong performance can charge more because they deliver higher returns.

This may be explained by investor willingness to pay for skill. In an efficient market, higher-skilled managers will make more than those with lesser skills. Berk and Green theorize that the AUM can predict a mutual fund manager's skill. Berk and Green hypothesize that investors are only willing to invest, given a fund has a positive managerial burden greater than zero. Devoid of value-added, investors would not invest in managers whose burden was less than zero (equates to negative return).

$$\text{Managerial Burden} = \text{Manager Compensation} \pm \text{Money Made (or taken) from investors}$$

Employing Berk and Green's equation for active funds, we estimate that until the critical juncture where returns are nullified by managerial burden, active managers will be compensated based on skill. We also conclude that an active mutual fund manager's compensation is more representative of their skill level than an active ETF manager. This is primarily because our regression results are based on fewer funds over a shorter period.

For passive mutual funds, our expense ratio coefficients are near zero and not significant (**Table 6**). This can be interpreted in several ways. One interpretation may be that passive fund managers do not add any value. Alternatively, it can be interpreted as a manager doing their job neither outperforming nor underperforming their benchmark index. One would expect excess portfolio returns to be at or near zero if a fund is tracking the broader market.

Lastly, we looked at recessionary periods to determine if active funds performed better than passive funds, with the hypothesis being that active funds can take actions to mitigate losses. We found that, overall, actively managed mutual funds have fewer losses, suggesting that it may be wise to switch into actively managed mutual funds during these periods. More work needs to be done to fully examine this strategy, though we show that it is worth studying. We do not have sufficient information to reach a

conclusion about active ETFs. They did not exist during the 2000's recession and there were only two in our sample at the time of the Great Recession. Lastly, we note that all funds were affected by the recessions; thus, none of the fund strategies are recession-proof.

CONCLUSION

Our focus on expense ratios and performance show that high expense funds must yield higher returns to offset costs born to investors. From 1998 to 2019, there has been downward pressure on mutual fund expense ratios coinciding with the introduction and rise of ETFs as investment alternatives. On average, ETFs outperform mutual funds and passive index-based funds outperform actively managed funds. Individually, active ETFs have generated the highest median returns; however, they were not introduced until 2009. It remains to be seen whether the strong performance of active ETFs will continue as the industry matures.

Our results show that passive funds perform better than active funds, which is why we believe money is flowing from underperforming actively managed domestic equity mutual funds into index domestic ETFs and mutual funds. ETFs have seen the largest inflows and have the highest returns. Passive mutual funds have outperformed active mutual funds but have seen fewer returns than passive ETFs. Thus, it follows that domestic equity index ETFs have had higher outflows.

Since 2003 mutual fund expenses have declined (**Figure 4**). We believe this to be an attributable effect of the introduction of ETFs into the marketplace. ETFs, on average, have much lower expenses than mutual funds. The downward trend of mutual funds is best explained by the exit of high-cost funds from the market and a shift from high-cost active mutual funds to passively managed mutual funds.

More research is needed to determine performance under extreme market pressures. From our data, we show a slight divergence in performance between mutual active and passive funds during periods of recession. This is likely attributable to an actively managed fund's ability to change investment strategies, whereas passive index funds follow persistent investment strategies despite market fluctuations. Further research should look to characterize other indicators of manager skill and their effect on active management strategies.

ACKNOWLEDGEMENTS

The author thanks mentor Dr. Shastri Sandy as well as Ray Bacon and Anna Wojcicki for their contributions to this work. The author thanks the University of Missouri McNair Scholars Program for its support in making this project possible.

REFERENCES

1. Agapova, A. (2011) Conventional mutual index funds versus exchange-traded funds, *Journal of Financial Markets* 14(2), 323–343. <https://doi.org/10.1016/j.finmar.2010.10.005>
2. Investment Company Institute. (2019) *2019 Investment Company Fact Book* 59th ed., Investment Company Institute. https://www.ici.org/pdf/2019_factbook.pdf
3. Guedj, I., & Huang, J. C. (2009) Are ETFs Replacing Index Mutual Funds? In *AFA 2009 San Francisco Meetings Paper*, <https://dx.doi.org/10.2139/ssrn.1108728>
4. Poterba, J. M., & Shoven, J. B. (2002) Exchange-Traded Funds: A New Investment Option for Taxable Investors, *American Economic Review* 92(2), 422–427. <https://www.aeaweb.org/articles?id=10.1257/000282802320191732>
5. Bergstresser, D., Chalmers, J. M. R., & Tufano, P. (2009) Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry, *The Review of Financial Studies* 22(10), 4129–4156. <http://doi.org/10.1093/rfs/bhp022>
6. Council of Economic Advisers (2015) The effects of conflicted investment advice on retirement savings, <http://purl.fdlp.gov/GPO/gpo55500> (accessed Oct 2019)
7. Securities and Exchange Commission, Press Release: SEC Adopts Rules and Interpretations to Enhance Protections and Preserve Choice for Retail Investors in Their Relationships With Financial Professionals, sec.gov/news/press-release/2019-89 (accessed Oct 2019)
8. 115th Congress (2017) Public Law 115–97, <https://www.congress.gov/115/plaws/publ97/PLAW-115publ97.pdf>
9. The White House, FACT SHEET: Middle Class Economics: Strengthening Retirement Security by Cracking Down on Backdoor Payments and Hidden Fees, <https://obamawhitehouse.archives.gov/the-press-office/2015/02/23/fact-sheet-middle-class-economics-strengthening-retirement-security-crack> (accessed Oct 2019)
10. United States Securities and Exchange Commission, Mutual Fund Fees and Expenses, <https://www.sec.gov/fast-answers/answersmfteeshtm.html> (accessed Oct 2019)
11. Jaffe, C. (2008) Bear Stearns Current Yield (YYY), *The Seattle Times*, <https://www.seattletimes.com/business/bear-stearns-current-yield-yyy/> (accessed Oct 2019)

12. French, K. R. (2008) Presidential Address: The Cost of Active Investing, *Journal of Finance*, 63(4), 1537–1573. <https://doi.org/10.1111/j.1540-6261.2008.01368.x>
13. Chung, J., Sensoy, B., Stern, L., Weisbach, M. (2012) Pay for Performance from Future Fund Flows the Case of Private Equity, *Review of Financial Studies* 25(11), 3259–3304. <https://doi.org/10.1093/rfs/bhr141>
14. Holmström, B. (1999) Managerial Incentive Problems: a Dynamic Perspective, *The Review of Economic Studies* 66(1), 169–182. <https://doi.org/10.1111/1467-937X.00083>
15. Berk, J. B., Green, R. (2004) Mutual Fund Flows and Performance in Rational Markets, *Journal of Political Economy* 112(6), 1269–1295. <https://www.journals.uchicago.edu/doi/abs/10.1086/424739>
16. Berk, J. B., Van Binsbergen, J. H. (2015) Measuring skill in the mutual fund industry, *Journal of Financial Economics* 118(1), 1–20. <https://doi.org/10.1016/j.jfineco.2015.05.002>
17. Lim, J., Sensoy B., Weisbach, A. (2016) Indirect Incentives of Hedge Fund Managers, *The Journal of Finance* 71(2), 871–918. <https://doi.org/10.1111/jofi.12384>
18. Choi, J. J., Laibson, D., Madrian, B. C. (2010) Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds, *The Review of Financial Studies* 23(4), 1405–1432. <https://doi.org/10.1093/rfs/bhp097>
19. Pástor, L., Stambaugh, R. F., Taylor, L. A. (2015) Scale and Skill in Active Management, *Journal of Financial Economics* 116(1), 23–45. <https://doi.org/10.1016/j.jfineco.2014.11.008>
20. Hortaçsu, A., Syverson, C. (2004) Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds, *The Quarterly Journal of Economics* 119(2), 403–456. <https://doi.org/10.1162/0033555041382184>
21. Wharton Research Data Services, CRSP Mutual Funds, wrds.wharton.upenn.edu (accessed Oct 2019)
22. Brown, S. J., Goetzmann, W., Ibbotson, R. G., Ross, S. A. (1992) Survivorship Bias in Performance Studies, *The Review of Financial Studies* 5(4), 553–580. <https://doi.org/10.1093/rfs/5.4.553>
23. Brown, S. J., Goetzmann, W. (1995) Performance persistence, *Journal of Finance* 50(2), 679–698. <https://doi.org/10.1111/j.1540-6261.1995.tb04800.x>
24. Carhart, M. (1997) On persistence in mutual fund performance, *Journal of Finance* 52(1), 57–82. <https://doi.org/10.1111/j.1540-6261.1997.tb03808.x>
25. Daniel, K., Grinblatt, M., Titman, S., Wermers, R. (1997) Measuring Mutual Fund Performance with Characteristic-Based Benchmarks, *The Journal of Finance* 52(3), 1035–1058. <https://doi.org/10.1111/j.1540-6261.1997.tb02724.x>
26. Evans, R. B. (2010) Mutual Fund Incubation, *Journal of Finance* 65(4), 1581–1611. <https://doi.org/10.1111/j.1540-6261.2010.01579.x>
27. First Bridge Data, List of active ETFs (US Listed ETFs), <https://firstbridgedata.com/page/list-of-active-etfs> (accessed Oct 2019)
28. NASDAQ, Nasdaq's List of Fifth Character Symbol Suffixes, <https://www.nasdaqtrader.com/content/technicalsupport/specifications/dataproducts/nasdaqfifthcharactersuffixlist.pdf> (accessed Oct 2019)
29. Center for Research in Security Prices, *Survivor-bias-free U.S. Mutual Fund Database Guide*, http://www.crsp.com/files/MF_Sift_Guide.pdf (accessed Sep 2019)
30. Morningstar Inc., Morningstar Investing Glossary, <https://www.morningstar.com/InvGlossary/> (accessed Oct 2019)
31. Stock, J., Watson, M. (2015) Interactions between Independent Variables, in *Introduction to Econometrics* 3rd ed., 278–287, Pearson Education, New Jersey.
32. Fama Jr., E., Multifactor Investing, https://www.ifa.com/academic-papers/multifactor_investing_eugene_fama_2006.pdf (accessed Oct 2019)
33. Fama, E., French, K. (1993) Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33(1), 3–56. [https://doi.org/10.1016/0304-405X\(93\)90023-5](https://doi.org/10.1016/0304-405X(93)90023-5)
34. Wharton Research Data Services, Fama-French Factors, wrds.wharton.upenn.edu (accessed Oct 2019)
35. French, K., Data Library, https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html (accessed Oct 2019)
36. Bernstein, W. J., Rolling Your Own: Three-Factor Analysis, <http://www.efficientfrontier.com/ef/101/roll101.htm>
37. Foundation for Statistical Computing (2019) R, Version 3.6 for Windows
38. StataCorp (2015) Stata Statistical Software, Release 14
39. National Bureau of Economic Research, US Business Cycle Expansions and Contractions, <https://www.nber.org/cycles.html> (accessed Oct 2019)

ABOUT THE STUDENT AUTHOR

James Monroe Gamble IV is a senior at the University of Missouri pursuing a Bachelor of Science in Business Administration with an emphasis in Economics and a minor in mathematics. After graduating, James plans to attend graduate school to obtain a Ph.D. in economics with interests in financial markets and household finance. Long-term, James sees himself becoming a professor and renowned economics researcher.

PRESS SUMMARY

In this study, we examine the impact of the emergence of Exchange Traded Funds (ETFs) as alternative investment vehicles to mutual funds. We compare ETF returns, performance, and expense ratios to those of mutual funds. We find over time an increased correlation between mutual fund fees and ETF fees and attribute this relationship to increased competition between mutual funds and ETFs. We also find up to a certain threshold, actively managed funds are worth their costs.

Acoustic Identification of Wild Gray Wolves, *Canis lupus*, Using Low Quality Recordings

Cara B Hull, Caitlin M McCombe, & Angela M Dassow*
Department of Biology, Carthage College, Kenosha, WI

<https://doi.org/10.33697/ajur.2020.005>

Students: chull1@carthage.edu and cmccombe@carthage.edu
Mentor: adassow@carthage.edu*

ABSTRACT

Invasive trapping and radio-collaring techniques are currently used by conservation biologists to study the population dynamics of gray wolves (*Canis lupus*). Previous research has found wolf howls can be used to determine individual identity on high quality recordings from captive animals, offering an opportunity for non-invasive monitoring of packs. We recorded wild wolves in Central Wisconsin to determine the effectiveness of these features in determining individuality in low quality recordings. The wolf howls analyzed were from two adult individuals from separate packs. Using a principle component analysis, maximum frequency and end frequency of the calls were determined to be most individualistic. Using these features in a discriminant function analysis, howls were able to be identified from individuals with 100% accuracy. Gray wolves play an important role in ecosystem maintenance, however, the current monitoring techniques are costly and invasive. The creation of an easily accessible, non-invasive technique that can be used by individuals with a variety of technical backgrounds is necessary to address concerns faced by conservation efforts. To address these issues, all analyses performed used free or low-cost software, making this method of individual identification a useful alternative for conservation biologists.

KEYWORDS

Canis lupus lycaon; Gray Wolf; Acoustic Signatures; Howls; Tracking Method; Conservation; Vocal Individuality

INTRODUCTION

Populations of gray wolves, *Canis lupus*, are monitored by conservation biologists due to this species' importance in maintaining healthy ecosystems.¹ However, tracking methods used to accomplish this are often highly invasive and costly.² Radio-collaring techniques used to track these populations are limited due to a small range of detection, time requirement, and high cost.³ The trapping techniques required for radio-collars can have a negative effect on wolves, such as avoidance of the capture area, increased susceptibility to disease, increased susceptibility to predation, as well as a loss of reproductive success.² Less invasive techniques have been suggested to track animals by analyzing their vocal individuality.^{4,5} Wolf packs occupy large home ranges predominated by coniferous-hardwood forests.⁶ Gray wolf individuals can be identified based on acoustic features of their howls, a type of vocalization characterized by their low frequency and high amplitude that allows them to propagate over long distances in highly forested areas.^{7,8} In some mammals, size and social status has been shown to affect the individual's vocalizations.² Wolf populations rely on howling as a form of communication within and across packs to signal territorial boundaries and social status.⁸ These attributes allow howls to be detected up to 10 km depending on terrain and weather conditions.¹⁰

The characteristics of wolf howls such as their maximum frequency, amplitude, duration, and harmonics are used to identify individuals.^{7,8} A prior study on captive Eastern gray wolves (*Canis lupus lycaon*) found that changes in the amplitude of the first four harmonics and the fundamental frequency of howls resulted in 95.5% accurate identification.⁷ The authors extracted amplitude harmonics on top of frequency and duration measurements from high-quality recordings of 89 howls of 6 captive adults to improve the identification accuracy of individual wolf vocalizations. A custom-made code in Matlab (Mathworks Inc. 2005) analyzed these features and classified individuals with high-levels of accuracy using a discriminant function analysis (DFA) compared to the free-source speech program Praat (95.5% vs 82.8%). Despite their success, these programs may not be feasible to learn by wildlife biologists needing to assess yearly pack compositions as they require extensive programming experience. Additionally, for the program that yielded the highest accuracy was very expensive which may limit widespread use. Also, some of the features that were measured can become attenuated in the low quality recordings these biologists would record in the wolves' natural environment outside captivity.⁶

This study aims to record wild gray wolves to determine whether individuality can be classified when howls become attenuated. Previous work identifying wild wolves achieved a 95.8% accuracy using fundamental frequency with Raven Pro 1.3 (Cornell

Laboratory of Ornithology), another expensive program.¹¹ The authors identified individualistic features that differed from those used to classify captive animals. Our study tests whether the acoustic features used to identify individual gray wolves in low-quality recordings will differ from features used in high-quality recordings. We propose a cost-effective method for acoustic monitoring of gray wolves using low-quality recordings that is able to be used by researchers with a variety of analytical backgrounds.

METHODS AND PROCEDURES

Study site and subjects

We studied four wolf packs located in Central Wisconsin that included: East Mather, West Mather, Meadow Valley, and South Bluff. These packs have home ranges in close proximity to one another that varied in size from 20 to 120 square miles. Packs occupied sandy, upland conifer-hardwood forests of oak, aspen, and jack pine.⁶ Beyond this base habitat structure, each pack's territory held unique habitat types. Oak savannas predominated East Mather, urban development fragmented West Mather forests, and large marshes and flowages characterized South Bluff and Meadow Valley. The Wisconsin DNR recently surveyed these packs and provided the authors with the locations of their home ranges based on radio-collar, winter snow-tracking data, and howl surveys.

Fieldwork

Daily surveys began by searching for scat and tracks down drivable roads within a pack's territory to use as an indicator for wolf presence. Tracks indicated the direction the wolf travelled and its size. Scat and tracks indicated wolf activity which determined the route of nighttime howl surveys used to obtain vocalizations from individuals in the wild. Fieldwork was conducted throughout August 2017, late in the pup rearing season when wolves were more vocally active, but still unable to travel far distances quickly.⁸ We estimated wolf pack composition of adults, yearlings, and new pups of the year by the frequency of their howls since larger animals with longer vocal tracts emit lower frequency sounds.¹²

We conducted howl surveys from 9:00 p.m.–12:00 a.m. when wolves are most responsive. Around 9:00 p.m., the assembled recording equipment consisting of a DR-44WL recording device (TASCAM, Santa Fe Springs, CA, USA) with built-in XY microphones and an AT8035 shotgun condenser microphone (Audio-Technica U.S., Inc, Stow, OH, USA). During the howl survey, one researcher would howl following the protocols established in Harrington and Mech in order to encourage howl responses from wolf packs by imitating a lone wolf.^{13,14} Wolves responsive to simulated howls primarily stand their ground, while unresponsive wolves (predominantly lone wolves or small packs) are more likely to retreat.^{9,14} Wild wolves may also approach human howlers to identify its source or when several kilometers away better assess features that are attenuated by distance.^{8,14} This suggested wolves considered imitated howls to be howls from conspecifics, as live howls elicit greater response than playback calls.^{8,9}

A second researcher would record the entirety of the session in mono channel at a 32 bit depth and 44.1 kHz sampling rate. If there was a positive response at any location we established the route to obtain closer proximity to the wolves, in order to acquire better recordings. At each of the howling spots we recorded the following data: meteorological data, GPS location, the number of wolf responses, approximate ages of wolves, and response distance and direction. The howl survey ended when vocal responses decreased around 12:00 a.m. Each night we rotated among the four packs, and also alternated researchers who howled to decrease the probability of wolves habituating to our voices. Wolves responded equally to all researchers.

Labwork

We filtered ambient noise from our recordings with Audition v. 2017.0.2 (Adobe, San Jose, CA, USA) by using Chebychev filters, noise reduction, and sound remover processes along with amplifier effects that increased the amplitude of the howl. Chebychev filters create a more immediate filtering effect on filtered out sounds. We used a high-pass filter on the Chebychev to remove sounds lower than a specific cut-off (around 300 Hz) of ambient noise. Measurements from Audition spectrograms included: the fundamental frequency (F_0), maximum frequency (F_{\max}), dominant frequency (F_{dom}), start frequency (F_{start}), middle frequency (F_{middle}), end frequency (F_{end}), call duration, intercall interval (ICI), dynamic range, minimum frequency (F_{\min}), bandwidth, and rises and falls from the (Table 1). In the current study, we selected these acoustic features based on past studies of canids and other organisms.^{8, 15, 16} We performed a Fast Fourier Transform (FFT) of 1024 and a window size of 0.5 seconds for call analysis. All features used to analyze distinctions between individuals were determined based off previous studies conducted on monk seals, gray wolves, and frogs in the genus *Thoropa*.^{7, 15, 16} Two researchers took measurements independently and compared results to minimize observer bias. We statistically analyzed the data using the program PAST3 v. 3.15.¹⁷ Welch's F test determined whether the amount of variation among an individual's howls were less than the amount of variation between individuals. Using PAST, a principal component analysis (PCA) clustered individuals howls based off feature similarity while also determining which features are most individualistic. With these results a discriminant function analysis (DFA) set the level of correct identification using the two most individualistic features. To better visualize the results from the DFA, we created a scatter plot using Matlab v.

9.6.0. (Mathworks Inc., 2019).¹⁸ We removed three howls from the analysis: one from the Meadow Valley individual and two from the South Bluff individual. The end of the howl from the Meadow Valley adult cracked on the microphone, causing it to be removed from skewed frequency measurements. The low amplitude of two howls from the South Bluff individual resulted in inaccurate measurements over 100Hz below the actual frequency, prompting their removal.

RESULTS

From the howl responses recorded during the month of August, the pack size and ages of individuals within the pack was approximated. All packs comprised of adults, yearlings, and juveniles used to identify howl responses (Table 2). We recorded a total of 20 minutes of usable howls containing howls overall from the four packs over several nights. Because a large portion of the recordings were chorus howls that overlapped in both time and frequency, the howls were unable to be used in the clustering analysis to identify individuals. Measurements included eight howls from a Meadow Valley adult wolf (Fig. 1) and thirteen howls from a South Bluff adult wolf (Fig. 2). Howls belonged to recordings of a single night for each individual. The measurements for minimum frequency, bandwidth, and rises and falls were removed from the analysis due to an inability to obtain accurate measurements from the spectrogram.

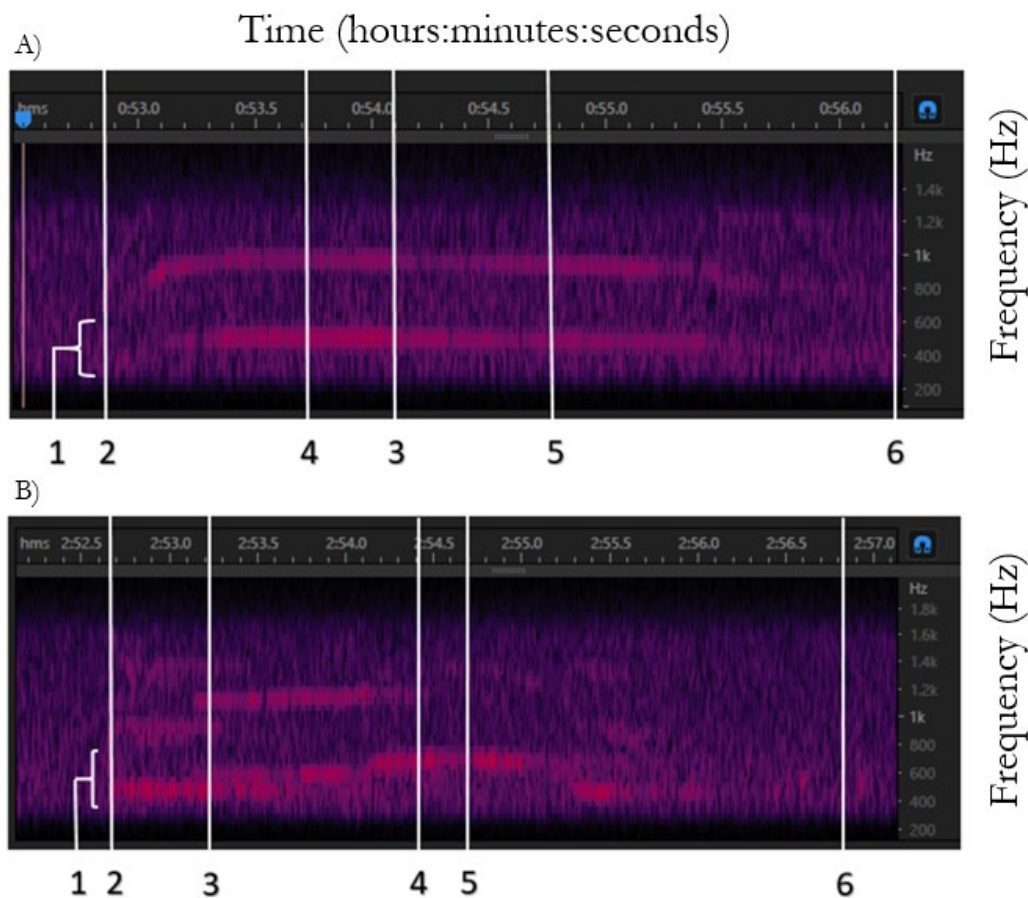


Figure 1. Sample of a howls from individual adult members of the A) Meadow Valley and B) South Bluff wolf packs in Central Wisconsin. The frequency is measured along the y-axis in Hz, time along the x-axis in the hours:minutes:seconds into the recording file. Amplitude is depicted by variation in the intensity of color, with warmer color indicating higher amplitude. Each of the measured frequency units are indicated above as follows: F_0 (1), F_{start} (2), F_{dom} (3), F_{max} (4), F_{mid} (5), and F_{end} (6).

A PCA determined the most individualistic features. We excluded F_{middle} because it negatively affected the correct classification of individuals. The results of this analysis clustered the howl into two separate ellipses that somewhat overlapped. This analysis found that the most individualistic feature was F_{max} (56.16% variance), followed by F_{end} (18.11% variance), and F_{start} (14.43% variance). F_{max} corresponded to the highest eigenvalue, thus representing the principal component with the greatest weighting; F_{start} and F_{dom} have principal components with the second and third highest weighting (Table 3).

When comparing the features using the discriminant function analysis, F_{max} and F_{end} gave the highest percentage of accurate identification, 100%, and eigenvalue, 4.7502, between the howls of the two adult wolves (Figure 4). We removed two howls from

the South Bluff individual and one howl from the Meadow Valley individual from the analysis because of a microphone crack and low amplitude. Of the remaining howls, 11/11 were correctly classified as the South Bluff individual and 7/7 were correctly classified as the Meadow Valley individual. Welch's F test was used to compare the means of F_{end} and F_{max} between wolves to determine if they were significantly different. The p-value for F_{end} was found to be 0.0072 and the p-value for F_{max} was 0.045.

Feature Analyzed	Description of Feature	Range of Measurements
F_0	The fundamental frequency of the entire call	430–570 Hz
F_{max}	The maximum frequency of the entire call	455–665 Hz
F_{dom}	The dominant frequency of the entire call	450–550 Hz
F_{start}	The frequency at the start of the call where there is an initial major rise in frequency	330–540 Hz
F_{middle}	The frequency at the middle of the call measured based on the start and stop time of the call, this was determined by finding the accurate measurement closest to the middle	410–630 Hz
F_{end}	The frequency at the end of the call before there was a consistent sudden drop in frequency	345–500 Hz
Call Duration	The time from the first spike in frequency to the last drop in frequency where the frequency is consistent	2.5–4.5 s
Intercall Interval (ICI)	The time between the end of one howl and the beginning of another determined by start and stop times	1.5–17.0 s
Dynamic Range	The level of variation in amplitude between the start and stop time of each call	6.80–14.10 dB
F_{min}	The minimum frequency measured within the continuous call including the start and end frequency	No accurate measurements could be obtained
Bandwidth	The maximum change in frequency throughout the call determined by subtraction of F_{min} from F_{max}	No accurate measurements could be obtained
Rises and Falls	Rises and falls were counted based on the number of changes in the fundamental frequency of at least 25 Hertz ¹⁹	No accurate measurements could be obtained

Table 1. List of features that were measured in Adobe Audition v 2017.0.2. All of the features listed were determined based on past literature about determining vocal individuality, however, some were removed from the analysis due to ineffective measurements that are explained in the discussion below. These features were measured in Adobe Audition with the frequency analysis and amplitude statistics tools.

Wolf Pack	Adults	Yearlings	Pups	Number of howl responses
Meadow Valley	1–2	1–2	3–4	9
South Bluff	1–2	0	1	4
East Mather	2	1	3–4	7
West Mather	2	1	2	4

Table 2. Estimated number of adults, yearlings, and pups, and the number of howl responses from four packs in Central Wisconsin. Of all the responses, we analyzed only two recordings, one from a Meadow Valley adult and one from a South Bluff adult.

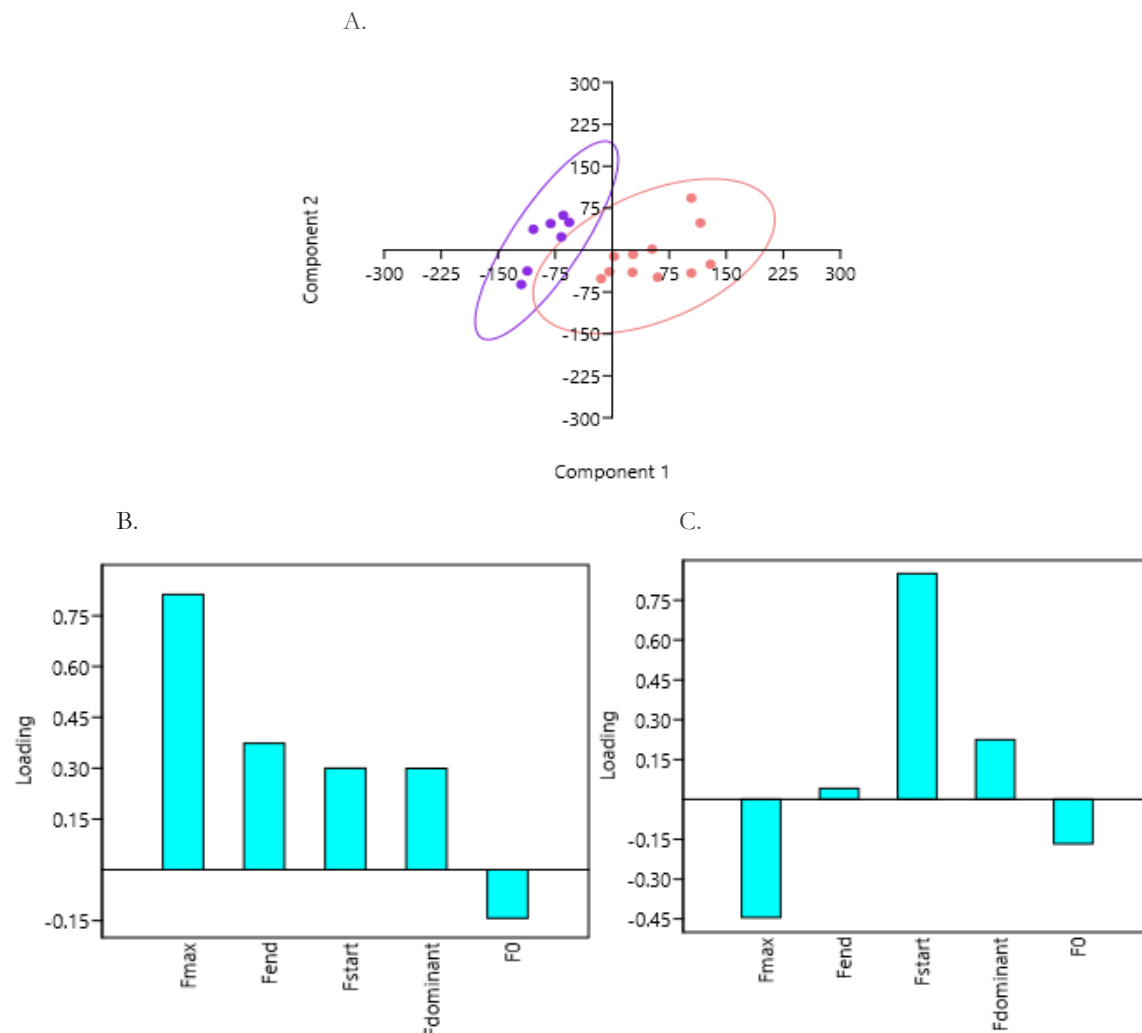


Figure 3. Principal component analysis (PCA) conducted using analyzable frequency measurements: F_{max} , F_{end} , F_{start} , and F_0 . A) Scatter plot for the South Bluff adult wolf howls (pink) and the Meadow Valley adult wolf howls (purple) with 95% confidence interval ellipses surrounding the howls for each individual. The loading plots use eigenvalues to show the weighting of each acoustic feature in B) PC1 and C) PC2.

PC	Eigenvalue
1	6812.34
2	2196.56
3	1749.73

Table 3. Chart of eigenvalue from principal component analysis. Results of principal component analysis with the principal component number (PC) and associated eigenvalue of the frequency measurements used for the PCA.

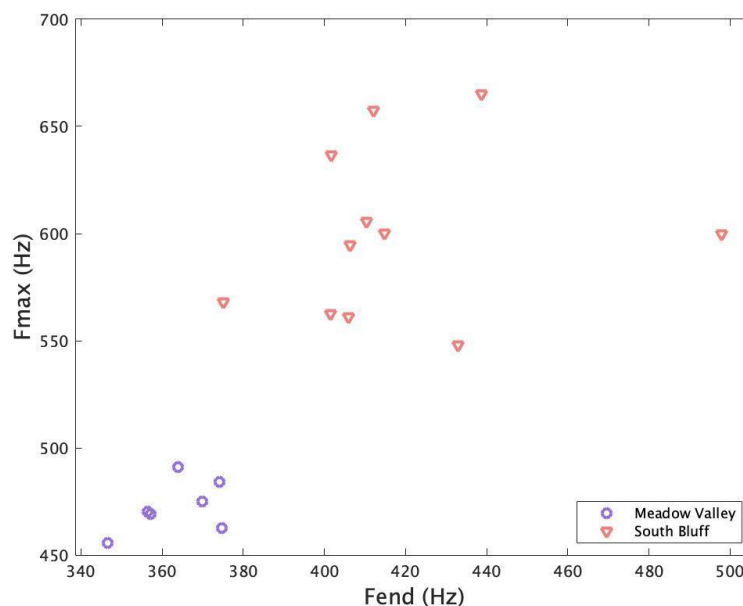


Fig 4. Scatterplot of the discriminant function analysis (DFA) using F_{\max} and F_{end} . Given the sets of howls from two individuals, DFA trains a classifier to the data to maximize the difference between sets. Each point represents a howl at a particular F_{end} (x-axis) and F_{\max} (y-axis). Pink depicts howls from the South Bluff individual and purple depicts howls from the Meadow Valley individual. The DFA analysis correctly identified all 18 calls compared across the two adult individuals with 100% accuracy at an eigenvalue of 4.7502. The DFA correctly assigned 11/11 howls from the South Bluff individual and 7/7 howls from the Meadow Valley individual.

DISCUSSION

This study determined whether the best identifying acoustic features in low-quality recordings differed from those used for high-quality recordings. Using a PCA we determined that F_{\max} , F_{end} , and F_{start} had the greatest amount of variance. We were able to achieve 100% correct identification of howls in a DFA using the parameters F_{\max} and F_{end} . Acoustic individuality as a form of identification for conservation efforts has been used previously in other organisms (i.e. Common bottlenose dolphin (*Tursiops truncatus*), European bittern (*Botaurus stellaris*), and Italian wolf (*Canis lupus italicus*)).^{2, 10} However, a cost effective, user-friendly method to identify gray wolves for conservation efforts has yet to be determined. Other studies on gray wolf individuality used amplitude and fundamental frequency to best discriminate howls from high-quality recordings. We tested a method to analyze howls of gray wolves from low-quality recordings and produced results that could accurately identify individuals using different acoustic features.

We filtered and measured howls using Audition and later differentiated individuals with 100% accuracy from low-quality recordings. In contrast to the expensive price of a MATLAB license, we were able to achieve these results at the low monthly cost of an Audition license. Audition together with the free PAST3 statistical software showcases how wildlife managers can effectively assess pack composition without the need for invasive radio collaring equipment and expensive software. These programs also do not require prior coding knowledge and therefore are more user-friendly and broadly accessible.

Throughout the field season, we recorded howls from many different individuals. Overlapping frequencies between the howls of individuals in choruses prevented us from using these howls in our analysis. Alpha adults play an important role in the social responsiveness of the pack. Wolves other than the alpha pair did not respond to simulated howls alone.⁸ These results are therefore limited to the alpha adults. If future studies determine methods to isolate howls from a chorus, our research can be expanded to all pack members. Passive acoustic monitoring of alphas can provide wildlife managers with the turnover rate of pack leaders. As more data accumulates, present recordings can be compared to howls recorded in future years in order to see if the same alpha pair still leads a particular pack.

We recorded the howls from each individual on a single night (Meadow Valley on 08/25/17 and South Bluff on 08/10/17). Other conspecifics howling did not interfere with the recording, so we know that these howls are attributed to the adult. Because we did not radiocollar individuals, we ran the DFA on howls obtained from a single night to identify that these howls belonged to a particular individual. However, we still can use DFA over multiple nights. The DFA differentiated the howls with 100% accuracy, which indicates the reliability of utilizing these methods for individualistic classification.

Overlapping ellipses in PCA help determine the confidence of howls being categorized to a particular group. Although the PCA listed the middle frequency as highly individualistic, it resulted in increased overlap between the 95% ellipse of the individuals and was therefore excluded. This could have been due to variability in frequency change patterns across howls along with the duration of these frequency changes. In some howls sudden drops or rises in frequencies occurred at different portions within the call. The PCA found F_{\max} , PC1, and F_{end} , PC2 to have the highest eigenvalues and percent variance. None of the analyzed howls clustered in the overlap region between the 95% ellipses from these measurements. This signifies that 95% of the time the howls will likely cluster within these ellipses. Though the howls could possibly land within the overlap region, the howls did not fall within that area. When conducting the PCA three howls were removed, one from the Meadow Valley individual and two from the South Bluff individual. We removed the howl from the Meadow Valley adult because the microphone cracked at the end of the howl, causing skewed frequency measurements. We also removed the two howls from the South Bluff individual due to their low amplitude resulting in inaccurate measurements over 100Hz below the actual frequency.

We were able to differentiate howls between the two individuals with 100% accuracy using F_{\max} and F_{end} . The F_{end} was not expected to be highly useful when identifying individuals in low quality recordings due to the degradation of the acoustics at the end of the call, especially over long distances. However, our analysis found F_{end} to be one of the most useful features for identifying wolf howls. When further comparing the recording conditions with the two wolves' of adult age, they were similar, suggesting F_{end} is a highly individualistic feature. The recording of the wolf from South Bluff was taken on a night when the weather was partially cloudy with light rain earlier that day. This was similar to the partially cloudy weather conditions of the recording night for the Meadow Valley individual. These individuals were both reported to be producing solo howls apart from the rest of their pack members. The distances were similar with the South Bluff individual being 0.64km and the Meadow Valley individual being less than 0.32km away. Since there was a high level of similarity between recording conditions, wolf traits, and distances of the recordings, the level of sound propagation for each of the howls would not have affected the measurement. Both the Meadow Valley and South Bluff individual howled on partly cloudy nights. No precipitation occurred during the howling sessions, but it lightly rained from 7:00-7:30 pm on the day we recorded the South Bluff adult. The daily high and low temperatures and wind speed/direction were comparable between the howling days: 71°F high and 54°F low with 5mph NW winds for the Meadow Valley individual and 79°F high and 56°F low with 7mph W for the South Bluff adult. For these reasons, F_{end} seems to be influenced by the characters of the wolf such as body size and social class rather than environmental factors. For example, larger animals have larger vocal tracts resulting in lower overall vocal frequencies.¹² Adults therefore have lower F_{end} frequencies than pups.

During howl analysis, four features were immeasurable because of program limitations or low-quality recordings: F_{\min} , bandwidth, intercall interval (ICI), and rises and falls. We originally included F_{\min} based on literature focused on determining individuality in Eastern gray wolves (*Canis lupus lycaon*).^{7, 19} However, when using this feature in low quality recordings it was indistinguishable from the ambient noise. We were able to differentiate F_{end} from ambient noise, allowing us to use it as a measurable parameter due to a large frequency decrease following the end of the call. In the recordings, some portions of the howls were too low amplitude to be measured. This resulted in a high level of variability within the minimum frequencies across calls. This level of variability was attributed to the program detecting only ambient noise rather than the actual wolf howl. For this reason, the feature was determined to be unmeasurable. This inability to measure F_{\min} in the howls resulted in the loss of other features for analysis, including bandwidth since the minimum is necessary to calculate this measurement. A third feature that we determined to be unusable was duration between howls. Duration between howls was chosen because it was found to be useful in differentiating between six species of frogs within the genus *Thoropa*.¹⁵ Though the vocal production mechanisms vary greatly between frogs and wolves, this feature was explored because if successful, this could be used for solo and chorus howls. Due to the sample size of howls collected per individual and the feature only being able to be measured between two consecutive howls it resulted in a

decreased number of measurements per individual wolf with some being disrupted by barks or unmeasurable howls. Aside from this, the howls showed no form of consistency in how often they were emitted, therefore, this feature was removed from the analysis.

Frequency modulation throughout the call of organisms was also found to be highly useful in past studies. In a study analyzing timber wolf howls, measurements were made on the number of abrupt changes in frequency that were different than the fundamental frequency by more than 25 Hz.¹⁷ More recently, in a study on dolphins' signature whistles, frequency modulation throughout the call was calculated by using algorithms to compare frequency changes at various points in time.²⁰ The method needed for quantitative analysis of the frequency modulation across the entire call is highly complex. Literature that uses this form of analysis relies heavily on programs that are costly, making the method less accessible for conservation efforts that have limited budgets. This method also requires a strong background in computer programming, removing the practicality of its use for people with a large assortment of backgrounds and skill sets. When applying the simpler method of measuring the number of abrupt changes, it becomes difficult to define set parameters because of the high level of variability in temporal and frequency aspects within howls. This leads to a high level of observer bias being introduced into the measurements, therefore, rises and falls were removed from the analysis.

CONCLUSIONS

The use of bioacoustics in conservation efforts has become a topic of rising interest because of its ability to offer a noninvasive method to track vocalizing populations.²¹ Thus far research on the topic of bioacoustics in conservation efforts has shown promising results, especially within gray wolf populations where research has been able to identify adult wolves using howls with 100% accuracy.⁷ However, current analyses of animal vocalizations use complex methods and expensive programs that are not user-friendly. With current methods, the use of vocal identity to track populations becomes challenging for widespread use by conservation biologists due to an inability to access many of the supplies and a lack of training in skills needed for these methods. Through this research we propose an accessible and inexpensive method to track wolf populations, including turnover rates and pack size. Low-cost programs were selected for all of the analyses that are targeted towards use by people with a variety of backgrounds in computational training. The most individualistic features have also been identified within the howls so analyses can be conducted in shorter amounts of time with only the features that were found to be reliably indicative of individuality. To further test the viability of these methods in the future we will increase the sample size of the analysis and conduct the study over multiple years on the same packs. This will allow us to determine the usefulness of these methods to track turnover rates.

ACKNOWLEDGEMENTS

The authors would like to express their appreciation to Dick Thiel, Norma Donovan, Ray Leonard, Linda Nelson, Bob Sherman, and the Timber Wolf Information Network for advising on wolf pack information, Sandhill Wildlife Area for the utilization of their facilities, Carthage College Natural and Social Science Division for sponsoring the Summer Undergraduate Research Experience (SURE) program, and the Carthage College Biology Department for supporting undergraduate research.

REFERENCES

- Fortin, D., Hawthorne, B., Boyce, M., Smith, D., Duchesne, T. and Mao, J. (2005) Wolves influence elk movements: behavior shapes a trophic cascade in Yellowstone National Park. *Ecology* 86: 1320-1330. <https://doi.org/10.1890/04-0953>
- Terry, A., Peake, T. and McGregor, P. (2005) The role of vocal individuality in conservation. *Frontiers in Zoology* 2: 10. <https://doi.org/10.1186/1742-9994-2-10>
- Harris, S., Cresswell, W.J., Forde, P.G., Trehwella, W.J., Woollard, T. and Wray, S. (1990) Home-range analysis using radio-tracking data—a review of problems and techniques particularly as applied to the study of mammals. *Mammal Review* 20: 97-123. <https://doi.org/10.1111/j.1365-2907.1990.tb00106.x>
- Hartwig, S. (2005) Individual acoustic identification as a non-invasive conservation tool: an approach to the conservation of the African Wild Dog, *Lycaon pictus* (Temminck, 1820). *Bioacoustics* 15: 35-50. <https://doi.org/10.1080/09524622.2005.9753537>
- Darden, S., Dabelsteen, T. and Pedersen, S. (2003) A potential tool for swift fox (*Vulpes velox*) conservation: individuality of long-range barking sequences. *Journal of Mammalogy* 84: 1417-1427. <https://doi.org/10.1644/BEM-031>
- Brown, R.T. and Curtis, J.T. (1952) The Upland Conifer-Hardwood Forests of Northern Wisconsin. *Ecological Monographs*. 22(3):217-234. [doi:10.2307/1943566](https://doi.org/10.2307/1943566)
- Root-Gutteridge, H., Bencsik, M., Chebli, M., Gentle, L.K., Terrell-Nield, C., Bourit, A. and Yarnell, R.W. (2014) Improving individual identification in captive Eastern grey wolves (*Canis lupus lycaon*) using the time course of howl amplitudes. *Bioacoustics-the International Journal of Animal Sound and Its Recording* 23: 39-53. <https://doi.org/10.1080/09524622.2013.817318>
- Suter, S.M., Giordano, M., Nietlispach, S., Apollonio, M. and Passilongo, D. (2017) Non-invasive acoustic detection of wolves. *Bioacoustics-the International Journal of Animal Sound and Its Recording* 26: 237-248. <https://doi.org/10.1080/09524622.2016.1260052>
- Harrington, F. and Mech, D. (1979) Wolf howling and its role in territory maintenance. *Behaviour*, 68: 207-249. [doi:10.1163/156853979X00322](https://doi.org/10.1163/156853979X00322)

10. Joslin P.W.B. (1967) Movements and home sites of timber wolves in Algonquin park. *American Zoologist* 7: 279–288. <https://doi.org/10.1093/icb/7.2.279>
11. Passilongo D., Buccianti A., Dessi-Fulgheri F., Gazzola A., Zaccaroni M. and Apollonio M. (2012) The acoustic structure of wolf howls in some eastern Tuscany (central Italy) free ranging packs. *Bioacoustics*. 19:159–175. <https://doi.org/10.1080/09524622.2010.9753622>
12. Faragó, T., Townsend, S. and Range, F. (2014) The information content of wolf (and dog) social communication, in *Biocommunication of Animals*. (Witzany, G., Ed.) 1st ed., 41-62. Springer, Dordrecht.
13. Harrington, F. and Mech, D. (1982) An analysis of howling response parameters useful for wolf pack censusing. *The Journal of Wildlife Management* 46: 686-693. doi: 10.2307/3808560
14. Leblond, M., Dussault, C. and St-Laurent, M.H. (2017) Space use by gray wolves (*Canis lupus*) in response to simulated howling: a case study and a call for further investigation. *Canadian Journal of Zoology*. 95(3): 221-226. doi:10.1139/cjz-2016-0191
15. Charrier, I., Marchesseau, S., Dendrinis, P., Tounta, E. and Karamanlidis, A.A. (2017) Individual signatures in the vocal repertoire of the endangered Mediterranean monk seal: new perspectives for population monitoring. *Endangered Species Research* 32: 459-470. <https://doi.org/10.3354/esr00829>
16. Nunes-de-Almeida, C.H.L., Assis, C.L., Feio, R.N. and Toledo, L.F. (2016) Redescription of the advertisement call of five species of Thoropa (Anura, Cycloramphidae), including recordings of rare and endangered species. *Plos One* 11: e0162617. <https://doi.org/10.1371/journal.pone.0162617>
17. Hammer, Ø., Harper, D.A.T., Ryan, P.D. (2001) PAST: Paleontological Statistics software package for education and data analysis. *Palaeontologia Electronica* 4(1): 9.
18. MATLAB. (2019). MATLAB version 9.6.0. 1114505 (R2019a). The Mathworks Inc. Natick, Massachusetts.
19. Tooze, Z.J., Harrington, F.H. and Fentress, J.C. (1990) Individually distinct vocalizations in timber wolves, *Canis lupus*. *Animal behavior* 40: 723-730. [https://doi.org/10.1016/S0003-3472\(05\)80701-8](https://doi.org/10.1016/S0003-3472(05)80701-8)
20. Kershenbaum, A., Sayigh, L.S. and Janik, V.M. (2013) The encoding of individual identity in dolphin signature whistles: how much information is needed?. *Plos One* 8: e77671. <https://doi.org/10.1371/journal.pone.0077671>
21. Palacios V, Font E and Márquez R. (2007) Iberian wolf howls: acoustic structure, individual variation, and a comparison with North American populations. *Journal of Mammalogy* 88(3): 606–613. <https://doi.org/10.1644/06-MAMM-A-151R1.1>

ABOUT STUDENT AUTHORS

Cara Hull is a student research assistant graduating from Carthage College in spring 2019 with a Bachelor of Arts in Biology with a minor in Chemistry. She currently studies bioacoustics with the intention to focus on research in conservation and genetics in the future. In fall 2019 she will be continuing her education at the University of Southern California pursuing a PhD in Molecular Biology.

Caitlin McCombe is a student research assistant who will be graduating in spring 2020 from Carthage College with a Bachelor of Arts in Biology and a minor in Chemistry. She currently studies bioacoustics and intends to pursue this research area into graduate school, but within the field of marine biology. In summer 2019, she will be working with the Read Lab at Duke University studying pilot whales.

PRESS SUMMARY

Current monitoring of gray wolves utilizes invasive trapping and radio-collaring methods to track population dynamics. Individual identity can be determined using wolf howls with high quality recordings of captive animals. Utilizing this knowledge, the following study aims to create a low cost method to track wild gray wolves in Central Wisconsin using acoustic features previously identified as individualistic in captive wolf analyses. Upon completion of future testing, the following method will provide a novel and non-invasive technique for monitoring gray wolves for use in wildlife management.

The Allele Frequency of the HFE gene mutation H63D (rs1799945) and Its Relationship to a Hereditary Hemochromatosis Diagnosis in Metabolic Nutrition Students at Virginia Tech

Tyler R. Ferqueron, Angela S. Anderson, & Deborah J. Good*

Department of Human Nutrition, Foods, and Exercise at Virginia Polytechnic Institute and State University, Blacksburg, VA 24061

<https://doi.org/10.33697/ajur.2020.006>

Student: tylerf4@vt.edu

Mentors: goodd@vt.edu; asphay@vt.edu

ABSTRACT

Hereditary hemochromatosis (HH) is a disease that causes excess iron absorption from the diet. This excess iron can be stored in the liver, skin, heart, pancreas, and joints, and then can lead to other health conditions, as the human body has no way of actively excreting iron. The human hemochromatosis protein (HFE protein) is encoded by the HFE gene, and mutations in this gene can lead to a dysfunction of the protein resulting in HH or iron overload later in adulthood. The objective of this study was to analyze the mutant allele frequency and the penetrance of the H63D mutation (SNP rs1799945) of the HFE gene in a cohort of Virginia Tech students. This study had a total of 69 participants. Fifty-two participants provided saliva samples, genomic data from 23andMe®, and surveys with phenotypic information. Of these, 6 were genotyped using the RFLP technique and served as controls for genotype confirmation. An additional 17 participants provided saliva samples, but did not provide 23andMe® data; genomic DNA from these participants were genotyped using the RFLP technique. Our results showed that although none of the participants had been diagnosed with HH, the mutant allele frequency of this population was 13.04%. In conclusion, as HH is usually diagnosed in older adults, we could not identify any students with a phenotype of HH, even though we could detect the mutant allele. This data suggests that affordable and accessible genetic ancestry and health kits such as the 23andMe® kit, could provide an efficient way to identify, prevent, and manage HH and other genetic diseases before symptoms arise.

KEYWORDS

Hereditary Hemochromatosis; Iron Absorption; Single Nucleotide Polymorphism; Restriction Fragment Length Polymorphism; College Student Population; 23andMe®; Population Analysis; Survey Results

INTRODUCTION

Hemochromatosis (HH) is a condition that causes excess absorption of the mineral iron. Iron works in the body as an electron donor and acceptor. Therefore, it is not surprising that the symptoms of HH, due to iron toxicity, are metabolic dysfunctions and imbalances from the production of reactive oxygen species (ROS).^{1,2} Excess iron produces ROS through the Fenton reaction, forming hydroxyl radicals from superoxide or hydrogen peroxide. Lipid peroxidative products which can form more toxic radicals are also formed in iron overload. Radicals formed from both the fenton reaction and from lipid peroxidative products can react with carbohydrates, proteins, and most importantly nucleic acids.¹ Not only does excess iron lead to reactive oxygen species, it can exacerbate other conditions that cause the dysregulation of iron distribution among the tissues and organs of the body. The liver is one of the first organs to be affected by excess iron absorption. The liver stores the largest amount of iron in the body, and there are no pathways for the excretion of iron.³ Too much iron in the liver can potentially lead to cirrhosis, fibrosis, inflammation, cancer, and more minor to severe liver conditions.^{1,2,4,5} Excess iron can also lead to cardiac disease including congestive heart failure, arrhythmias, hypertrophy, and endocarditis.¹

Currently, HH is associated with 5 different genes including HFE (hemochromatosis), Tfr2 (transferrin receptor 2), HJV (hemojuvelin), HAMP (hepcidin), and SLC40A1 (solute carrier family 40 member 1).³ This study focused on the HFE gene as mutations in this gene account for the most common form of HH, affecting nearly 1 million people.⁶ The most severe HFE gene variant that is associated with HH is a mutation at position 282 in which the amino acid cysteine is replaced by tyrosine, written as Cys282Tyr or C282Y (rs1800562).³ However, the most common HFE gene variant is due to the single nucleotide polymorphism (SNP), rs1799945, that results in a mutation at amino acid position 63 of the HFE protein, where the amino acid histidine is replaced with an aspartate, written as Hisp63Asp or H63D. In the gene, the wild-type allele is a cytosine (C) and the mutant allele is a guanine (G) (**Figure 1**).³ The rs1799945 variant has been linked with a mild form of HH when the homozygous genotype (GG) is present. According to the National Center for Biotechnology Information (NCBI) SNP database, Northern European populations (locations include the United Kingdom, Ireland, Guernsey, Finland, Germany, Austria, Belgium, France,

Luxembourg, Netherlands, Switzerland, Denmark, Faroe Islands, Iceland, Norway, and Sweden according to 23andMe® have a higher H63D mutant allele frequency than other populations.⁵ Reports suggest that the frequency of the H63D mutant allele among the American population is from 11.5 to 13.5%.^{5,7} For this study, we decided to focus on the most common variant, rs1799945 (**Figure 1**).

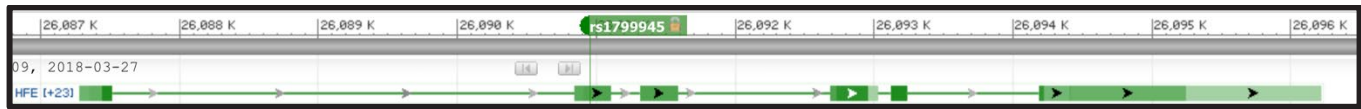


Figure 1. The location of rs1799945 within the HFE gene on chromosome six. As shown, rs1799945 is located within exon two. The HFE gene contains six total exons.

Transferrin is the transport protein of iron in the blood. The HFE protein normally interacts with the transferrin receptor (TfR) lowering its affinity for transferrin, and ultimately lowering the total amount of iron absorbed.⁸ When mutated, transferrin levels continue to signal to the intestinal cells to export iron into the blood stream. The C282Y and H63D variants are thought to disrupt the normal interaction of the HFE protein with TfR, leading to the disruption of the regulation of iron absorption and resulting in an excessive increase of iron absorbed from the diet.⁹ This leads to the condition known as type 1 hemochromatosis.^{6,8} Type 1 hemochromatosis typically affects men more than women, and first show symptoms in those around 40 to 60 years of age.⁹ The HFE protein also regulates the production of hepcidin, which is produced by the liver and decreases the amount of iron absorption, recycling, and storage.⁹

The objective of this study was to evaluate the allele frequency of the H63D mutation and its relationship to a hereditary hemochromatosis diagnosis and population ancestry in nutrition students of a junior/senior level metabolic nutrition course, using raw data from the 23andMe® ancestry kit and a Restriction Fragment Length Polymorphism (RFLP) technique.

METHODS AND PROCEDURES

Participant enrollment

A total of sixty-nine students in an upper level Metabolic Nutrition course were enrolled in this study. The participants ranged from 20 to 23 years of age. The study protocol was approved by the Institutional Review Board at Virginia Tech. Student saliva samples were deidentified and considered biospecimens. Saliva data was linked to survey and 23andMe® raw data using unique identifiers that were not linked to any identifying information.

Each participant was given a 23andMe® ancestry kit (Mountain View, CA) and completed two surveys. In the first survey, participants were given a variety of questions, one of which was “Have you or a family member been diagnosed with any of the following genetic conditions? (check all).” One of the possible responses was hereditary hemochromatosis. In the second survey, among other questions included in the study, participants were asked to rank their ancestry from their 23andMe® ancestry results. The genotypes of the participants were compared to their previous diagnoses and ancestry. The frequency of the alleles was also calculated.

DNA isolation

The collected saliva samples were mixed with a DNA stabilization buffer and stored at 4 °C. The DNA was isolated and purified from 2.5 mL samples of saliva donated by each participant using a published protocol.¹⁰ Briefly, cell lysis solution was added to degrade the cells and expose the DNA. RNA was removed using and RNaseA solution (Thermo-Fisher Scientific, Waltham, MA), while proteinase K solution (Thermo-Fisher Scientific, Waltham, MA) was used to digest the proteins. A protein precipitation solution was used to help separate the proteins from the genomic DNA. Centrifugation separated the proteins and lipids from the DNA. Addition of isopropanol and a glycogen solution followed by an ethanol wash was used to isolate and purify the genomic (g)DNA. The gDNA was rehydrated in Tris-EDTA. The concentration of DNA (ug/uL)(OD 260 x 50 ng/uL) yielded from the isolation was evaluated using a NanoDrop spectrophotometer (Thermo-Fisher Scientific, Waltham, MA), and DNA quality analyzed by agarose gel electrophoresis.

Primer preparation and Restriction Fragment Length Polymorphism (RFLP) design

Primers for the polymerase chain reactions were created using NCBI's Primer-BLAST program¹¹ (**Figure 2A**). The product target size was set from 200 bp to 500 bp. From here, a set of primers was returned that put the variant allele right in the middle. The 3' primer was reversed so it was converted to the reverse complement. Standard primers were ordered (Eurofins MWG Operon, Luxembourg). A primer stock solution was made by diluting with milli-Q (autoclaved water) to make a 100 μM solution. The forward primer (34.4 nmol) was diluted with 344 μL of pure water, and the reverse primer (39.0 nmol) was diluted with 390 μL of pure water. Twenty μL of the diluted forward and twenty μL of reverse primer were added into a small tube with 60 μL of pure water to create a 10 μM stock solution. The stocks were stored at -20 °C.

The RFLP genotyping technique is an older, but still a useful technique. RFLP viability was determined for the H63D mutation using the sequence amplified by the primers developed above and the Restriction Mapper program (<http://www.restrictionmapper.org/>).¹¹ The program determined an RFLP after inputting 10 base pairs of the sequence on both sides of the mutation. The program showed that the BclI restriction enzyme would cut the wild-type allele, but not the variant allele (Figure 2B).

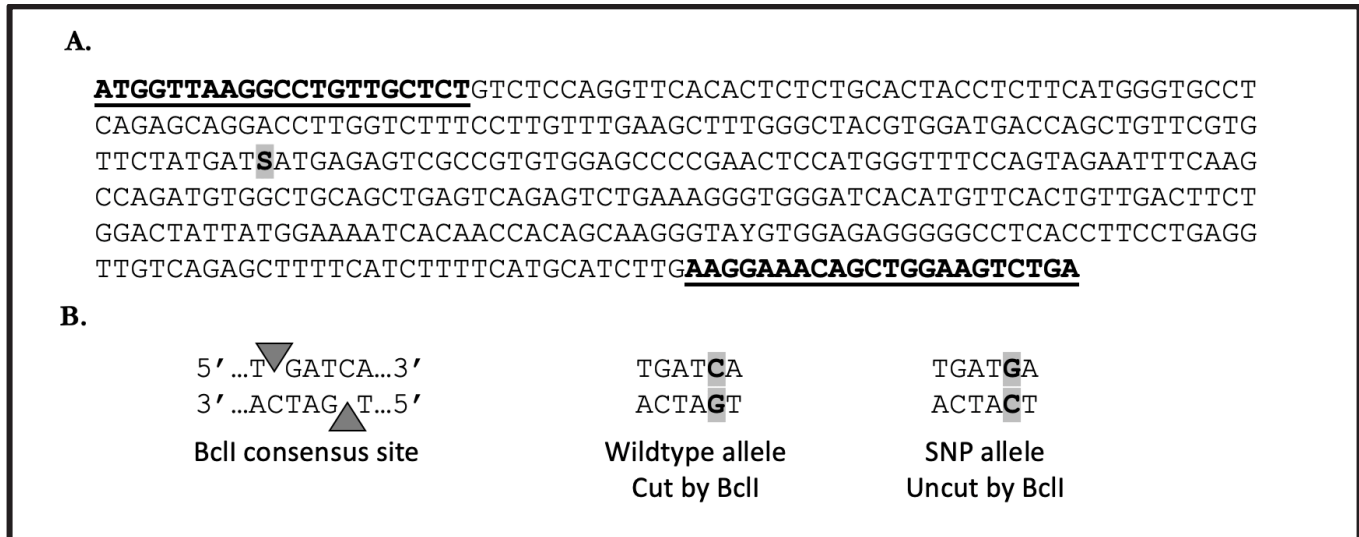


Figure 2. Primers and restriction fragment polymorphism design. **A.** The forward and reverse primers surrounding the location of the rs1799945 C/G variant (indicated by the IUPAC symbol “S”, shown in bold, gray highlight) are indicated by bold, underlined font. **B.** The consensus BclI sequence is shown, with the normal location of cut sites indicated by the arrowheads. The sequence surrounding the variant in the mutant and wildtype alleles are shown. Only the wildtype (normal) allele is cut by BclI, as the variant sequence no longer matches the consensus.

Polymerase Chain Reaction (PCR)

The polymerase chain reaction was used to amplify the HFE region surrounding rs1799945 using the primers shown in **Figure 2A**. Each PCR tube contained 10.0 µL of 2X PCR master mix (GoTaq, Promega, Madison, WI), 0.5 µL of primer stock, 2.0 µL of template genomic DNA from individual volunteers and 7.5 µL of nuclease-free water. Following mixing, the tubes were centrifuged for a few seconds to insure all reagents were at the bottom of the PCR tube. The samples were then run in an Eppendorf (Hauppauge, NY) PCR machine with an initial denaturation cycle of 95 °C for 10 minutes, followed by 30 cycles of a denaturation cycle of 95 °C for 30 seconds, an annealing cycle of 55 °C for one minute, an extension cycle of 72 °C for one minute. After a final extension of 72 °C for 10 minutes, the samples were stored at a temperature of 4 °C.

Agarose gel electrophoresis

After the PCR amplification, the samples were examined for a correctly-sized PCR amplification using a 1.5% agarose TAE gel containing ethidium bromide. Six (6) µL of each sample was mixed with 2 µL of loading dye and dispensed into their respective wells. The gel electrophoresis was carried out at 83 V for around 25 minutes. Once the bands had appeared to separate, the gels were viewed under a UV light using the BioRad Chemidoc System (Hercules, CA). Visualization of a 380bp band indicated a positive PCR amplification.

RFLP genotyping

Amplified samples from the PCR reaction were digested by the BclI enzyme (Thermo-Fisher Scientific, Waltham, MA) by adding 10 µL of the DNA, 18 µL of nuclease-free water, 2 µL of 10x Buffer G, and 1.5 µL of the BclI enzyme to a 0.5 mL centrifuge tube. The mixtures were mixed gently and incubated for approximately 1.5 hours at 37 °C. After incubation, digested samples were analyzed using the gel electrophoresis protocol previously described. The homozygous wild-type variants showed two bands of 244 bp and 136 bp respectively, while the homozygous mutant variants showed a single band of 380 bp. Heterozygous samples showed all three bands. Visualization of the bands using the BioRad Chemidoc System (Hercules, CA), allowed for the determination of the genotypes. This method of genotyping was used for the samples of the 17 participants that did not provide 23andMe® data or the second survey. Six samples from the 52 participants that provided 23andMe® data and the second survey were also genotyped using this method as controls to confirm the RFLP results.

Survey data analysis

Data on the participants' gender, ancestries, HH diagnoses, and genotypes were collected from the survey results. This information was broken down into the following groups: "Diagnosis with HH," "Northern European Ancestry," "Non-Northern European ancestry," "Unknown ancestry," "Male," and "Female." These groups were compared to the wild-type, heterozygous, and mutant genotypes.

Statistical Analysis

The significance of the genotypes relation to a HH diagnosis was assessed using an online genotype analysis program (http://www.physics.csbsju.edu/cgi-bin/stats/exact_form.sh?nrow=2&ncolumn=3)¹², which used Fisher's Exact Test. Significance was set a priori at $P \leq 0.05$.

RESULTS

PCR Results

Figure 3A shows representative PCR results that were obtained for 23 participants who originally provided saliva for genomic DNA isolation. Seventeen were from volunteers who provided saliva samples, but who did not provide genotype data from 23andMe®. An additional 6 samples served as study controls as they were from volunteers who provided both saliva samples, and genotype data from 23andMe®.

RFLP genotyping

Twenty three samples that successfully amplified were genotyped with representative samples shown in **Figure 3B**. Of the 23 samples, 17 were genotyped as WT and contained only the normal alleles (CC), 4 were genotyped as heterozygous (HET) containing one normal allele and one variant allele (CG), and 2 were genotyped as mutant (MUT), or homozygous for the mutant alleles (GG). These genotypes are presented in **Table 1**.

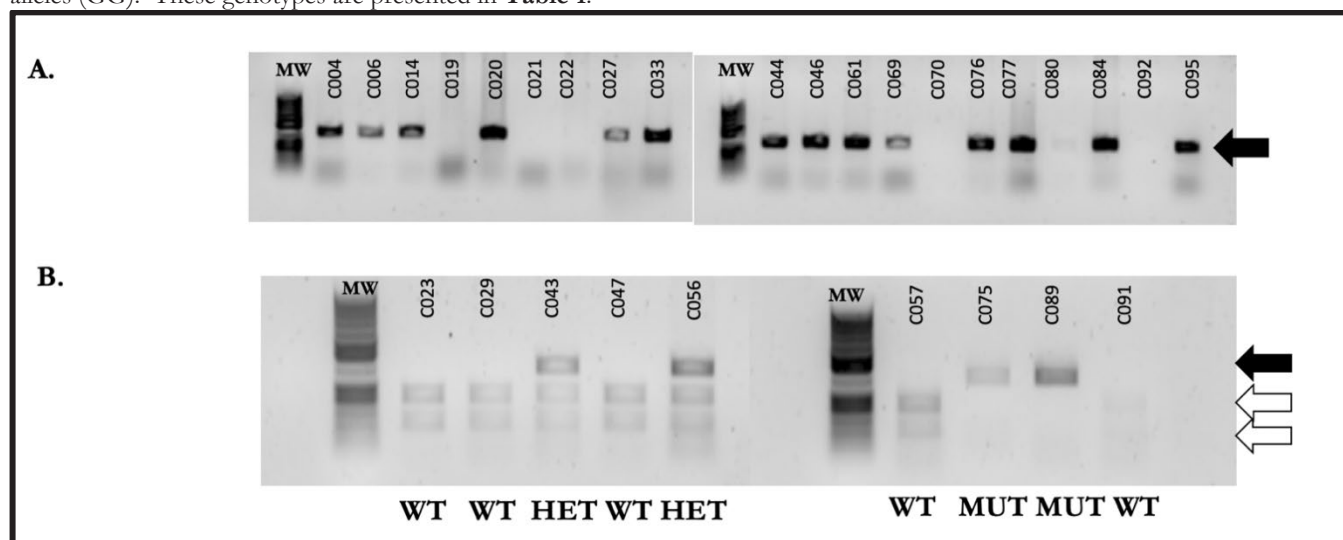


Figure 3. Representative PCR and RFLP results. **A.** PCR results. The 380 bp PCR amplicon band is indicated by the arrow. Sample numbers are provided above each lane and correspond to the listed samples in **Table 1**. PCR amplifications that were unsuccessful were excluded from the study. **B.** RFLP results. As above, sample numbers above each lane correspond to the sample identifiers in **Table 1**. The genotype for each lane is indicated above. The filled arrow indicates the allele containing the variant and is uncut at 380 bp. The open arrows show the WT alleles of 244bp and 136 bp respectively. MW-molecular weight marker; WT-homozygous the normal allele only; HET-heterozygous and contains both alleles; MUT-homozygous for the mutant allele containing SNP rs1799945.

In silico analysis and survey results

In silico analysis included a total of 69 participants. Fifty-two participants provided saliva samples, genomic data from 23andMe®, and surveys with phenotypic information. Seventeen participants provided saliva samples, but did not provide 23andMe® data; genomic DNA from these participants were genotyped using the RFLP technique. Genotype results for males and females are shown in **Table 2**. None of the volunteers indicated that they had ever been diagnosed with HH, and therefore no significant relationship between the mutant allele and an HH diagnosis was found ($P=1.000$). However, for those volunteers that responded to the second survey, a significant relationship was found in that the mutant allele was significantly related to those of Northern European descent ($P=0.0066$). For those individuals, the mutant allele frequency among this population was found to be 13.04% (**Table 3**).

Sample Code	Genotype	Sample Code	Genotype
C004	CC	C056*	CG
C006	CC	C057*	CC
C014	CC	C061	CC
C020	CC	C069	CC
C023*	CC	C075*	GG
C027	CC	C076	CG
C029	CC	C077	CC
C033	CC	C084	CC
C043*	CG	C089*	GG
C044	CC	C091	CC
C046	CG	C095	CC
C047	CC		

Table 1. DNA samples and associated and their genotypes determined by RFLP analysis.

*These samples were controls, and were genotyped by 23andMe® and the RFLP technique.

	TOTALs per category (percent)	C:C Homozygous WT	C:G Heterozygous HET	G:G Homozygous MUT	Fisher's Exact Test (<i>P</i> -value)	Mutant Allele Frequency (percent)
N (Total)	69 (100%)	53	14	2	N.A.	13.04
% of Population		76.8	20.3	2.9	N.A.	
Diagnosis with HH	0 (0 %)	0	0	0	<i>P</i> = 1.0 (Not significant)	0
Females	54 (78.2%)	41	11	2	N.A.	18.3
Males	15 (21.8%)	12	3	0	N.A.	12.5

Table 2. Genotypes from both 23andMe® and RFLP genotyping were combined to quantify the genotype frequency of the H63D mutation in males versus females.

	TOTALs per category (percent)	C:C Homozygous WT	C:G Heterozygous HET	G:G Homozygous MUT	Fisher's Exact Test (<i>P</i> -value)	Mutant Allele Frequency (percent)
N (Total)	69 (100%)	53	14	2	N.A.	13.04
% of Population		76.8	20.3	2.9	N.A.	
Diagnosis with HH	0 (0 %)	0	0	0	<i>P</i> = 1.0 (Not significant)	0
Northern European Ancestry (responded to 2nd survey)	47 (68%)	33	12	2	<i>P</i> = 0.0066 (highly significant)	17.02
Non-Northern European ancestry (responded to 2nd survey)	5 (7.2%)	5	0	0	N.A.	0
Unknown ancestry (did not respond to 2nd survey)	17 (N.A.)	15	2	0	N.A.	6.7

Table 3. Genotypes from the both 23andMe® and RFLP genotyping were combined to quantify the genotype frequency of the H63D mutation. Ancestry analysis was obtained by self-report of the volunteers using 23andMe generated data.

DISCUSSION

Type I hemochromatosis is one of the most common genetic conditions in the United States and worldwide.⁶ In our results, 20.3% of individuals in a college nutrition course were mutant allele carriers, and 2.9% of the individuals were homozygous for the allele. These results are similar to published reports of the H63D mutant allele frequency among Americans (11.5 to 13.5%),^{5,7} as well as a recent study of a Greek population for rs1799945, which found that up to 25% of individuals were carriers of the mutant allele.¹³ We found no individuals of non-Northern European descent that were carriers. This low carrier rate among those of non-Northern European descent is consistent with previously published reports.^{5,7,13,14} For example, as shown by a study examining rs1799945 in Russian and Asian groups, the frequency is only 3-8%.¹⁴ Our data also indicates a significant correlation with Northern European ancestry, as determined by the Fisher's Exact Test. Overall the study verifies the validity of studying single gene variants within small populations using both laboratory-based and in-silico based analyses.

Among all the participants, none of them had ever been diagnosed with hemochromatosis. As expected, there was no association between the mutant (G) allele and a hemochromatosis diagnosis. However, there may be reasons for why neither of the two individuals who were homozygous for the mutant allele have ever been diagnosed with HH. First, both these participants were female. This is relevant, because most females menstruate every month. With blood loss comes iron loss, so this release of blood from the body might be enough to reduce the overall levels of iron to prevent a diagnosis of HH.^{15,16} Also, this population was relatively young (20-23 years of age). Typically, signs and symptoms of HH do not appear until around 40 to 60 years of age, meaning that this population might be too young to have severe enough symptoms of HH to warrant testing for a diagnosis.⁹

The results of this study shows that determining participants genotypes using the 23andMe® data can provide genetic risk information prior to onset of HH symptoms. While current treatments for HH are effective in lowering the iron levels within acceptable ranges, diseases caused by HH such as hypogonadism, cirrhosis, arthritis, and insulin-dependent diabetes are often irreversible.¹⁷ By simply using an direct-to-consumer DNA kit, people of this age could predict their predisposition to HH and other genetic disorders before the symptoms arise at a relatively low cost. This would help individuals with HH or other genetic disorders, begin to manage and limit the possible harmful effects of genetic disorders.

There are a few improvements that could be made upon this study. First, not all the samples that underwent PCR were successfully genotyped due to technical errors in isolation, storage, of analysis. Due to the biospecimen nature of our study, volunteers could not be re-contacted to provide saliva samples for re-analysis. A new direction for this research would be to compare female populations before and after menopause, as previous findings suggest that males present greater phenotypic effects correlating with the mutant genotype than females.^{3,18,19} Also, future studies should further explore direct-to-consumer DNA kits as a means to predict HH and other harmful genetic disorders. Both of these suggestions could be done using a different recruiting strategy, such as recruiting from the entire campus including both students and faculty/staff, rather than from a single class. Finally, measuring blood serum levels of iron and transferrin may help strengthen the findings and show a better relationship between H63D and iron overload, even without the diagnosis of HH.

ACKNOWLEDGEMENTS

The authors thank the Office of Undergraduate Research at Virginia Tech for a grant that ASA and DJG secured, to fund our course-based undergraduate research program.

REFERENCES

1. Kohgo Y, Ikuta K, Ohtake T, Torimoto Y, Kato J. (2008) Body iron metabolism and pathophysiology of iron overload. *International journal of hematology* 88(1), 7-15. <https://doi.org/10.1007/s12185-008-0120-5>
2. Nahon P, Sutton A, Rufat P, et al. (2008) Liver Iron, HFE Gene Mutations, and Hepatocellular Carcinoma Occurrence in Patients With Cirrhosis. *Gastroenterology* 134(1), 102-110. <https://doi.org/10.1053/j.gastro.2007.10.038>
3. Gallego Carlos J, Burt A, Sundaresan Agnes S, et al. (2015) Penetrance of Hemochromatosis in HFE Genotypes Resulting in p.Cys282Tyr and p.[Cys282Tyr];[His63Asp] in the eMERGE Network. *The American Journal of Human Genetics* 97(4), 512-520. <https://doi.org/10.1016/j.ajhg.2015.08.008>
4. Fracanzani AL, Fargion S, Stazi MA, et al. (2005) Association between heterozygosity for HFE gene mutations and hepatitis viruses in hepatocellular carcinoma. *Blood Cells, Molecules, and Diseases* 35(1), 27-32. <https://doi.org/10.1016/j.bcmd.2005.03.007>
5. Steinberg KK, Cogswell ME, Chang JC, et al. (2001) Prevalence of C282Y and H63D Mutations in the Hemochromatosis (HFE) Gene in the United States. *JAMA* 285(17), 2216-2222. <https://doi.org/10.1001/jama.285.17.2216>
6. Genetics Home Reference, US National Library of Medicine. HFE gene, <https://ghr.nlm.nih.gov/gene/HFE>. (Accessed Feb 2020)
7. dbSNP, National Center for Biotechnology Information. Reference SNP rs1799945, <https://www.ncbi.nlm.nih.gov/snp> (Accessed Feb 2020)

8. Feder JN, Penny DM, Irrinki A, et al. (1998) The hemochromatosis gene product complexes with the transferrin receptor and lowers its affinity for ligand binding. *Proceedings of the National Academy of Sciences of the United States of America* 95(4), 1472-1477. <https://doi.org/10.1073/pnas.95.4.1472>
9. Genetics Home Reference, US National Library of Medicine. Hereditary Hemochromatosis, https://ghr.nlm.nih.gov/condition/hereditary-hemochromatosis?_ga=2.1100785.1187020442.1551651207-1118969514.1551651207#genes. (Accessed, Feb 2020)
10. Goode MR, Cheong SY, Li N, Ray WC, Bartlett CW. (2014) Collection and extraction of saliva DNA for next generation sequencing. *Journal of visualized experiments : JoVE* (90), 51697. <https://doi.org/10.3791/51697>
11. Primer Blast, US National Library of Medicine, <https://www.ncbi.nlm.nih.gov/tools/primer-blast/> (Accessed Feb 2019)
12. Physics Department, College of Saint Benedict & Saint John's University. Data Entry, Exact 2×3 Contingency Table, http://www.physics.csbsju.edu/cgi-bin/stats/exact_form.sh?nrow=2&ncolumn=3 (Accessed Apr 2019)
13. Katsarou MS, Latsi R, Papasavva M, et al. (2016) Population-based analysis of the frequency of HFE gene polymorphisms: Correlation with the susceptibility to develop hereditary hemochromatosis. *Mol Med Rep* 14(1), 630-636. <https://doi.org/10.3892/mmr.2016.5317>
14. Mikhailova SV, Babenko VN, Ivanoshchuk DE, et al. (2016) Haplotype analysis of the HFE gene among populations of Northern Eurasia, in patients with metabolic disorders or stomach cancer, and in long-lived people. *BMC Genet* 17(1), 83. <https://doi.org/10.1186/s12863-016-0396-x>
15. Milman N, Byg K-E, Ovesen L, Kirchhoff M, Jørgensen KSL. (2003) Iron status in Danish women, 1984–1994: a cohort comparison of changes in iron stores and the prevalence of iron deficiency and iron overload. *European Journal of Haematology* 71(1), 51-61. <https://doi.org/10.1034/j.1600-0609.2003.00090.x>
16. Warne CD, Zaloumis SG, Bertalli NA, et al. (2017) HFE p.C282Y homozygosity predisposes to rapid serum ferritin rise after menopause: A genotype-stratified cohort study of hemochromatosis in Australian women. *Journal of Gastroenterology and Hepatology* 32(4), 797-802. <https://doi.org/10.1111/jgh.13621>
17. Pietrangelo A. (2010) Hereditary Hemochromatosis: Pathogenesis, Diagnosis, and Treatment. *Gastroenterology* 139(2), 393-408.e2. <https://doi.org/10.1053/j.gastro.2010.06.013>
18. Allen KJ, Gurrin LC, Constantine CC, et al. (2008) Iron-Overload–Related Disease in HFE Hereditary Hemochromatosis. *New England Journal of Medicine* 358(3), 221-230. <https://doi.org/10.1056/NEJMoa073286>
19. Kaczorowska-Hac B, Luszczek M, Antosiewicz J, et al. (2017) HFE Gene Mutations and Iron Status in 100 Healthy Polish Children. *J Pediatr Hematol Oncol* 39(5), e240-e243. <https://doi.org/10.1097/MPH.0000000000000826>

ABOUT STUDENT AUTHOR

Tyler Ferqueron graduated in May 2019 with a B.S. in Human Nutrition, Foods, and Exercise, science option, from Virginia Tech. He plans to pursue a M.D. or M.D./Ph.D.

PRESS SUMMARY

Hereditary hemochromatosis (HH) results in a faulty iron sensing mechanism and causes the human body to absorb too much iron from the diet. Stored excess iron in the body can lead to several health conditions. Certain genetic mutations have been shown to lead to this disease. One of these mutations (referred to as H63D) within the HFE gene was studied in students aged 20-23 years old in an upper-level metabolic nutrition course to determine how frequently it occurred, with what population ancestry it occurred in, and whether carriers of the variant gene had received an HH diagnosis. It was found that there was no relationship between having this mutation and having been diagnosed with HH in this population, likely due to the young age of the population. However, a significant correlation was found between those of European decent and being a H63D variant carrier, consistent with other published findings. Direct-to-consumer DNA kits similar to the one used in this study may be helpful in identifying HH before symptoms arise.

Meningococcal Meningitis in College Students at United States Universities

Mikafui Dzotsi*

College of Public Health, Ohio State University, Columbus, OH

<https://doi.org/10.33697/ajur.2020.007>

Student: dzotsi.2@buckeyemail.osu.edu*

Mentor: harris.44@osu.edu

ABSTRACT

The purpose of this study was to review current United States (US) university meningitis prevention and awareness efforts for college students with the objective of finding improved methods for meningitis control on college campuses. Meningococcal meningitis cases occurring amongst the students at 45 US universities, reported by the National Meningitis Association between 2013-2017, were reviewed. Apart from analyzing the incidence of meningococcal serotype cases and prevention protocols at the 45 US universities, interviews were conducted with university health directors and health center staff to assess the nature of meningitis control programs on college campuses. Of the 45 US universities reported between the years, 2013-2017, 20 universities had cases of *Meningococcal meningitis* serotype B (Men B) while 25 universities had cases of serotype A (Men A), C (Men C), W (Men W), or Y (Men Y). Among 80 cases across all US universities, there were 11 deaths for a case fatality rate of 11/80 (14%). While all universities adhere to state requirements of immunization against serotypes A, C, W and Y, the vaccine for Men B was only recently FDA approved and is not widely used. Further review of some university meningitis prevention and awareness efforts reveal a trend in more passive (e.g. posters, pamphlets, health portal guidance) approaches, while the uptake of active campaign efforts (e.g. vaccination drives, presentations) are not always prioritized until outbreaks occur.

KEYWORDS

Awareness; Epidemiology; Disease; Meningococcal; Prevention; Serotype; University; Vaccination

INTRODUCTION

Meningococcal meningitis is a highly contagious *Neisseria meningitidis* (*N. meningitidis*) bacterial infection (case fatality- 10-15%), that causes inflammation of the brain and spinal cord. Given the relatively high case fatality rate and disability ratio of 11-19%,¹ efficient surveillance methods and fast response efforts are a necessity. Infection of the meninges of the brain and spinal cord can cause severe chronic effects including nerve damage, paralysis, hearing loss and life-threatening septicemia.² Based on characteristics of *N. meningitidis* polysaccharide capsules, the bacteria can be classified into different serogroups. Although there are at least 13 knowledgeable serogroups, this study only focused on serogroups A, B, C, W-135, and Y. In the US, meningococcal Serogroups B and C account for 60% of meningococcal cases each year.³ While new strains of serogroup A meningococcus (Men A), serogroup C meningococcus (Men C), serogroup W meningococcus (Men W), and serogroup Y meningococcus (Men Y) all remain a threat, serogroup B carries high importance given its antigenic mimicry of human neural tissue antigens.⁴

Transmission of *N. meningitidis* is mainly through direct contact and droplets.⁵ *N. meningitidis* colonizes mucosal surfaces of the nasopharynx primarily through horizontal genetic exchange. To mitigate the transmission rate of *N. meningitidis*, a common method for treatment and control of the disease, is the use of antibiotics. Nevertheless, increasing virulence seen in *N. meningitidis*, is a direct result of evolved genetic mechanisms that have led to antibiotic resistance.⁵ As a result, there is no guarantee that antibiotics prescribed for treatment (cefotaxime or ceftriaxone)⁶ or even current meningococcal vaccines are effective. Transmission through droplets occurs by the diffusion of large- droplets through respiratory tract or throat secretions from host to host, including those of asymptomatic carriers.¹ Unvaccinated and immunocompromised individuals are the most susceptible.² Even though meningococcal disease is rather rare and on the decline, college students who often live in dormitories or close quarters, remain one of the most susceptible populations.⁷

Over the years, the World Health Organization (WHO) and Centers for Disease Control (CDC) have worked alongside state health departments to improve meningococcal disease vaccination programs. The meningococcal conjugate vaccines (*Menactra*, *Menveo*, and *MenHibrix*), instituted in the US in 2005, have continued to help decrease meningococcal disease cases and outbreaks by 80%.¹ The US Health Department recommends that all adolescents receive a dose of Men ACWY at 11 or 12 years old, followed by a second dose or booster at the age of 16. In addition, the Advisory Council on Immunization Practices (ACIP) recommends administration of Men ACWY booster vaccines at times of increased meningococcal incidence (e.g. first year

university students living in residence halls, high risk individuals).⁸ Vaccination proves to be the best prevention for disease, and as a result mass vaccination could be the best solution. Nevertheless, vaccination campaigns can be very expensive, as seen by the potential administration of 300,000-500,000 doses per year costing almost 1.8 million dollars.⁸ It is important to note, however, that before the use of the meningococcal conjugate (Men ACWY) vaccines in adolescents was recommended, the overall annual incidence of meningococcal disease had already decreased by approximately 64% between 1996 and 2005.¹ Thus, while vaccination may aid in decreasing rates of meningococcal disease, other preventive measures such as better sanitation and decreased crowding, have played a more critical role.⁸

In addition to addressing meningococcal A, C, W, or Y prevention, outbreaks of serotype meningococcal B at prestigious US universities such as Princeton University (New Jersey) and the University of California Santa-Barbara (California) campus between 2013-2014, inspired a push for serotype B vaccination. Individuals at increased risk for meningococcal serotype B are encouraged to take the serogroup B meningococcal vaccines (*Bexsero* and *Trumenba*).⁹ Although both vaccines have been deemed effective by the FDA, to date, there is still little data available on vaccine effectiveness against clinical disease or duration of protection against clinical disease.¹⁰

Even though cases of meningococcal disease also occur outside of university environments, factors within university settings make certain university populations more susceptible to contraction of the disease. In addition, living arrangements, also make university students vulnerable. According to the US News Annual Survey on college student living numbers, in 2015, on average, approximately 39% of university students lived on campus.¹¹ Although some students are supplied with single room accommodation, many students are forced to share a room with their peers. As a result, majority of these rooms are limited in space, which may allow for easy transmission of bacterial *N. meningitidis* droplets. Increased risk factors for meningococcal disease may also include social life and habits (e.g. drinking, more than one kissing partner, sharing drinks), HIV infection, or travel to places where meningococcal disease is common.¹² Over the past few years, cases and outbreaks have continued to occur at universities including cases at Rutgers University, University of Oregon and many others (see **Table 2**).

Even though meningococcal disease is on the decline as a whole,¹ this study works to identify incidence of cases and US university prevention methods, so as to decrease future cases.

METHODOLOGY

In this study, universities used as case studies were specifically universities that reported meningococcal cases to the National Meningitis Association (NMA) between March 2013 and November 2017.¹³ US university meningococcal cases reported to the NMA between these years, totaled 80 cases. University and state immunization requirements and outbreak protocols were compared and analyzed to determine effectiveness in preventing meningococcal serotypes (e.g. Men A, B, C, W, Y). In addition, university active or passive prevention efforts were also analyzed. Active approaches were defined as efforts that included campaign or vaccination drives and/ or health presentations. Passive approaches were defined by the presence of posters, information on health center portals, and distribution of pamphlets.

For this study 45 (NMA Official Meningococcal Outbreak Reported) US universities were analyzed, all spanning 22 US states (see **Table 1 and Table 2**). Of the 45 universities, 20 universities had cases of viral meningococcal B serotype, while 25 of the universities had cases reported as either Men A, C, W, or Y serotype cases. Meningococcal outbreak information for each university were retrieved from the NMA, in addition to the CDC and directly from university health centers. University meningitis immunization requirements were gathered through university health center websites and student mandatory immunization lists, while state meningitis immunization requirements were retrieved through the National Meningitis Association and State Health Department Immunization Requirement Records.

For further analysis of university prevention and awareness efforts, university health directors or health center staff (of the 45 universities analyzed), were all contacted and sent a recruitment letter to serve as case studies for the cases that occurred on their campuses. This was approved through exempt IRB permissions (00006378). Only ten university health directors or health center staff members were included in the study. These university health directors or health center staff members were from a series of seven US universities (University of Oregon, University of California Santa Barbara, Missouri University, Kalamazoo College, Palomar Community College, Georgia Tech, and West Chester University). All these universities were chosen as case studies. Each of the university health directors or health center staff members agreeing to partake in an interview were asked five questions over a 30-40 minute interview period. All interviews were conducted over phone conference between Summer 2016 and Winter 2017. To identify methods for improved prevention of meningococcal disease prevention and response on university campuses, international meningitis organizations and national university health organizations were also interviewed. Organizations interviewed included Meningitis Now, National Meningitis Research Foundation, American College Health Association, and the

National Meningitis Association. Each individual interviewed in this study chose to remain anonymous. To protect their rights and privacy no profiles will be released.

RESULTS

As shown in **Table 2**, between 2013 and 2017, 20 US universities had Men B cases, while 25 had non-serogroup B cases or unspecified serotype cases. Overall, this amounted to 80 cases across the 45 US universities, displayed in **Table 1**. Of all the 45 universities reported to have had meningococcal disease cases, five universities were reported to have had outbreaks. All outbreaks reported were caused by the spread of serotype B meningococcal disease. Out of all states with university meningococcal disease cases, it is evident that California had the highest number of Men B cases (N=16/52, 31%). Further, there were more Men B cases than non-serogroup B cases amongst Californian universities (see **Table 1**).

U.S. State	# of Cases of A, C, W, Y/ Unknown Serotype	# of Cases of B Serotype	Total Meningococcal Cases
Alabama	1	0	1
Arizona	1	0	1
Atlanta	1	0	1
California	4	16	20
Connecticut	1	0	1
Connecticut	2	0	2
Florida	1	0	1
Idaho	1	0	1
Illinois	1	1	2
Iowa	1	0	1
Massachusetts	1	0	1
Massachusetts	0	2	2
Michigan	0	1	1
Missouri	0	1	1
New Jersey	0	10	10
New York	2	0	2
North Carolina	2	0	2
Ohio	1	0	1
Oregon	0	12	12
Pennsylvania	5	2	7
Rhode Island	0	2	2
South Dakota	1	0	1
Virginia	1	0	1
Washington D.C.	0	1	1
Wisconsin	1	4	5
Totals	28	52	80

Table 1. Comprehensive Summary of Cases of Meningococcal Meningitis in College Students at United States Universities by State (2013-2017) ¹³

Protocol/ Vaccination Requirements

Based on much of the data collected from interviews and research literature, vaccination requirements amongst the majority of US universities affected by meningococcal disease cases between 2013 and 2017 were similar. According to data collected from specific university meningococcal disease protocol and immunization requirements, approximately 84% (N=38) of the 45 universities studied, required Men ACWY conjugate vaccination. It is important to note, however, that while some universities may have required the Men ACWY vaccination, it may not have been required for all the undergraduate student population. An example of this was University of Missouri where Men ACWY vaccination was only required for students living in university housing, ¹⁴ rather than the whole undergraduate student population. While all universities followed their state health policies and immunization requirements, not all universities had Men ACWY vaccination requirements.¹³ As a result, some universities chose to follow ACIP vaccination guidelines and require Men ACWY vaccination, regardless of whether their state required it or not.¹⁵ According to National Meningitis Association geographical maps, the requirement of meningococcal vaccination is most prominent amongst mid-western and eastern states.¹³ In addition, while some states require universities to mandate Men ACWY vaccination before entering college (66%, N=33), other states either require awareness and prevention education or have no requirements at all (see **Figure 1**).

Name of US University	State	Outbreak Year	# of Cases at University	Serotype
Auburn University	Alabama	2016	1	A,C, W, Y/ Unknown
Bucknell University	Pennsylvania	2017	2	A,C, W, Y/ Unknown
Cal Poly	California	1 case (2015) 2 cases (2017)	3	B
Dakota Wesleyan University	South Dakota	2015	1	A,C, W, Y/ Unknown
Drexel University	Pennsylvania	2014	1	B
Elon College	North Carolina	2017	1	A,C, W, Y/ Unknown
Georgetown University	Washington D.C.	2014	1	B
Georgia Tech *	Atlanta	2013	1	A,C, W, Y/ Unknown
Humboldt State University	California	2014	1	A,C, W, Y/ Unknown
John Tyler Community College	Virginia	2015	1	A,C, W, Y/ Unknown
Kalamazoo College *	Michigan	2013	1	B
Kutztown University	Pennsylvania	2014	1	B
Loyola University	Illinois	2013	1	A,C, W, Y/ Unknown
Marquette University	Wisconsin	2015	1	A,C, W, Y/ Unknown
Missouri University *	Missouri	UNKNOWN	1	B
Northeastern University	Massachusetts	2016	1	A,C, W, Y/ Unknown
Northern Arizona University	Arizona	2017	1	A,C, W, Y/ Unknown
Ohio University	Ohio	2017	1	A,C, W, Y/ Unknown
Oregon State University	Oregon	2017	5	B
Palomar Community College *	California	2014	1	B
Penn State University	Pennsylvania	2016	2	A,C, W, Y/ Unknown
Princeton University	New Jersey	2013	8	B
Providence College	Rhode Island	2015	2	B
Rutgers University	New Jersey	2016	2	B
Santa Barbara City College	California	2017	1	B
Santa Clara University	California	2016	3	B
San Diego State University	California	2014	1	B
Seminole State College	Florida	2014	1	A,C, W, Y/ Unknown
St. Ambrose University	Iowa	2016	1	A,C, W, Y/ Unknown
St. Mary's College	California	2015	1	A,C, W, Y/ Unknown
State University College of Oswego	New York	2013	1	A,C, W, Y/ Unknown
University of California, Santa Barbara *	California	2013	6	B
University of California, Berkeley	California	2017	1	A,C, W, Y/ Unknown
University of California, Davis	California	2015	1	B
University of California, Long Beach	California	2013	1	A,C, W, Y/ Unknown
University of Hartford	Connecticut	2014	1	A,C, W, Y/ Unknown
University of Idaho	Idaho	2014	1	A,C, W, Y/ Unknown
University of Illinois	Illinois	2017	1	B
University of Massachusetts, Amherst	Massachusetts	2017	2	B
University of North Carolina, Charlotte	North Carolina	UNKNOWN	1	A,C, W, Y/ Unknown
University of Oregon *	Oregon	2015	7	B
University of Rochester	New York	2015	1	A,C, W, Y/ Unknown
University of Wisconsin Madison	Wisconsin	2010 (1 case) 2014 (3 cases)	4	B
West Chester University*	Pennsylvania	2013	1	A,C, W, Y/ Unknown
Yale University	Connecticut	2016	2	A,C, W, Y/ Unknown

*Interviewed Universities

Table 2. Cases of Meningococcal Meningitis in College Students at United States Universities (2013-2017) ¹³

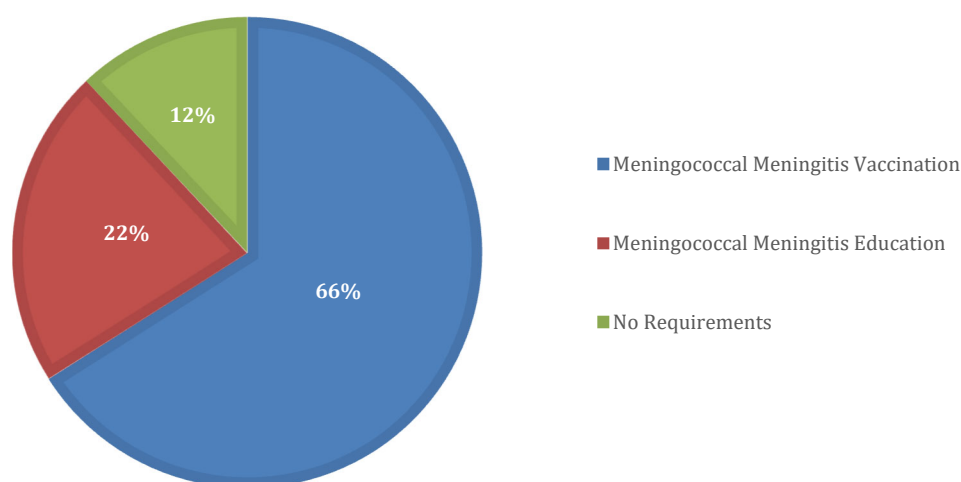


Figure 1. Meningococcal Meningitis State Health Department Prevention Requirement Distributions Across All United States, National Meningitis Association ¹³

Apart from requiring general meningococcal vaccination, knowledge of the types of meningococcal vaccination required is critical. While many universities required the Men ACWY vaccination, Men B vaccination was also recommended. Even though all the universities that were specifically interviewed and utilized as case studies recommended Men B vaccination to their students, many university health directors and staff stated that it would be unlikely for them to require Men B vaccination in addition to Men ACWY vaccination. Other explanations for only recommending the Men B vaccination might have been due to ACIP guidelines.¹ According to the ACIP, although several small outbreaks occurred on college campuses after 2013, college students in general were not at higher risk of meningococcal serogroup B disease than persons of the same age who are not college students.¹ Consequently, ACIP did not routinely recommend Men B vaccination for college students. However, college students could choose to receive Men B vaccination to reduce their risk of meningococcal serogroup B disease.¹⁶ Interviews with Meningitis Now, the Meningitis Research Foundation, and university health directors and staff revealed lower Men B vaccination uptake given personal out of pocket expenses and limited research on the effectiveness of Men B vaccination.

Although immunization protocols were a large focus for prevention efforts amongst case study universities, university health staff recognized that more needed to be done in terms of prevention and awareness for students, given evident case occurrences, irrespective of having immunization requirements or not (see **Figure 2**).¹⁷⁻²³ While 21% of case events amongst US universities (2013-2017), occurred at universities without Men B and/or Men ACWY vaccination recommendations or requirements, 76% of case events occurred at universities with Men B and/or Men ACWY vaccination recommendation or requirements.

Awareness/ Campaign Result

After conducting interviews with student health center directors and staff, amongst universities used as case studies, it became evident that the majority of awareness campaigns for the prevention of meningococcal disease on campuses is through passive approaches as opposed to active ones. Five of eight case study universities voiced using more passive approaches, while the remainder utilized additional active approaches (**Figure 3**).¹⁷⁻²³ Specific passive approaches tended to include sharing of meningococcal disease symptoms and vaccination information through online portals. When asked why more passive approaches were taken in regard to prevention and awareness of meningococcal disease, the majority (57%, N=4) of case study universities argued that passive approaches were cheaper and lowered the amount of competing health related information and messages already present within university communities.¹⁷⁻²³ In addition, four of case study universities stated that due to the fact that the incidence rate of meningococcal disease was significantly lower to that of other health conditions or risks such as depression amongst university students, an active approach of prevention and awareness would take away from other more prevalent issues.¹⁷⁻²³

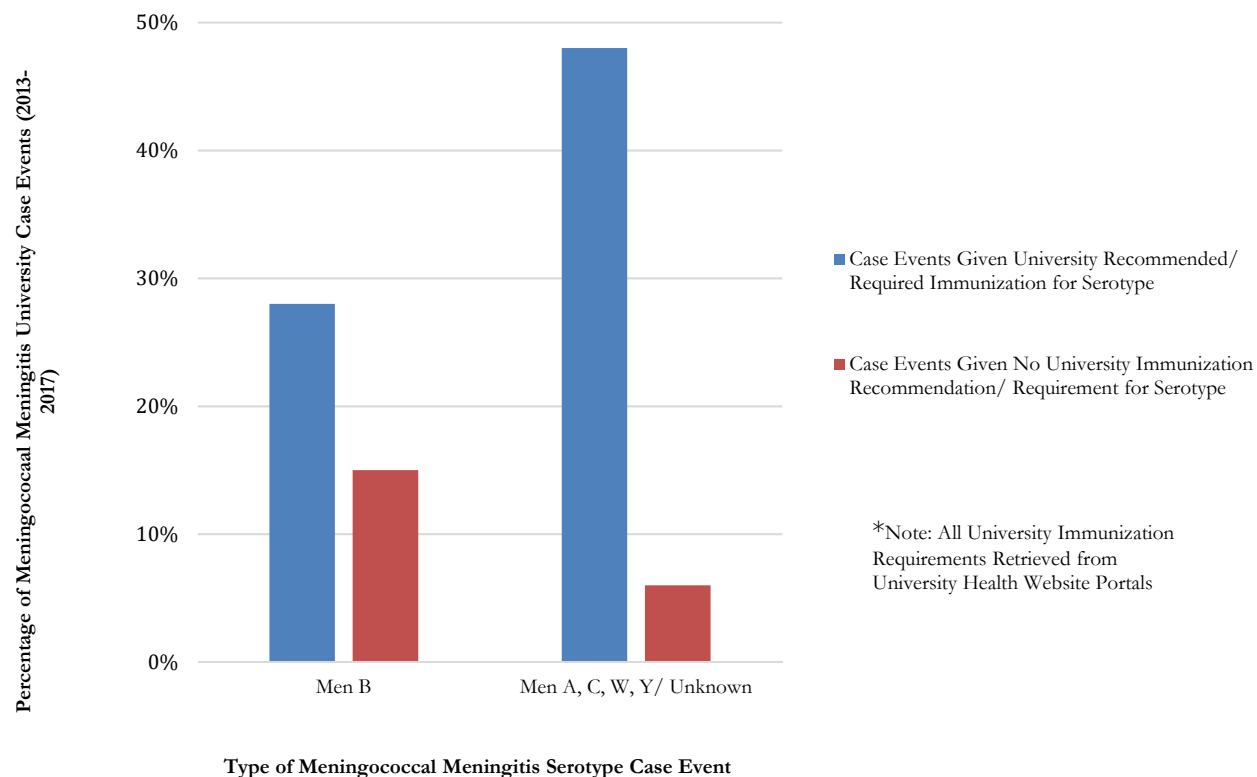


Figure 2. US University Meningococcal Meningitis Case Events (2013-2017) Given University Immunization Requirements¹³

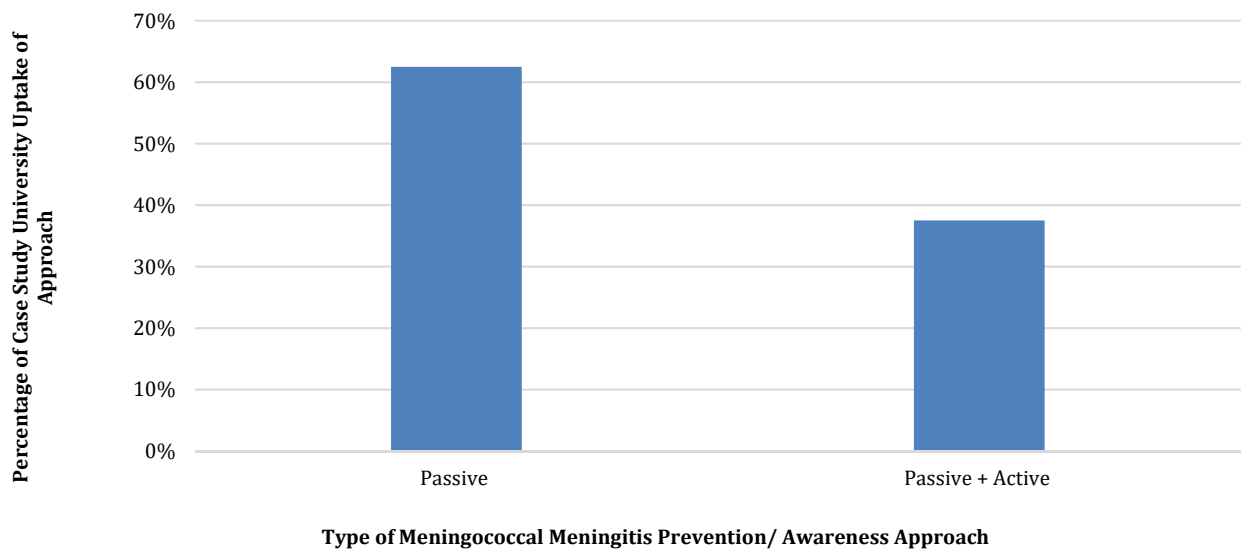


Figure 3. Distribution of Prevention and Awareness Approaches across US Case Study Universities. ¹⁷⁻²³

Active approach universities included Kalamazoo College, who post outbreak, made an effort to educate more of their student population on the risks and early symptoms of meningococcal disease. Specific methods used included active distribution of meningococcal disease cards and display of meningitis prevention videos and presentations to students at orientation, in addition to hosting a health forum. Apart from active approaches alone, Kalamazoo College also used passive approaches in spreading awareness by sending out an informative email during the summer months. According to some Kalamazoo College student health center staff, distribution of mail during summer months also targeted and informed parents, and not just students.¹⁷

In the aftermath of most outbreaks or cases occurring on campuses, the majority of university student health center staff stated that their universities were pushed to improve their preventative measures for meningococcal disease on their campuses. Specific improvements apart from mass vaccination campaigns at large outbreak institutions, included postings of CDC protocol and recommendations for meningococcal disease vaccination and prevention on electronic student health center portals. Although majority of universities stated that they took more of a passive approach in spreading awareness and preventing meningococcal disease, it is important to note that during or immediately after specific outbreaks or cases occurred, active responses were taken.¹⁷⁻²³

Apart from current approaches to awareness and prevention of meningococcal disease on university campuses, case study universities were also asked to list creative new approaches that they thought might be more effective. In response, 80 % (N= 6) of suggestions were targeted at spreading awareness about the specific symptoms of meningococcal disease. Specific strategies centered on symptom awareness included educating parents, posting flyers around campus, and contacting students through their emails.¹⁷⁻²³ Some suggestions included the posting of flyers not only in large public spaces like the student union, but also within entry ways of residence halls. In addition to spreading awareness of meningococcal symptoms to university communities, other suggestions included working with other outside providers and companies to not only improve student health care but also the stocking of specific ACWY Conjugate vaccines at student health centers. In stocking vaccination supplies, students would have easy vaccine accessibility upon consultation.

DISCUSSION

After gathering case specific data and interviewing health directors and health center staff, it was evident that the majority of university prevention and awareness programs were retrospective and passive rather than prospective and active. Due to money and the higher prevalence of other health concerns on campuses (e.g. depression, alcoholism, opiate use), active campaign efforts to control meningococcal disease were not a priority until outbreaks occurred on their campuses.¹⁷⁻²³

Although some university approaches to awareness and prevention were more passive, it did not mean that universities were not prioritizing the importance of meningococcal disease on their campuses. Results revealed that while all US universities were required to follow state health regulations, some universities went above and beyond in choosing to follow ACIP guidelines, in addition to providing awareness and vaccination education, before or after meningococcal meningitis case occurred have been present on their campuses. An example of this can be seen by the response efforts during the University of Oregon outbreak of 2015.

Between early January and June 2015, University of Oregon had an outbreak of serotype B *Neisseria meningococemia*. Although the university had not experienced quite as large an outbreak previously, as cases were confirmed, University of Oregon and state health department officials worked together to prevent further spread of the disease.²⁴ As soon as cases were confirmed, victims of the outbreak (see **Table 3**) were interviewed about their main contacts, so the university could identify susceptible students and staff. Overall the university identified more than 800 of its students and staff, believed to have been in contact with victims of the outbreak.²⁴ Individuals having close contact with victims for a least a four-hour cumulative period within a one-week period were also provided with preventative medications other than antibiotics to help further prevent the spread of disease.²⁵

Even though cases occurred gradually, University of Oregon staff and public health officials pushed for a mass vaccination campaign. The recent licensure of Men B vaccinations (*Bexsero* and *Trumenba*), meant that future cases of the serotype B *Neisseria meningococemia*, responsible for the university outbreak, could be prevented. After the sixth case was confirmed and public health officials realized that providing antibiotics would not fully prevent the spread of disease, University of Oregon parents were also contacted and encouraged to get their children vaccinated. While undergraduates and first year students living in dormitories were mainly targeted to receive vaccination, graduate students who lived in campus housing were also urged to become vaccinated.²⁶ By late February and early March of 2015, University of Oregon launched a mass vaccination campaign, where all 22,000 undergraduates would be encouraged to participate in a two-step vaccination regime. In order to cover costs of vaccination, the university worked with insurance companies to ensure complete vaccination coverage for their students. For uninsured students, University of Oregon health directors and staff attempted to get uninsured individuals on the Oregon Health Plan or provided

free shots from donated doses.²⁷ While vaccination was one of the best methods of prevention, marketing and awareness of the meningococcal disease became a necessity to encourage undergraduate students to become vaccinated. According to media coverage of the event, “around campus posters [featured] smiling student athletes showing off the arm where they were vaccinated and sporting a green or gold adhesive bandage with the school logo on it,” were one of many attempts to encourage vaccination.²⁸ In addition, short wait times combined with free swag and food, worked to encourage students to become vaccinated. Although the University of Oregon outbreak occurred between January and May of 2015, the university still attempts to spread awareness of the disease and its effects through online portals and information pages on their student health center website.²³

Case Number	Date*	Type of Meningococcal Disease Contracted	Demographics of Victim	Status
1	1/13/2015	<i>Neisseria meningococemia</i>	Residence Hall Resident	Recovered
2	2/2/2015	<i>Neisseria meningococemia</i>	Private Residence Resident	Recovered
3	2/8/2015	<i>Neisseria meningococemia</i>	UNKNOWN	Recovered
4	2/17/2015	<i>Neisseria meningococemia</i>	18 Years Old, Female, Acrobatics and Tumbling Team	Death
5	UNKNOWN	<i>Neisseria meningococemia</i>	UNKNOWN	Recovered
6	3/18/2015	<i>Neisseria meningococemia</i>	20 Year Old, lived in off campus fraternity, experienced flu-like symptoms	Recovered
7	5/29/2015	<i>Neisseria meningococemia</i>	52 Year Old father of University of Oregon student	Recovered

*Dates may not be the exact date victims started showing symptoms, but rather when the media was notified of confirmed cases occurring.

Table 3. Example Case Study of Meningococcal Meningitis: University of Oregon.²³⁻²⁸

Although large vaccination campaigns are effective (as seen by University of Oregon efforts), Men B vaccination was only licensed in 2015, explaining university trends to follow ACIP protocol and only recommend it to students. In this way, while some universities did not require Men B or sometimes Men ACWY vaccination, they still ensured that university students had access to vaccination programming. According to interviewed university health center staff, passive approaches to the prevention of meningococcal disease were more cost effective and less likely to cause fear amongst students.¹⁷⁻²³ However, if students were not aware of specific symptoms of the disease, there was the potential for college students to confuse meningococcal disease symptoms to the flu, or even perhaps a hangover.²⁹ According to a study conducted at Ohio University between 2008 and 2010, specific susceptibility factors on university campuses included individuals in Greek life, having more than one kissing partner, or drinking alcoholic beverages.²⁹ As a result, university culture and pre-existing conditions might have played a large role in not only contraction of meningococcal disease but also the identification and response time in seeking immediate care. Perhaps, this is where passive approaches can be made to be more effective.

Given the susceptibility of university students, who may live in dormitories, have multiple kissing partners, smoke, or frequently are victims of hangovers, it is important that universities are educating their university communities. Although meningococcal disease continues to be a rare disease, continued outbreaks of meningococcal disease reported to the National Meningitis Association between the years of 2013-2017, show that there are still approaches to prevention and awareness that can be improved upon. Universities that seemed to get good feedback from students and their university communities on meningococcal prevention efforts included that of Kalamazoo College. Reasoning behind positive feedback could include the fact that active meningococcal awareness and prevention efforts taken by Kalamazoo Health staff, involved more campus involvement. Specific events requiring more involvement included large health forums or orientation sessions that not only addressed meningococcal disease but also addressed other health and safety risks and offered resources to prevent them.¹⁷

In many ways, by involving more of the university community in organizing such events or presenting meningococcal disease information at events that reach a larger population of students, more university students might be more inclined to learn more about preventative methods and spread awareness about meningococcal disease. For a few universities in the United Kingdom, active methods such as spreading prevention and awareness strategies at orientation health information sessions and organizing health campaign weeks, have helped in educating university campus communities.³⁰ In comparison to many US universities, according to representatives from Meningitis Now and the Meningitis Research Foundation, some British universities have sought to improve their meningococcal prevention and awareness efforts by partnering with local meningitis organizations.^{29,30} One of

the largest set-backs with active prevention and awareness campaigns, voiced by case study universities, was cost. Nevertheless, by partnering with national meningitis organizations, universities could cut back on costs, and allow organizations to organize campaign programs and disburse awareness and prevention information to university students.

Another solution going forward, might be to include aspects of the Health Belief Model in prevention and awareness efforts on university campuses. In many ways, one of the most important parts of prevention and awareness is making sure that all parties involved, that being university students, health center workers, and state health department officials, all equally value and believe in the end goal of wanting to decrease the amount of meningococcal disease cases seen at university institutions. The Health Belief Model, a model designed to address the cognitive behaviors of reducing the outcome of a disease, would encourage individuals to acknowledge the need for improvement and engagement or action in regard to prevention efforts.³² Through the use of direct focus groups and belief and realization that meningococcal prevention can be improved on university campuses, protocols and active or passive approaches can be made to be more effective.

CONCLUSION

The findings from this research not only brought forth helpful insight into university meningococcal disease prevention and awareness efforts, but also further methods to improve them (e.g. university community participation and involvement in advocacy, health forums, targeted educational programming on meningococcal signs and symptoms, passive posters and incentives encouraging vaccination). Although meningococcal disease has been on the decline since before ACWY vaccines were legalized, efforts to improve meningococcal disease prevention and awareness on campuses have improved. Approaches to meningococcal disease prevention and awareness at most university institutions were passive in approach, however some universities still adopted active approaches. Even though active approaches are often sometimes more expensive, partnering with local health or meningitis organizations might be a way that universities can decrease cases occurring at their universities, while being cost effective. Even though many of the larger outbreaks occurring amongst universities between the years of 2013-2017 were serotype B, it should be noted that there were no Men B vaccinations available until 2015. Given the recent licensure of Men B it is understandable why universities followed ACIP and CDC guidelines of only recommending the Men B vaccine to susceptible populations. Going forward, analysis on the relation between the magnitude of state meningococcal prevention requirements (e.g. vaccination, education) and geographical case occurrences will be needed, so as to limit confounding variables and identify whether university prevention or state requirements have more of an effect on meningococcal disease incidence. In addition, more case studies will be needed to confirm the best preventative methods for meningococcal disease at U.S. universities. Nevertheless, case study universities utilized in this study provide a valuable case for passive approaches being more cost effective and more easily implemented into university communities. Meningococcal disease continues to be the most common amongst individuals in their late adolescence, which sadly includes majority of first year students entering university. While the effect of meningococcal disease can be detrimental to brain function and circulation, meningococcal disease does not have to be a threat to older adolescents if preventative methods and protocol are valued at university campuses.

ACKNOWLEDGEMENTS

The author thanks Dr. Randall Harris, OSU College of Public Health, Professor of Epidemiology. Many thanks is also given to all the interviewed US university health staff and international and national organizations, for their participation in the study.

REFERENCES

1. Centers for Disease Control and Prevention (March 2013). *Prevention and Control of Meningococcal Disease, Recommendation of the Advisory Committee on Immunization Practices (ACIP)*. Morbidity and Mortality Weekly Report, Center for Disease Control. 2 (62): RR02; 1-22. Retrieved from <https://www.cdc.gov/mmwr/preview/mmwrhtml/rr6202a1.htm>.
2. National Health Service (2016). *Meningitis*. Retrieved from <https://www.nhs.uk/conditions/meningitis/causes/>.
3. Center for Disease Control and Prevention. *Surveillance*. Retrieved from <https://www.cdc.gov/meningococcal/surveillance/index.html>.
4. Villena, R., Safadi, M., Valenzuela, M. T., Torres, J. P., Finn, A., & O'Ryan, M. (2018). *Global Epidemiology of Serogroup B Meningococcal Disease and Opportunities for Prevention with Novel Recombinant Protein Vaccines*. Human Vaccines & Immunotherapeutics. 14(5), 1042-1057. doi:10.1080/21645515.2018.1458175.
5. Rouphael, Nadine G., and David S. Stephens (2012). *Neisseria Meningitidis: Biology, Microbiology, and Epidemiology. Methods in molecular biology (Clifton, N.J.)* 799: 1-20. PMC. doi: 10.1007/978-1-61779-346-2_1.
6. El Bashir H, Laundry M, Booy R (2003). *Diagnosis and Treatment of Bacterial Meningitis*. Archives of Disease in Childhood. 88(615-620). doi:10.1136/adc.88.7.615.
7. Harrison LH, Dwyer DM, Maples CT, Billmann L (1999). *Risk of Meningococcal Infection in College Students*. JAMA. 281 (20), 1906-1910. doi:10.1001/jama.281.20.1906.

8. Advisory Committee on Immunization Practices (June 2000). *Meningococcal Disease and College Students*. Morbidity and Mortality Weekly Report, Center for Disease Control. 7(49): RR07; 11-20. Retrieved from <https://www.cdc.gov/mmwr/preview/mmwrhtml/rr4907a2.htm>.
9. MacNeil J.; Patton M. (2017). *Manual for the Surveillance of Vaccine-Preventable Diseases*. National Center for Immunization and Respiratory Diseases, CDC. Retrieved from <https://www.cdc.gov/vaccines/pubs/surv-manual/chpt08-mening.html>.
10. Kuhdari, P., Stefanati, A., Lupi, S., Valente, N., & Gabutti, G. (2016). Meningococcal B vaccination: real-world experience and future perspectives. *Pathogens and Global Health*, 110(4-5), 148–156. doi:10.1080/20477724.2016.1195072.
11. Friedman, Jordan (2016). *National Universities Where the Most Students Live on Campus*. US News and World Report. Retrieved from <https://www.usnews.com/education/best-colleges/the-short-list-college/articles/2016-11-15/11-national-universities-where-the-most-students-live-on-campus>.
12. Immunization Action Coalition (Feb 2018). *Meningococcal ACWY*. National Center for Immunization and Respiratory Diseases, CDC. Retrieved from http://www.immunize.org/askexperts/experts_meningococcal_acwy.asp.
13. National Meningitis Association. Retrieved from <https://www.nmaus.org/>.
14. University of Missouri Student Health Center (2018). *MU Immunizations Policy*. Retrieved from <https://studenthealth.missouri.edu/forms-policies/mu-immunizations-policy/>.
15. Advisory Committee on Immunization Practices (2011). *General Recommendations on Immunization --- Recommendations of the Advisory Committee on Immunization Practices (ACIP): Morbidity and Mortality Weekly Report*. Recommendations and Reports. 60(2): RR02; 1–64. Retrieved from <https://www.cdc.gov/mmwr/preview/mmwrhtml/rr6002a1.htm>.
16. Advisory Committee on Immunization Practices (October 2015). *Use of Serogroup B Meningococcal Vaccines in Adolescents and Young Adults: Recommendations of the Advisory Committee on Immunization Practices*. Morbidity and Mortality Weekly Report, Center for Disease Control: 64(41);1171-6. Retrieved from <https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6441a3.htm>.
17. Kalamazoo College Health Staff, personal communication, June 2017.
18. Westchester University Health Staff, personal communication, June 2017.
19. Palomar Community College Health Staff, personal communication, June 2017.
20. Missouri University Health Staff, personal communication, June 2016.
21. University of California Santa Barbara Health Staff, personal communication, July 2016.
22. Georgia Tech Health Staff, personal communication, July 2016.
23. University of Oregon Student Health Center Staff, personal communication, October 27, 2016.
24. KVAL News (Jan 2015). *UO Student Hospitalized with Potentially Deadly Infectious Disease*. KVAL. Retrieved from <https://kval.com/news/local/uo-student-hospitalized-with-potentially-deadly-infectious-disease>.
25. Herriman, Robert (Feb 2015). *University of Oregon outbreak: 3rd meningitis case reported*. Outbreak News Today. Retrieved from <http://outbreaknewstoday.com/university-of-oregon-outbreak-3rd-meningitis-case-reported-25954>.
26. Capitano B; Dillon K; LeDuc A; Atkinson B; Burman C (January 2019). *Experience Implementing a University-Based Mass Immunization Program in Response to a Meningococcal B Outbreak*. *Human Vaccines and Immunotherapies*. 3 (717-724). doi: 10.1080/21645515.2018.1547606.
27. Hammond, Betsy (Feb 2015). *University of Oregon meningitis vaccination will be largest in US since approval of new drug*. The Oregonian. Article Retrieved from http://www.oregonlive.com/education/index.ssf/2015/02/university_of_oregon_meningiti.html.
28. Hanrahan, Mark (March 2015). *6th Student in Oregon Infected with Meningitis Bacteria*. USA Today. Article Retrieved from <https://eu.usatoday.com/story/news/nation/2015/03/19/meningitis-outbreak-oregon/25049737/>.
29. Clark TA, Cohn AC, DiOrio M, Garcia J, Gaskell JR, King L, Machesky K, MacNeil JR, Mandal S, Messonnier NE, Mayer LW, Plikaytis BD, Quinn K, Schmink SE, Wang X, Wu HM (Aug 2013). *Prolonged university outbreak of meningococcal disease associated with a serogroup B strain rarely seen in the United States*. *Clinical Infectious Diseases*, National Center for Biotechnology Information.;57(3), 344-8. doi: 10.1093/cid/cit243.
30. Meningitis NOW, personal communication, June 22 2016.
31. Meningitis Research Foundation, personal communication, June 23 2016.
32. Wagner, A. L., Boulton, M. L., Sun, X., Mukherjee, B., Huang, Z., Harmsen, I. A., Zikmund-Fisher, B. J. (2017). *Perceptions of measles, pneumonia, and meningitis vaccines among caregivers in Shanghai, China, and the health belief model: a cross-sectional study*. *BMC pediatrics*, 17(1), 143. doi:10.1186/s12887-017-0900-2.

ABOUT THE AUTHOR

Mikafui Dzotsi is a recent graduate of The Ohio State University. Mikafui's interest in investigating the prevention and awareness of meningococcal disease on college campuses was sparked by her exposure to epidemiology as a Public Health Major during her undergrad. She is currently working towards completing her Master Degree in Epidemiology. She hopes that her degree will enable her to create sustainable solutions for the prevention of global health issues.

PRESS SUMMARY

Even though meningococcal disease is rather rare and on the decline, college students, who often live in dormitories or close quarters, remain one of the most susceptible populations. Recent outbreaks at some US universities have brought more attention to the importance of good vaccination policies and disease prevention efforts on university campuses. This study reviews the incidence of specific meningococcal meningitis serotypes occurring at US universities (2013-2017), in addition to university prevention and awareness efforts. Through means of this analysis, this study functions to identify more effective methods for meningitis control on college campuses.

Urban Wind Harvesting Using Flow-Induced Vibrations

Levon Ghabuzyan*, Christopher Luengas, & Jim Kuo

Department of Mechanical Engineering, California State University, Los Angeles, Los Angeles, CA

<https://doi.org/10.33697/ajur.2020.008>

Students: lghabuzyan@gmail.com, cluengas95@gmail.com

Mentor: jkuo7@calstatela.edu

ABSTRACT

The growing global interest in sustainable energy has paved the way to the rapid development of large-scale wind farms, consisting of dozens to hundreds of wind turbines. Although these large wind farms can generate enormous amount of power, they are also costly and require large areas of land or water, and thus are not suitable for urban environments. Smaller urban wind turbines have been developed for urban environments, but there are significant challenges to their widespread deployment. One of these challenges are their urban wind flows as they are strongly affected by complex building structures, producing highly turbulent flows. Any urban wind turbine would need to be designed to function efficiently and safely under these flow conditions; however, these unpredictable and turbulent winds can induce undesirable vibrations and cause early failures. Recently, bladeless wind turbines are gaining interest due to their reduced costs compared with conventional wind turbines such as the vertical-axis wind turbine and horizontal-axis wind turbine. These bladeless turbines convert flow wind energy into vibration energy, then converts the vibration energy into electricity. This paper examines the effects of force-induced vibrations on a cantilever beam system through wind tunnel experimentation. When fluid flows around a bluff body, periodic shedding of vortices may occur under the right conditions. The vortex shedding process creates an asymmetric pressure distribution on the body which causes the body to oscillate, known as vortex-induced vibrations. The purpose of the paper is to understand the factors affecting flow-induced vibrations and to improve wind energy harvesting from these vibrations. The first part of the paper focuses on wind tunnel experiments, by utilizing a cantilever beam configuration, conceptualized by previous research. Then, the experimental model was tested in different configurations, to determine the best setup for maximizing vibrations induced on the model. The long-term goal of the project was utilizing the model to optimize the system to improve efficiency of wind energy harvesting. The experimental results showed that the presence of an upstream cylinder will significantly improve the amplitude of vibration for energy harvesting, furthermore, the experiments showed that spacing in different directions also affect the amplitude of the vibrations. A two tandem cylinder system was used in this work, including a fixed rigid upstream cylinder and a downstream cylinder supported by a cantilever beam. Various configurations of these two cylinders in terms of spanwise and streamwise separation distances were studied and their maximum and root mean square displacements are reported for different wind speeds. Results showed that the presence of an upstream cylinder will significantly improve the amplitude of vibrations. This work verified that a wind energy harvester needs to consider the effects of wind speed and separation configuration of the cylinders in order to maximize the harvester's performance in urban environments.

KEYWORDS

Sustainable Energy; Energy Harvesting; Urban Environments; Bladeless Wind Turbines; Flow-Induced Vibrations; Cantilever Beam System; Wind Tunnel; Wake

INTRODUCTION

In the past century, the planet has witnessed an exponential growth in human population.¹ As the global population increases, so has the demands for energy and this trend is projected to increase with time.² In order to reduce the reliance on fossil fuels, many governments implemented policies by transitioning towards renewable energy sources.³ Wind energy is one of the fastest growing sources of renewable energy, which grew at 9.1% in 2018 to an overall capacity of 597 GW, covering close to 6% of the world's electricity demand.⁴ The primary form of harvesting wind energy is utilizing bladed turbines, as most of that energy is extracted through large wind turbines. However, conventional wind turbines are expensive, leading to the development of bladeless turbines that harvest wind energy through vibrations.^{5,6} These bladeless wind turbines harvest energy from flow-induced vibrations (FIV) and can be designed to extract from omnidirectional winds. Furthermore, winds in urban environments are affected by surrounding complex structures, inducing turbulence, making it difficult to harvest energy economically with conventional urban wind turbines.⁷ These turbulent and omnidirectional urban winds could lead to fatigue and thus early failures.⁸ In contrast, bladeless wind turbines are designed to harvest vibrational energy induced by flow instabilities.⁹ As wind flows through a bluff body, flow separations occur and vortices are induced in a periodic manner, creating pressure gradients in the transverse direction, orthogonal to the flow and undergoing vibrations.⁹ The vibrations induced by vortices are called vortex-induced vibrations (VIV).

A famous example of VIV is the Tacoma Narrows Bridge Collapse, which led to a catastrophic failure. However, instead of damping or reducing VIV, understanding this phenomenon is important to build an efficient bladeless wind turbine for urban communities through designing and testing a mass spring damper system in order to understand how VIV in wind could be harvested efficiently.

PAST EXPERIMENTAL WORK

There have been various studies focusing on harvesting wind energy through the use of VIV. In the work by Matsumoto,¹⁰ it was discussed that an excitation force was created through instabilities in the shear layer and that a cylinder can induce symmetrical vortices. Williamson *et al.*,¹¹ discussed the motion in which bluff bodies experience and how that motion generates different modes of vibration. Williamson *et al.*,¹¹ discussed the effects of vortex shedding on structures and how unsteady pressure causes bodies to undergo motion. In Sarpkaya's¹² review, he described the interactions between a bluff body and its surrounding fluid and highlighted the importance of flow separation around the cylinder as it creates eddies behind the body which causes a lift force to occur. The resulting transverse waves from multiple directions causes the bluff body to experience VIV, from which the mechanical energy then could be harvested.¹² Assi *et al.*,¹³ applied dimensional analysis to understand the vibration phenomenon and improve the efficiency of energy harvesting. In the work by Kim *et al.*,¹⁴ they conducted a series of tests using a mass spring damper system, using tripping wires to control the fluid boundary layer and thus VIV. In their experiment, it was found that when two cylinders are in tandem, the normalized distance between cylinders relative to their diameters will greatly influence amplitudes of vibration. Their experiment highlighted the importance of the spacing configuration between the cylinders. Papaioannou *et al.*,¹⁵ conducted wind tunnel and numerical experiments involving two cylinders in tandem, and they observed that large oscillations will occur if the natural frequency of the cylinder is close to the vortex shedding frequency. Furthermore, Mittal *et al.*,¹⁶ discussed how two cylinders in tandem will interact differently than just a single standalone cylinder. When the structural frequency of a cylinder is similar to the vortex shedding frequency, the cylinder experiences a lock-in, which causes the cylinder to vibrate at a much higher amplitude.¹⁶ In the experimental paper by Amandolese *et al.*,¹⁷ they examined the effects of VIV on a cylinder using a wind tunnel. The purpose of their experiment was to describe how the mass ratio of the structure affects the surrounding flow and how it causes the bluff body to experience oscillations.¹⁷ In the FIV work done by Kim,¹⁸ **Equation 1** shows that the motion of the downstream cylinder, y , can be modeled as,

$$m \frac{\partial^2 y}{\partial t^2} + c \frac{\partial y}{\partial t} + ky = F_L \quad \text{Equation 1.}$$

where m is the mass, c is the damping coefficient, k is the stiffness, and F_L is the lift force acting on the cylinder. The lift force induced by the upstream cylinder contributes to the damping of the system and could lead to oscillatory instability. This aerodynamic damping effect is described in detail in the work by Kim.¹⁸ In the work by Usman *et al.*,¹⁹ the authors conducted wind tunnel experiments with a cantilever beam system, where two cylinders are placed in tandem. The upstream cylinder is fixed, and the downstream cylinder was mounted on a cantilever beam where the flow was induced to cause vibrations on the cylinder.¹⁹ The setup by Usman *et al.*¹⁹ included an adjustable rail that would change the vertical distance between the downstream cylinder and upstream cylinder. This setup allowed them to study the effects of spacing to the behavior of the downstream cylinder. Their study also tested the specimen under different wind speeds, which changes the oscillation of the vibrating cylinder. Their configuration primarily focused on the effect on a fixed upstream cylinder placed directly upstream of the oscillating downstream cylinder. This paper focuses on the design conceptualized by Usman *et al.*, to validate and expand upon their work using a similar experimental setup. The system designed in this work will permit various configurations of two cylinders and allow them to move in the spanwise and streamwise directions in order to test and analyze the effects of configuration on vibrational motion. The next section discusses the methods and procedures of the experimental work.

METHODS AND PROCEDURES

In this section, an experimental setup based on the work Usman *et al.*, is discussed. Similar to the two tandem cylinder setup by Usman *et al.*,¹⁸ a rail system has been built to study the effects of positioning two cylinders in tandem. The upstream cylinder is fixed in position using a 3D printed mount. The position of the downstream cylinder can be adjusted in the spanwise and streamwise directions. The downstream cylinder is attached to a cantilever beam that is mounted on a 3D printed sliding mount which allows the cylinder to change position in the spanwise direction. The downstream cylinder contained an MMA 8452Q accelerometer that is connected to an Arduino Board, that will read the output data. **Table 1** has a detailed list of the parameters of the apparatus used in this work and its specifications are listed on **Table 2**. **Figure 1** illustrates the CAD model of the experimental setup and how the device was setup. Furthermore, this paper examines the effects of wind speed range of 2.5 m/s to 10 m/s (i.e. 2.5 m/s, 3 m/s, 4 m/s, 5 m/s, 6 m/s, 7 m/s, 8 m/s, 9 m/s, and 10 m/s), with similar spacing to the work proposed by Usman *et al.*; however, in addition to the streamwise spacing, the spanwise spacing was varied from 0D to 4D (i.e. 0D, 0.5D, 1D, 2D, 3D, and 4D). Furthermore, this work examines the behaviour of the downstream cylinder, when its spanwise spacing is changed as the spacing varies from 0D to 4D (i.e. 0D, 0.5D, 1D, 2D, 3D, and 4D).

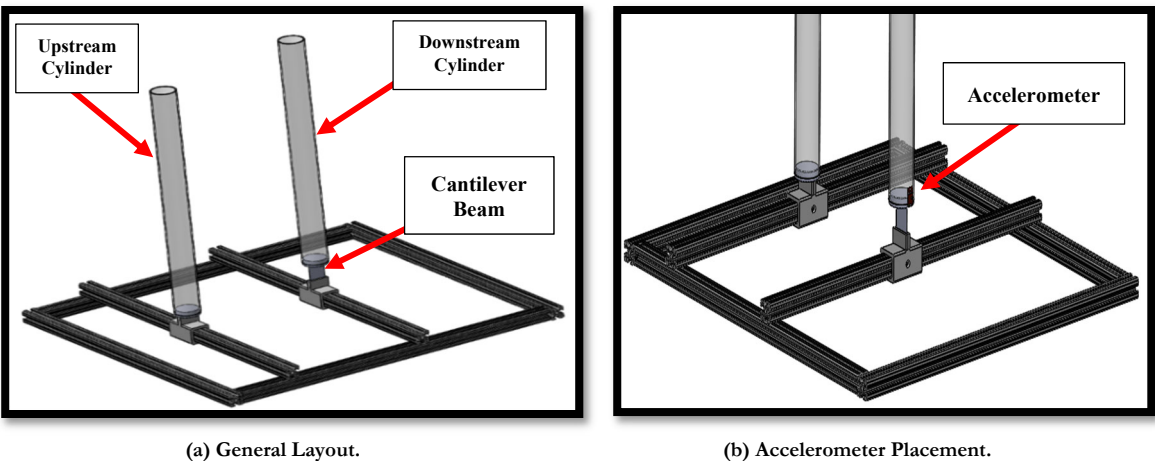


Figure 1. Proposed System Overview.

Parameters	Dimension Values
Downstream Cylinder Height	495.3 mm
Upstream Cylinder Height	533.4 mm
Diameter of Cylinders	44.45 mm
Downstream Cylinder Mass (With Mount Adapter)	80 g
Height of Cantilever Beam	107.95 mm
Width of Cantilever Beam	30 mm
Thickness of Cantilever Beam	0.3 mm
Material of Cantilever Beam	Aluminium
Young's Modulus of Cantilever Beam	69 GPa

Table 1. Apparatus Parameters.

Parameters	Capabilities
Dimension	3 mm x 3 mm x 1 mm
Data Scales	$\pm 4\text{ G}$
Output Data Rates	1.56 Hz to 800 Hz
Digital Output	8-bit and 12-bit
Axis	X, Y, Z Axes
Noise	$99\text{ }\mu\text{G}/\sqrt{\text{Hz}}$

Table 2. MMA 8452Q Accelerometer Specifications.

This paper observed the behaviour of the downstream cylinder at different spacings and wind speeds, and then find the best configuration for energy harvesting experimentally. Furthermore, this paper examined the behaviour of the downstream cylinder, using an accelerometer, which is connected to a microcontroller that is connected to a computer that outputs the reading through an interface. The acceleration data was used to study the amplitudes of oscillations of the downstream cylinder, at different flow speeds and configurations. **Figure 2** illustrates the basic overview of the experimental configuration and which parameters were changed in the experimental trials.

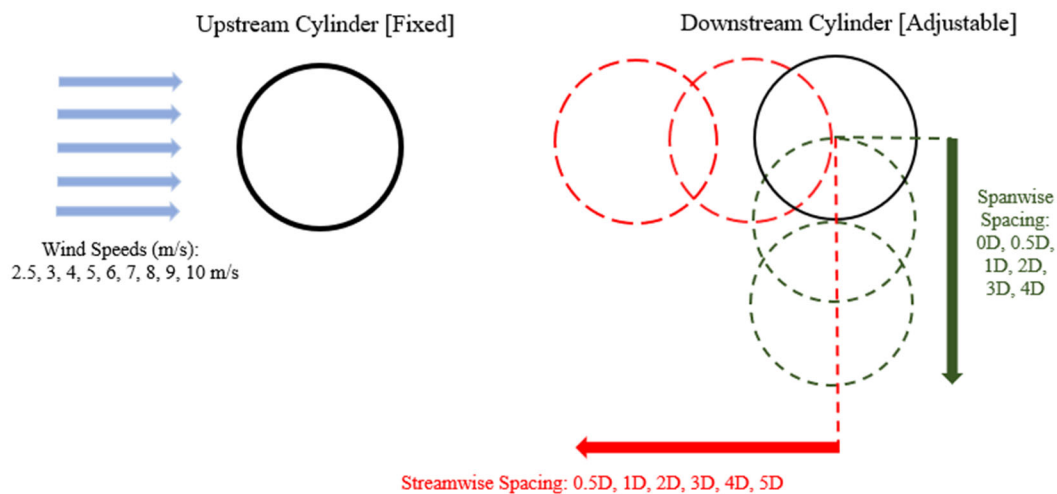


Figure 2. Experimental Configuration.

Experimental tests were conducted in the Aerodynamics Laboratory in California State University, Los Angeles using a subsonic wind tunnel. The wind tunnel has a test section area of $1.15 \times 0.71 \times 0.48$ (m) and is capable of reaching speeds up to 30 m/s where the wind flows from the right to left as shown in **Figure 3**. The experimental data of this study consists of an MMA 8452Q accelerometer and a TFI Cobra Probe to measure vibration and wind speed, respectively. The TFI Cobra Probe was attached to a mounting system positioned to measure wind speed from the inflow. As illustrated in **Figure 4a**, the downstream cylinder was connected to an aluminum sheet metal cantilever beam that was mounted to the proposed system. The MMA 8452Q accelerometer was attached onto the tip of the cantilever beam, to measure the acceleration in the direction perpendicular to the wind flow. The upstream cylinder was securely fixed onto a separate beam connected to the experimental platform. The wind tunnel setup is shown in **Figure 4b**.

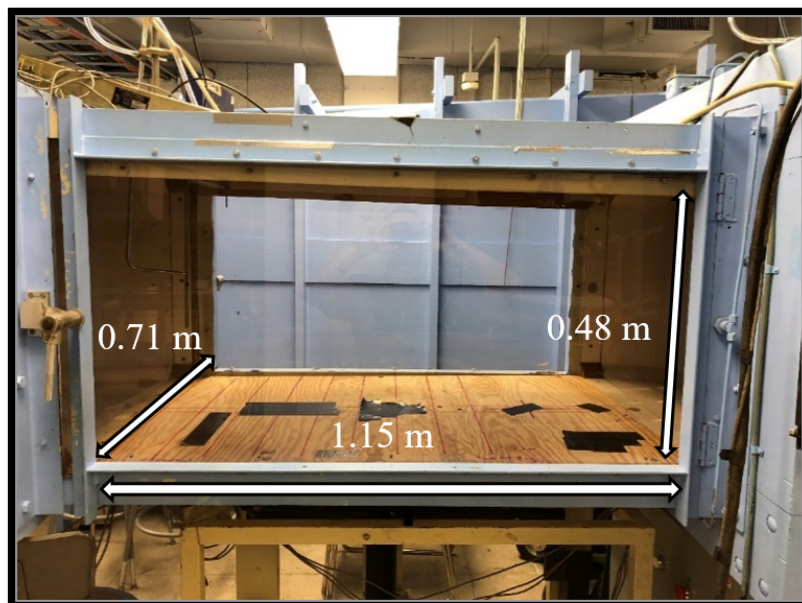


Figure 3. Wind Tunnel Cross-Sectional Area.

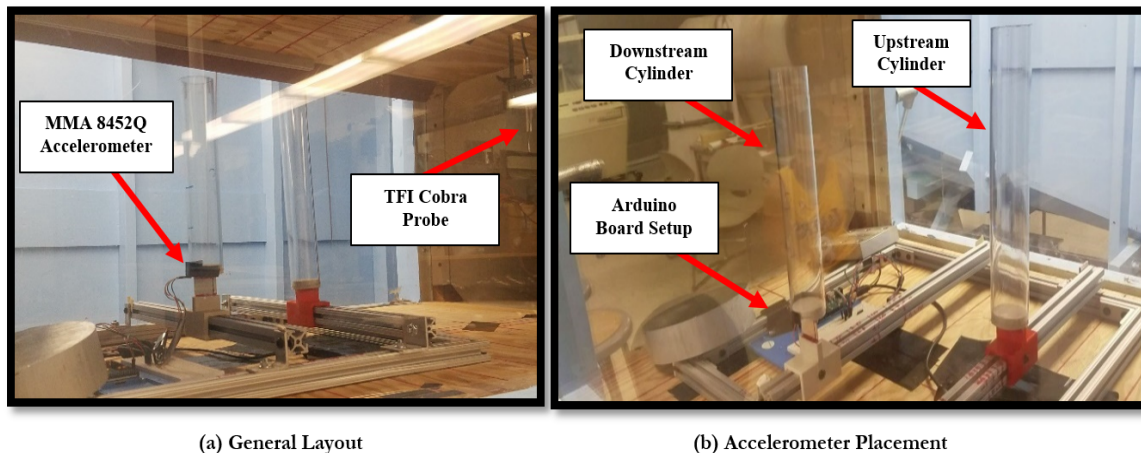


Figure 4. Wind Tunnel Setup.

A series of experiments were conducted to determine the optimal spacing between the upstream and downstream cylinders in both the streamwise and spanwise directions. The wind speed and acceleration were measured in all cases in order to determine the optimal displacement of the cantilever beam. Once the data was collected using the Arduino Board, it was then analyzed and complied to compare the performance of the downstream cylinder based on the different configurations and wind speeds. The next section presents the complied data and examine the results.

RESULTS AND DISCUSSION

Streamwise Configuration

This subsection discusses the findings of the streamwise experiments, where the distance between the tandem cylinders are varied in the streamwise direction. The MMA 8452Q accelerometer measured the acceleration of the downstream cylinder in unit of g (acceleration of gravity). Once the data was collected from the accelerometer, it was then uploaded to an in-house MATLAB code for post-processing. This code converted the acceleration data from g to SI units, then integration was performed to obtain velocity and displacement of the downstream cylinder. To perform integration over a fixed time period, the mean acceleration over that time period must be determined. Then the acceleration data should be subtracted by the mean value and then integration can be performed to obtain velocity. This procedure was then repeated for displacement. As displacement amplitudes can vary with time, two important quantities are root mean square (RMS) displacement and a maximum peak displacement.

Figure 5 illustrates the RMS displacement amplitude at various wind speeds as the downstream cylinder is aligned with the upstream fixed cylinder along the spanwise direction, with varying separation distances in the streamwise direction, ranging from 1.5D to 5D. As a benchmark, experiments without an upstream cylinder were conducted for comparison. As shown in **Figure 5**, the displacement of oscillation of the downstream cylinder increases with the presence of the upstream cylinder. The results show that at 3D separation, the downstream cylinder experiences larger oscillations at higher wind speeds, but the increase in displacement diminishes beyond 8 m/s or $Re = 13,656$. At wind speeds lower than 8 m/s, the RMS displacement at 1.5D and 2D separation seem to be the greatest, but quickly tapers off as wind speed increases. This phenomenon occurs because the strength of the oscillating vortices, downstream of the upstream cylinder, are functions of wind speed and downstream distance.

Furthermore, the results indicate that the 3D spacing generates the highest RMS displacement at the higher wind speeds, e.g. greater than 8 m/s, surpassing all other configurations tested. These results are similar to the findings from Usman *et al.*¹⁹, as both works show that a certain configuration outputs a greater magnitude of displacement, and that the results are repeatable. One phenomenon observed at 3D separation is that between wind speeds 3 m/s to 5 m/s, the vibrational amplitude increases rapidly. Then, between 5 m/s to 7 m/s, the rate of increase declines. Beyond 7 m/s, as flow speed increases the vibrational amplitude increases rapidly again as this may occur due to the interactions between the vortices from the upstream and downstream cylinders.^{20,21} Specifically, at the first cut-in speed of 3 m/s, vortex resonance or lock-in occurs, then at the second cut-in speed of 7 m/s, galloping takes over.²¹

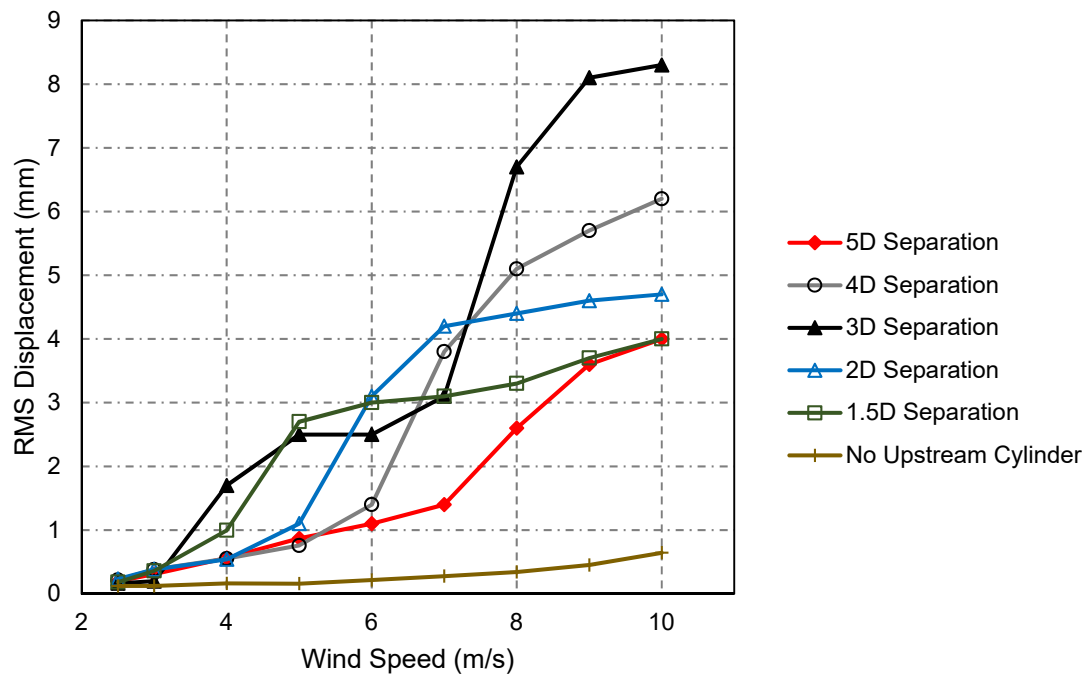


Figure 5. RMS Displacement for Streamwise Experiments.

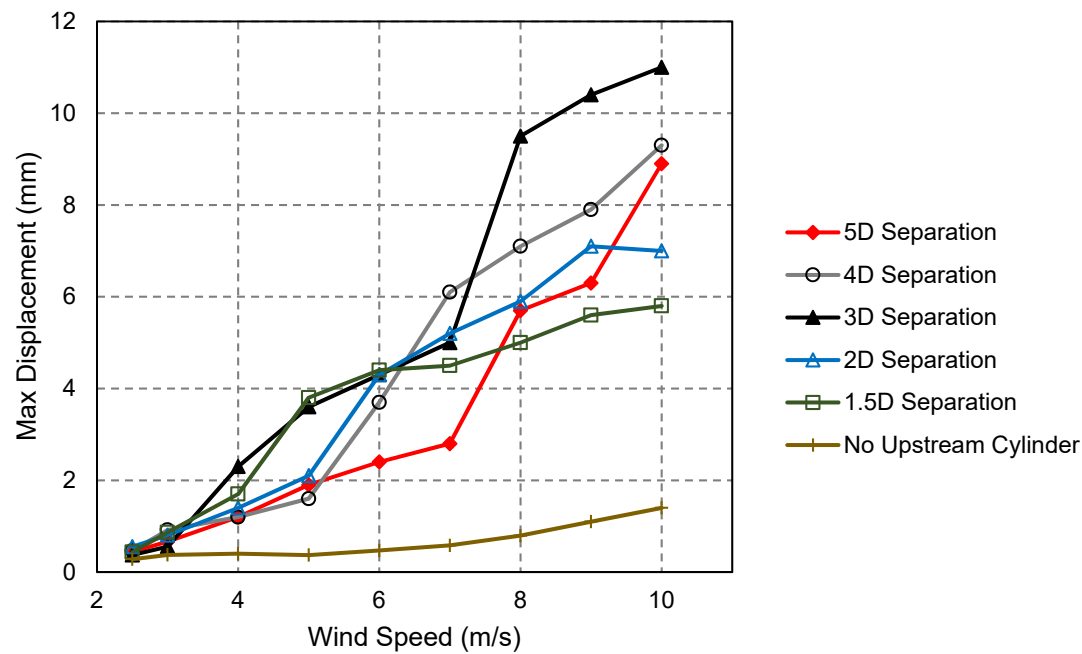


Figure 6. Maximum Displacement for Streamwise Experiments.

In addition to the RMS displacement, the maximum displacements are also analyzed, shown in **Figure 6**, as it displays similar trends as the RMS results, but showing key differences between the two sets of data. First, at 4D and 5D separation, the maximum displacement is larger relative to the trends seen in RMS displacement at greater wind speeds. That is, the trends observed for maximum displacement do not match that of RMS displacement at 4D and 5D separation. This is in contrast with 1.5D and 2D separation, which has trends similar to the RMS results. The maximum displacement observed at 3D separation is also consistent with the RMS displacement results. The next subsection will examine the results for the spanwise experiments and observe the behavior of the downstream cylinder, when the spanwise spacing is changed.

Spanwise Configuration

The following results compare between displacement amplitude and wind speed along the spanwise direction of the downstream cylinder as illustrated in **Figure 2**, when the streamwise separation distance is at 3D. Thus, spanwise separation distances were varied at 3D streamwise separation distance. Following the same procedures done in the streamwise direction by using MATLAB, both the RMS and maximum peak displacements were observed to see if the two displacements show consistency. When observing the RMS results, illustrated in **Figure 7**, the 0D spanwise separation (3D separation in the streamwise direction) experience the highest oscillating displacement. A decreasing RMS displacement trend is observed when separation increases, especially the 1D spanwise separation. The 1D separation experiences a significant decrease in displacement due to the wake effects caused by upstream cylinder, which caused the downstream cylinder to vibrate in a non-uniform pattern, thus resulting in the significant decrease in RMS displacement. This was also observed for the other configurations from 0.5D to 4D. It was further observed that when the spanwise separation reaches 4D, the effects of upstream cylinder on the downstream cylinder is negligible. These experiments demonstrated that 0D spanwise separation produces the highest displacement, but as soon as there are deviations in alignment, there will be significant decreases in RMS displacement.

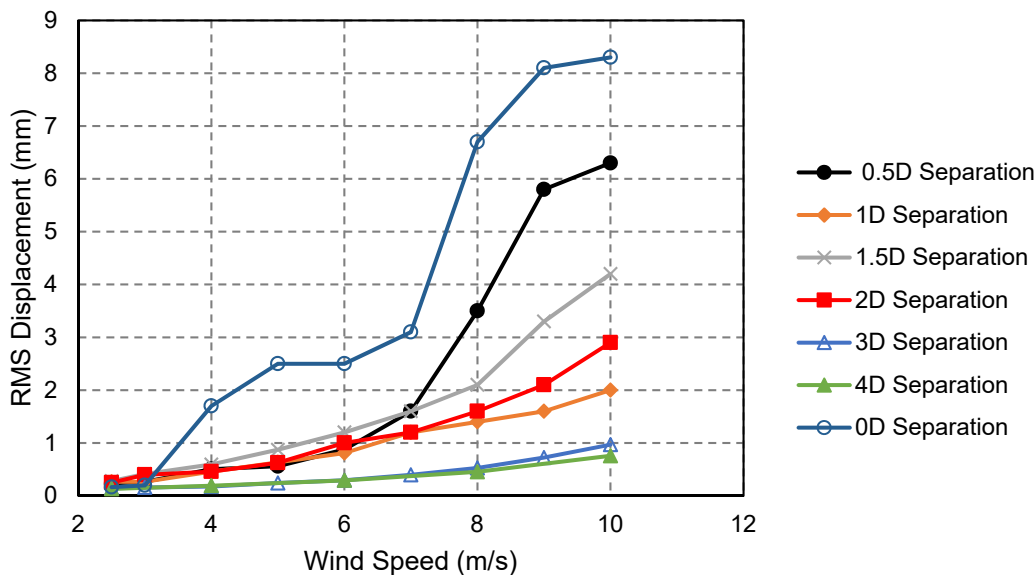


Figure 7. RMS Displacement for Spanwise Experiments.

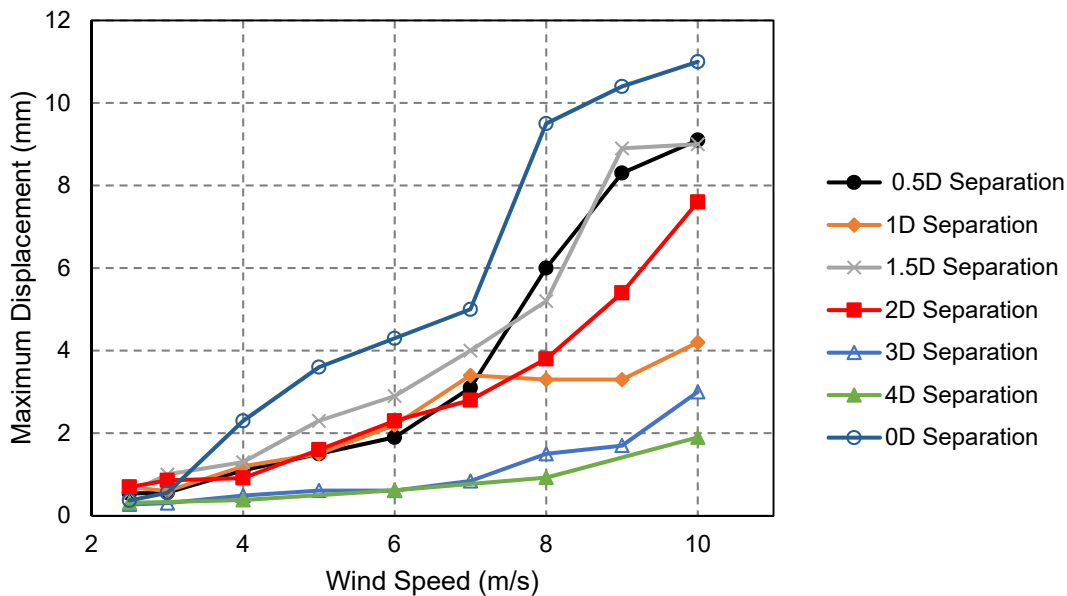


Figure 8. Maximum Displacement for Spanwise Experiments.

When observing the maximum displacement results, illustrated in **Figure 8**, the trends are different compared to the RMS results. The maximum displacement at 1.5D separation becomes similar to that of 0.5D separation. Another interesting observation is that for 1D separation at wind speed of 7 m/s, there is a drop in maximum displacement. This phenomenon is due to the effects of upstream wake and can be visually observed in the experiments, as the specimen was seen to vibrate in a non-uniform pattern. Moreover, the maximum displacement results for 0.5D to 4D, are consistent with RMS displacement, as they show similar trends.

CONCLUSIONS

A VIV harvester has been built to study the potential of harvesting wind energy from flow-induced vibrations. A two tandem cylinder system was used in this work, including a fixed rigid upstream cylinder and a downstream cylinder supported by a cantilever beam. Various configurations of these two cylinders in terms of spanwise and streamwise separation distances were studied and their maximum and RMS displacements are reported for different wind speeds. Results showed that the presence of an upstream cylinder will significantly improve the amplitude of vibration. Furthermore, 0D separation in the spanwise direction and 3D separation in the streamwise direction produces the highest displacement at wind speeds greater than 8 m/s. Any deviation from 0D spanwise separation will decrease displacement and as the separation increases to 4D, upstream effects become negligible and the downstream cylinder behaves as if there was no upstream cylinder. This work verified that a wind energy harvester needs to consider the effects of wind speed, separation configuration of the cylinders in order to maximize the harvester's performance in urban environments.

ACKNOWLEDGMENT

The authors thank the Department of Mechanical Engineering at California State University, Los Angeles, for providing access for the wind tunnel facility for experiments. The authors thank Luis Fernandez for his help and guidance in the fabrication of the model. The authors also thank Kevin Pan for his contribution in the experiments and for always willing to help. Lastly, the authors thank Daniel Lopez for his constant support and willingness to help at any time, but most importantly for being a great friend who never stopped encouraging the efforts of this project.

REFERENCES

1. Keinan, A., Clark A. G. (2012) Recent Explosive Human Population Growth Has Resulted in an Excess of Rare Genetic Variants, *Science* 336, 740–743. <https://doi.org/10.1126/science.1217283>
2. Edelenbosch, O.Y., Kermeli, K., Crijns-Graus, W., Worrell, E., Bibas, R., Fais, B., Fujimori, S., Kyle, P., Sano, F., Van Vuuren, D.P. (2017) Comparing projections of industrial energy demand and greenhouse gas emissions in long-term energy models, *Energy J.* 122, 701–710. <https://doi.org/10.1016/j.energy.2017.01.017>
3. Renewable Energy Policy Network for the 21st Century, Adib, R. (2015) *Renewables 2015 Global Status Report Distributed Renewable Energy for Energy Access*, https://cleanenergysolutions.org/sites/default/files/documents/gsr2015_cesc-webinar_9-10-2015.pdf (accessed June 2018)
4. World Wind Energy Association, Press Releases Statistics. (2019) *Wind Power Capacity Worldwide Reaches 597 GW, 50.1 GW Added in 2018*, <https://nwindea.org/blog/2019/02/25/wind-power-capacity-worldwide-reaches-600-gw-539-gw-added-in-2018/> (accessed July 2018)
5. Poulsen, T., Hasager, C.B. (2016) How Expensive is Expensive Enough? Opportunities for Cost Reductions in Offshore Wind Energy Logistics, *Energies* 9(6), 437. <https://doi.org/10.3390/en9060437>
6. Chizfahrn, A., Azadi Yazdi, E., Eghtesad, M. (2018) Dynamic modeling of vortex induced vibration wind turbines, *Renewable Energy* 121, 632–643. <https://doi.org/10.1016/j.renene.2018.01.038>
7. Walker, S.L. (2011) Building mounted wind turbines and their suitability for the urban scale-A review of methods of estimating urban wind resource, *Energy and Buildings* 43, 1852–1862. <https://doi.org/10.1016/j.enbuild.2011.03.032>
8. Sharpe, T., Proven G. (2010) Crossflex: Concept and early development of a true building integrated wind turbine, *Energy and Buildings* 42, 2365–2375. <https://doi.org/10.1016/j.enbuild.2010.07.032>
9. Wu, N., Wang, Q., Xie, X. (2013) Wind energy harvesting with a piezoelectric harvester, *SMS*.22, <https://doi.org/10.1088/0964-1726/22/9/095023>
10. Matsumoto, M. (1999) Vortex shedding of bluff bodies: A review, *J Fluid Struct.* 13, 791–811. <https://doi.org/10.1006/jfls.1999.0249>
11. Williamson, C. H. K., Govardhan, R. (2004) Vortex-induced vibrations, *Annu. Rev. Fluid Mech.* 36, 413–455. <https://doi.org/10.1146/annurev.fluid.36.050802.122128>
12. Sarpkaya, T. (2004) A critical review of the intrinsic nature of vortex-induced vibrations, *J. Wind Eng. Ind. Aerodyn.* 19, 389–447. <https://doi.org/10.1016/j.jfluidstructs.2004.02.005>
13. Assi, G.R.S., Meneghini, J.R., Aranha, J.A.P., Bearman, P.W., Casaprima, E. (2006) Experimental investigation of flow-induced vibration interference between two circular cylinders *J. Wind Eng. Ind. Aerodyn.* 22, 819–827. <https://doi.org/10.1016/j.jfluidstructs.2006.04.013>

14. Kim, S., Mahbub Alam, Md., Sakamoto, H., Zhou, Y. (2009) Flow-induced vibration of two circular cylinders in tandem arrangement. Part 2: Suppression of vibrations, *J. Wind Eng. Ind. Aerodyn.* 97, 312–319. <https://doi.org/10.1016/j.jweia.2009.07.003>
15. Papaioannou, G.V., Yue, D.K.P., Triantafyllou, M.S., Karniadakis, G.E. (2008) On the effect of spacing on the vortex-induced vibrations of two tandem cylinders, *J. Wind Eng. Ind. Aerodyn.* 24, 833–854. <https://doi.org/10.1016/j.jfluidstructs.2007.11.006>
16. Mittal, S., Kumar, V. (2001) Flow-induced oscillations of two cylinders in tandem and staggered arrangements, *J. Wind Eng. Ind. Aerodyn.* 15, 717–736. <https://doi.org/10.1006/jjls.2000.0376>
17. Amandolèse, X., Hémon, P. (2010) Vortex-induced vibration of a square cylinder in wind tunnel, *Comptes Rendus Acad. Sci.* 338, 12–17. <https://doi.org/10.1016/j.crme.2009.12.001>
18. Kim, H. (2009) Mechanism of wake galloping of two circular cylinders, *A Dissertation Submitted to Nagoya University*
19. Usman, M., Hanif, A., Kim, I., Jung, H. (2018) Experimental validation of a novel piezoelectric energy harvesting system employing wake galloping phenomenon for a broad wind spectrum, *Energy J.* 153. 882–889. <https://doi.org/10.1016/j.energy.2018.04.109>
20. Zdravkovich, M.M. (1988) Review of interference-induced oscillations in flow past two parallel circular cylinders in various arrangements, *J. Wind Eng. Ind. Aerodyn.* 28. (1-3) 183-199. [https://doi.org/10.1016/0167-6105\(88\)90115-8](https://doi.org/10.1016/0167-6105(88)90115-8)
21. Bokaian, A., Geoola, F. (1984) Wake-induced galloping of two interfering circular cylinders, *J. Fluid Mech. J.* 146. 383-415. <https://doi.org/10.1017/S0022112084001920>

ABOUT STUDENT AUTHORS

Levon Ghabuzyan and Christopher Luengas just completed their Bachelor of Science in Mechanical Engineering at the California State University, Los Angeles. Currently both authors are enjoying their summer break and studying for the Fundamentals of Engineering Exam (FE) as both are planning to pursue their Master of Science in Mechanical Engineering as Levon will continue his education at California State University, Los Angeles, this upcoming Fall 2019 and Christopher will return in Fall 2020.

PRESS SUMMARY

The expansion of urban area has brought about effects on the winds that pass through and the need for more energy, highlights an opportunity for wind energy. These urban winds can be used to harvest energy by the means of vortex-induced vibrations. The purpose of the paper is to understand the factors affecting VIV and to improve energy harvesting from these vibrations, by using a cylinder-cantilever beam system to determine the most efficient setup.

Exploring the Relationship between Dystopian Literature and the Activism of Generation Z Young Adults

Aysa Jerald

Department of English, University of Georgia, Athens, GA

<https://doi.org/10.33697/ajur.2020.009>

Student: aij81511@uga.edu*

Mentor: ekraft@uga.edu

ABSTRACT

Some recent research has posited that the independent and revolutionary traits of Generation Z can be traced to the circumstances of their births, specifically the 9/11 attacks and the Great Recession. While there has been research examining the effect of these events on the type of behavior Generation Z exhibits towards political and societal issues, there has been little research that examines the literary culture in which they grew up. Did popular dystopian works such as *Catching Fire* by Suzanne Collins (2009), *Divergent* by Veronica Roth (2011), and *The Maze Runner* by James Dashner (2009) have an impact on their political identities and behaviors? This paper examines that question by using a mixed method approach: a public questionnaire, thirteen in-depth interviews with a select group of Generation Z students from the University of Georgia, and direct content analyses of the key works under consideration. This study argues that the relationship between dystopian literature and young adult activism may offer insight into the ways literature can be used as a revolutionary tool. This study also hopes to add to the literature exploring the characteristics of Generation Z and the significance dystopian literature may have not only on a young adult's thoughts but also their actions.

KEYWORDS

Dystopian Literature; Dystopian Literary Media; Generation Z; Youth Activism; Literary Influence; Activist Typology; Aspects of Literary Response: A New Questionnaire; College Students; *Divergent*; *Catching Fire*; *The Maze Runner*; Literary Culture, The Hunger Games

INTRODUCTION

Young adult activism is not a foreign concept to the United States. In the 1960s, the Civil Rights movement and the Vietnam War draft drew the attention of many young objectors, leading to the development of historical protests, marches, and organizations. Nearly sixty years later, with issues involving race and politics still prevalent, young adults are still actively on the move. This time, however, Generation Z stands at the forefront, armed with technological confidence and mastery.

Generation Z

Generation Z has no precise starting date. However, a rough estimate places the members to be born between 1996 and 2012. Marked with a vague concept of the American Dream, Generation Z cannot truly remember a time period before terrorism and environmental concerns existed in the world. Yet despite the dark stereotypes that surround the technologically absorbed generation, they are very interested in political and societal issues. For example, research has shown that because they do not trust elected officials, their feelings about the United States are pessimistic and dismal. However, they still believe voting is important due to strong concerns regarding inequality and the direction of the economy.¹

Only recently has documentation shown Generation Z finally beginning to amplify their voices to make a difference, amidst the gun violence, gender and racial inequality, and other injustices of the current times.² The question of where this motivation stems from has vaguely been answered, as some researchers suggest that the growing use of social media has proved itself an easier basis for activism.³ However, the literature lacks any other explanation and ignores the possibility of literature and/or literary media as a motivator as well.

As the first wave of Generation Z entered adulthood, the popularity of dystopian literature increased in parallel. A dystopia, first coined in the 1700s, is an imagined society depicted as undesirable, evil, or oppressive.⁴ The most recognizable dystopian literature within recent years has been found primarily within the young adult genre and has since been successful in the literary market. According to an article published in 2017 by *The New Yorker*, dystopian novels seem to be in a "Golden Age."⁴ Its

attraction may be due to the reader's ability to relate to the characters, and the material being synonymous with the issues of the present.

The purpose of this study is to explore a possible connection between the activism of Generation Z young adults and the popularity of dystopian themes in literature and the literary media. It attempts to discover if dystopian literature has a relationship with Generation Z's mannerisms, beliefs, or values. In general, this research project re-evaluates the importance of literature and the power the Humanities may have in shaping the attitudes and behaviors of generational readers and film-goers.

Readers' Dilemma

Literature's relationship to social action may begin with examining literature's relationship to cognitive development. In 2017, a study on "readers' dilemma" examined how fictional media could shape a child's behavior. Readers' dilemma occurs when a reader must determine if a piece of information found in a fictional text or media can also apply to the real world.⁵ It is the process of looking past the fantastical or implausible elements of fiction to discover the realistic implications. Once the deeper theme or motif of the fiction is found, the reader may be indifferently or directly affected. If directly affected by the theme, the research suggested that the reader's real-world behavior was subject change.⁵ Thus, fictional media was seen not only as a means to entertain, but also a means to educate or enlighten.

Literary Influence

Psychiatrists and educated professionals have also seen and admitted to literature's influence on cognitive and social change. Based on the idea that fictional novels can provoke strong feelings of compassion and empathy within the reader, several professors have admitted to the potential use of those emotions in eliciting action. For example, results from a survey of 22, 562 professors at 372 colleges found that undergraduate educators felt certain novels involving adversity could cultivate compassion, thus cultivating enlightenment and action.⁶ Specifically, they believed reading fictional texts combating social injustice could increase a student's likelihood to fight against a similar injustice in the real world, due to feelings the students inhabited while reading about victims of calamity. Therefore, fictional texts were seen to have the same power as factual texts, in that they both could inform the reader on previously unknown concepts and beliefs. If reading a newspaper could inspire change, then it was possible a speculative fiction novel could as well.

Dystopian Literary Influence

Specifically, dystopian literature is the genre in question for this project. Can dystopian literature have a relationship with social change? While the answer has yet to be confirmed, previous research suggests a possible relationship. In an article published by Melissa Ames in 2013, it is proposed that educators should begin using dystopian literature to engage apolitical students.⁷ She first argues that dystopian literature may be used as a coping mechanism for modern-day students, who grew up with a more than understandable definition of terrorism, hence the 9/11 attacks.⁷ Despite its dark content, modern dystopian literature appeals to young adults because they can relate to the metaphorical world. Themes like terrorism, dominion, isolation, and climate change found in dystopian literature can also be found in the real world. The young adult reader can connect with the characters in the novel, who suffer through the same things that they do. The novels are a type of "emotional security" blanket, which allows young adults to bear with the trauma of the real world.⁷

Ames also argues that reading this material may spark the revolution the authors intended to ignite within young adults. She advocates for the teaching in the classroom of social-political thematic undertones found in the novels so that students can better understand the issues portrayed.⁷ Then students may not only be comforted but also enlightened. In hopes that their enlightenment may lead to social action, Ames hypothesizes the positive relationship between the dystopian trend and social-political engagement, especially within young adults.⁷

In addition to exploring the typology of Generation Z activism, this research study will also explore Melissa Ames' hypothesis. The following research questions will be addressed:

1. What is Generation Z's orientation towards literary texts?
2. Is there a relationship between the typology of Generation Z young adult activists and the typology of activists seen in popular dystopian novels?
3. Is there a relationship between dystopian literature and the activism of Generation Z young adults?

METHODS AND PROCEDURES

To find the answers to the research questions, a mixed method study was conducted, utilizing a Sequential Exploratory Design. The qualitative data were used to further explain the quantitative findings, which served as the forefront of the data collection. Second, a direct content analysis was conducted on three dystopian novels to discover a theoretical comparison.

Quantitative Procedures

For the quantitative portion of this mixed method study, four sections of the *Aspects of Literary Response: A New Questionnaire* were used.⁸ Prior to distribution, the proposed research and survey instrument received approval from the University of Georgia Institutional Review Board (IRB). SPSS statistical software was used to analyze the data.

The survey was administered through Qualtrics to Generation Z participants from the University of Georgia. Permission to assist with the recruitment, data collection, and to utilize data from this study was granted by the University of Georgia's Division of Academic Enhancement and Department of English. All Generation Z students were recruited through classroom presentation sampling or through email that included a hyperlink to the online questionnaire. The classrooms were provided by the University of Georgia's Department of English and Division of Academic Enhancement faculty members/professors who had agreed to participate in this study after being contacted by email. The list of student emails was provided by the university club/journal entitled The Bell and the TRIO McNair Scholars Program, as both organizations had agreed to their members' participation in this study after also being contacted by email. Before beginning the questionnaire, students were provided a consent form in which they had to agree or disagree to participating in the survey. The students' identities remained anonymous, and they were free to quit the questionnaire at any time. The questionnaire lasted approximately ten to twelve minutes. Students were offered no incentive but thanked for their willing participation.

Aspects of Literary Response: A New Questionnaire

The *Aspects of Literary Response: A New Questionnaire* was used to inquire about Generation Z's views, opinions, and perceptions of fictional literature/literary media. The questionnaire has seven subscales; however, only four of these subscales were used. Those scales were Insight, Empathy, Imagery Vividness, and Rejection of Literary Values. Insight contained fourteen questions, Empathy contained seven questions, Imagery Vividness contained nine questions, and Rejection of Literary Values contained nine questions. Each question was measured on a five-point Likert Scale, where 1 = "Not True At All," 2 = "Slightly True," 3 = "Moderately True," 4 = "Very True," 5 = "Extremely True." The Insight section pertained to the degree to which the reader/viewer became enlightened about things they did not previously know either about themselves or the world. Statements included, "Reading literature makes me sensitive to aspects about my life that I would usually ignore" and "I often find my shortcomings explored through characters in literary texts." The Empathy section pertained to the degree to which the reader identified with the characters. Statements included, "Sometimes I feel that I've 'become' a character I've read about in fiction" and "When I read fiction, I often think of myself as one of the characters in the story." The Imagery Vividness section pertained to the degree to which the literary world became vividly present to the senses. Statements included, "I can readily visualize the persons or places described in a novel or short story." Finally, the Rejection of Literary Values section referenced the participant's rejection of literature's influence altogether. Statements included, "I think people should spend less time talking or writing about literature." All four sections were used to answer the first research question and gain an understanding of Generation Z's overall literary opinions and cognizance.

Quantitative Participants

A total of a hundred Generation Z students from the University of Georgia participated in taking the questionnaire. Of the hundred students, 70% were female and 30% were male. The sample size consisted of 50% White students, 21% African American students, 6% Hispanic/Latina students, 16% Native American/American Indian/Asian students, and 7% of students categorized as Other. The Generation Z age range consists of those twenty-three and below; however, the sample size primarily consisted of 46% eighteen-year-olds, 24% nineteen-year-olds, 17% twenty-year-olds, and 13% twenty-one-year-olds and above.

Qualitative Procedures

Of the hundred participants who took the ALR questionnaire, thirteen students volunteered to participate in the interview portion of this study. The participants were Generation Z student volunteers who took the ALR questionnaire prior to the interview. Permission to assist with the recruitment, data collection, and to utilize data from this study was also granted by the University of Georgia's Division of Academic Enhancement and Department of English. All students were recruited via classroom presentation or email where a sign-up sheet to participate was handed out/attached. Before beginning the interviews, students were provided a consent form in which they had to agree or disagree to participating in the interview and being recorded. The interviews lasted approximately thirty to forty-five minutes. They were recorded using the Voice Memos application on an iPhone and instantly transferred to a flash drive for safe-keeping. Students were offered no incentive but thanked for their willing participation.

Participants were asked a total of sixteen questions broken up into three different sections. The first section inquired about the student's activism and social-political opinions. Questions included, "Are you personally concerned or indifferent about social-political issues" and "If you are an activist, what are your reasons for being one?" The second section inquired about the participant's literary perceptions. In hopes to understand their *Aspects of Literary Response: A New Questionnaire* responses in relation to dystopian literature specifically, questions included, "How do you feel when you are exposed to dystopian themes in literature or literary media?" and "Would you say you become inspired by the characters in dystopian fiction?" The interview responses from this section were used to provide justification for the quantitative findings. Finally, the third section inquired about the participants' characteristics, behaviors, and experiences that shaped who they are. Questions included, "Would you say that fear wakes you up or shuts you down?" and "Over the course of your life, do you feel as if you've become more defiant or more subservient towards society?" The third section of the interviews was a key component in discovering the participants' insurgent tendencies. It allowed for a comparison between certain Generation Z personalities and the personalities found in popular dystopian works with which they grew up with.

Qualitative Participants

Participants were thirteen undergraduates from the University of Georgia. The demographics for the interviews include seven females and six males, and seven White students, three African American students, and three Asian American students.

Direct Content Analysis

A direct content analysis was done on three popular dystopian novels of the past ten years: *Divergent* by Veronica Roth, *Catching Fire* by Suzanne Collins and *The Maze Runner* by James Dashner. In addition to being top selling and grossing, the novels were used because they have film adaptations, which increased the probability of their exposure to Generation Z young adults. The novels were also chosen because the individuality of the protagonists, grouped with their unique dystopian environments and the rise of revolution, made for three distinct heroic archetypes. This was essential for the creation of an activist typology.

The content of the novels was coded chapter by chapter to find key elements of themes or traits relating to heroism and activism. Three different variations of activism emerged from each of the three protagonists in the novels. Using those three variations, a comparison was made between the typology of activists that exist in the novels and the typology of activists that exist within Generation Z young adults, or, rather, those who were interviewed. To simplify, first, the traits, experiences, and mannerisms of each of the main protagonists in the novels were identified and listed. Then, the traits, experiences, and mannerisms of each of the interviewed participants were identified and listed. If the traits of an interviewed participant matched with the traits of a main protagonist, the interviewed participant was grouped with the fictional character and their activist type. This typology comparison was then used to examine Generation Z's activist tendencies in relation to the activist tendencies found in the characters in dystopian fiction. It also allowed and gave grounds for the third research question to be investigated in more depth. However, those theoretical comparisons will be further explored in a later part of this paper.

RESULTS

The means and standard deviations were calculated for each of the four subscales of the *Aspects of Literary Response: A New Questionnaire*.

Quantitative Results

The descriptive statistics for the Insight subscale showed in **Figure 1** that the possible range was between 21 and 70. The mean score was 49.78, and the standard deviation was 11.78, suggesting that the responses were somewhere between "Moderately True" and "Very True." This meant that the students affirmed the belief that literature may guide the recognition of previously unrecognized qualities. Based on the similar responses from the interviews, this data was confirmed. The majority of the interviewed participants admitted to seeing the relationship between reality and fiction.

The descriptive statistics for the Empathy subscale showed in **Figure 1** that the possible range was between 7 and 35. The mean score was 18.66, and the standard deviation was 7.45, suggesting that the mean average for the Empathy subscale, which explained the reader's ability to identify to the characters, ranged between "Slightly True" and "Moderately True."

The descriptive statistics for the Imagery Vividness subscale showed in **Figure 1** that the possible range was between 11 and 45. The mean score was 29.51, and the standard deviation was 9.35. Thus, the mean average for the Imagery Vividness section, which expressed the extent to which the literary world became vividly present to the reader, ranged between "Moderately True" and "Very True."

Finally, the descriptive statistics for the Rejection of Literary Values subscale showed in **Figure 1** that the possible range was between 9 and 39. The mean score was 15.00, and the standard deviation was 5.09. The mean average for Rejection of Literary

Values ranged between “Not True At All” and “Slightly True.” This meant that the majority of participants did not reject literature or its values/purpose.

		Statistics			
		Rejection	Imagery Vividness	Empathy	Insight
N	Valid	100	100	100	100
	Missing	0	0	0	0
Mean		15.0000	29.5100	18.6600	49.7800
Median		14.0000	31.0000	17.0000	50.0000
Mode		14.00	32.00 ^a	15.00	52.00 ^a
Std. Deviation		5.09307	9.35981	7.45115	11.78390
Variance		25.939	87.606	55.520	138.860
Range		30.00	34.00	28.00	49.00
Minimum		9.00	11.00	7.00	21.00
Maximum		39.00	45.00	35.00	70.00
Sum		1500.00	2951.00	1866.00	4978.00
Percentiles	25	11.0000	22.2500	13.0000	42.0000
	50	14.0000	31.0000	17.0000	50.0000
	75	17.0000	36.0000	24.7500	57.7500

a. Multiple modes exist. The smallest value is shown

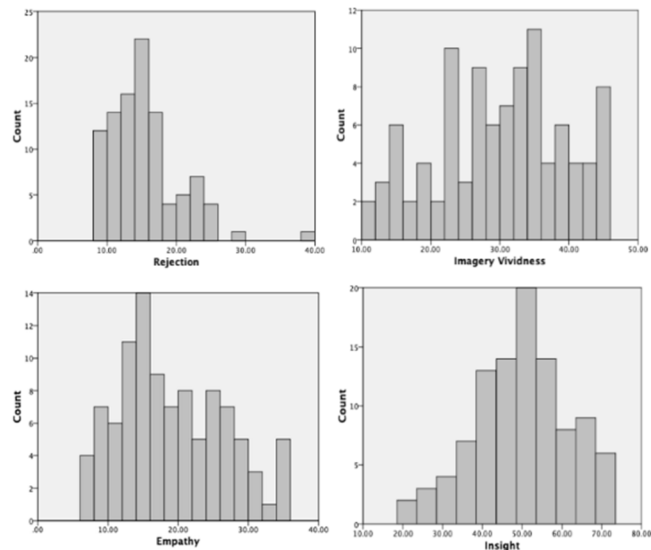


Figure 1. Reported outcomes of the subscales Insight, Empathy, Imagery Vividness, and Rejection of Literary Values with their mean averages.

A Pearson Product-Moment correlation analysis was conducted to investigate if there was any relationship between Insight and Empathy. Guided by Hopkins and Weisberg's research on cognitive development and literary influence, the suggestion was made that there was a relationship with a reader's ability to be enlightened by and/or connect with the literature.⁵ As shown in **Figure 1**, the Pearson coefficient was .573 and statistically significant at the .01 alpha level. This suggests a moderately strong positive relationship between Insight and Empathy. The extent to which the participant discovered things that they did not previously know about themselves had a connection to the extent to which the participant identified with the character.

Similarly, a Pearson Product-Moment correlation analysis was run to investigate if there was any relationship between Imagery Vividness and Empathy. **Figure 2** shows the Pearson coefficient to be .639 and statistically significant at the .01 alpha level. The data suggests that the extent to which the literary world became vividly present to the participant had a positive correlation to the extent to which the participant identified with the character.

		Correlations	
		Insight	Empathy
Insight	Pearson Correlation	1	.573**
	Sig. (2-tailed)		.000
	N	100	100
Empathy	Pearson Correlation	.573**	1
	Sig. (2-tailed)	.000	
	N	100	100

** . Correlation is significant at the 0.01 level (2-tailed).

		Correlations	
		Empathy	Imagery Vividness
Empathy	Pearson Correlation	1	.639**
	Sig. (2-tailed)		.000
	N	100	100
Imagery Vividness	Pearson Correlation	.639**	1
	Sig. (2-tailed)	.000	
	N	100	100

** . Correlation is significant at the 0.01 level (2-tailed).

Figure 2. Reported outcomes of the Pearson Product-Moment correlation analysis for the subscales Insight and Empathy and Imagery Vividness and Empathy.

The Pearson Product-Moment correlation found weak correlations as well. In regards to the relationship between Rejection of Literary Values and Imagery Vividness, **Figure 3** shows the Pearson coefficient to be .056 and without significance. This implied that the participants who rejected the values found in literature were not shown to have feelings of enlightenment or a sensory connection towards the text. Similarly, the Pearson Product-Moment correlation ran to also expose the relationship between Rejection of Literary Values and Insight had a coefficient of -.142 as displayed in **Figure 3**. This showed that the participants who rejected the values found in literature did not necessarily have a connection to the characters found in fiction.

Correlations				Correlations			
		Imagery Vividness	Rejection			Rejection	Insight
Imagery Vividness	Pearson Correlation	1	.056	Rejection	Pearson Correlation	1	-.142
	Sig. (2-tailed)		.579		Sig. (2-tailed)		.158
	N	100	100		N	100	100
Rejection	Pearson Correlation	.056	1	Insight	Pearson Correlation	-.142	1
	Sig. (2-tailed)	.579			Sig. (2-tailed)	.158	
	N	100	100		N	100	100

Figure 3. Reported outcomes of the Pearson Product-Moment correlation analysis for the subscales Rejection of Literary Values and Imagery Vividness and Rejection of Literary Values and Insight.

In summary, the quantitative data provided a small view of Generation Z's literary values. Through the descriptives shown in the Insight, Empathy, and Imagery Vividness sections, it is inferred that they have a "Moderately True" opinion towards literature's influential and relational capabilities. Additionally, they do not generally reject literature's purpose or values, and the interview responses provided a more in-depth explanation as to why, with dystopian fiction in mind specifically. Eleven out of the thirteen participants agreed that dystopian literature had significance. Yet the question of personal significance was not specifically answered. It was not until a direct content analysis on three popular dystopian novels was conducted that the qualitative data began to suggest an influential relationship. It also began to suggest a typology.

Qualitative Results: An Activist Typology

Three categories of Generation Z activism arose from this study, all of which can be compared to the typology of activism seen in recent dystopian literature and literary media. The three categories are as follows: the Selfless Activist, the Reluctant Activist, and the Instinctual Activist. Thirteen Generation Z students were interviewed, yet only twelve fit into a category that provided an explanation for their drive and provocations. The outlier participant of this study will be further investigated and explained in the discussion section of this paper.

The Selfless Activist

Three out of the thirteen participants of this study were categorized as selfless activists. Lawrence A. Lengbeyer describes selflessness as "setting aside concerns for oneself, out of concern for others."⁹ It is the notion that encompasses self-sacrifice, whether mentally or physically, so that another can be comforted or provided for. Selflessness is more or less an aspect of heroism, as it consists of bravery that results from a lack of self-regard. As opposed to the champion-like doctrine that typically defines heroism, the characteristics of the selfless activist refer to mediocre acts like standing up for someone in a distressed situation or speaking for a cause that does not personally affect them.¹⁰ It gives a platform for the "realistic hero," or someone who does the right thing without possessing outstanding traits and abilities.¹⁰ Instead, they use personal and moral values, but they allow their selflessness to dictate their actions.

Selflessness is also an emphasized theme in Veronica Roth's *Divergent* series. Throughout the first novel, the protagonist Tris Prior undergoes a variety of tests to analyze her mental and physical abilities. In all of the cognitive and physical assessments, Tris acts without regard for herself, essentially ignoring her fear to protect and/or save another.¹¹ The author means to imply that selflessness often leads to bravery, or that the two terms are synonymous.¹¹

The first novel was published in 2011 and the first film was produced in 2014. Nine out of the thirteen participants said they have seen the film and/or read the novel in their youth, including the three mentioned below. Overall, the participants' perceptions of the novel ranged from indifference to reflective. For example, Participant 4 gave her thoughts on what she gathered from the story:

I felt the idea of Factions was pretty reflective of our society ... I feel like how they divided themselves so strictly and kept themselves in boxes and the fight for leadership was kind of reflective of what I've seen today ... What's legal may not always be right; the laws may not be fair.

When asked about their reason for being involved in protesting or advocating against certain social-political issues, the three participants talked about various causes. Participant 2 spoke on her White Privilege and how it should be handled. Aware of the social advantages, she explained that those with privilege need to use their benefits to help others. They should not be selfish with their privilege and instead use their special right to help those without. Similar to Tris from *Divergent*, who sees "guilt as a tool,

rather than a weapon against the self,” Participant 2 recognized the use and value of guilt by association.^{11(p315)} For example, she states,

I know that I have privilege and that’s not something I can personally help, but it’s something I can use to other people’s advantage. I think it’s important that when we do have some sort of privilege to use it to help those who are disadvantaged by the same system that we’re advantaged by. And I think that being politically active is part of that and without people voting and [working] to change legislature and politicians then ... the privilege will never go away and people won’t have the same privilege that [I] do ... It’s my civic duty to help those whose voices aren’t heard as loud as mine.

The progressive attitudes of these Generation Z participants were reflected in their responses, as concepts like “White Privilege” and oppression became less of a quickly-avoided topic of discussion. Additionally, the participants openly acknowledged the necessity of change as a means to fix the indifference and mistakes of the generations that came before them. In Roth’s *Divergent*, an older character recollects, “Decades ago our ancestors ... divided into factions that sought to eradicate those qualities they believed responsible for the world’s disarray.”^{11(p42)} In Tris’s world, the older generation believed that conforming one’s personality was the best way to establish peace. However, Tris later discovers that her ancestors were wrong, and she intends to fix her outlook on separated personalities. She says, “I feel like someone breathed new air into my lungs. I am not Abnegation. I am not Dantless. I am Divergent.”^{11(p442)} Tris considers herself renewed and finds a sense of self-agency. She intends to go her own way instead of the way others have paved for her. This independence in thought, brought on by defiance, was also found within the participants of this study who were categorized as selfless activists:

I feel like it’s my responsibility to try and fight for better circumstances for other people. And I feel like, as a younger generation, we can’t just accept that the older generation knows everything, or that they know what’s best. We have to accept that they’ve made mistakes, and we might have to fix them. (Participant 4)

The selflessness that exists in these participants is often a product of their relationships. As the world gradually grows more and more diverse, diverse relationships are likely to occur. Through those relationships, perceptions are broadened and viewpoints are subject to change. As one begins to sympathize with the difficult experiences of their peers, they may also begin to advocate for change along with them. Thus, the selfless activists are those who are willing to put themselves in uncomfortable positions for the betterment of others despite having no direct connection to the issue. An advocate for taking action, Roth’s protagonist comes to believe “in the ordinary acts of bravery; in the courage that drives one person to stand up for another.”^{11(p207)} Gathered from the data, it was found that the selfless activists of this study believed that same notion:

One of my teachers, she gave a quote the other day in class, or maybe a couple of weeks ago, where she said, “I think your generation ... is the first the first in a while to care about something other than themselves.” And, I would agree with that, very strongly. I think we do care about stuff other than ourselves ... For example, if I had a teammate who felt called to kneel for the national anthem at a game and was under some type of ridicule, I would stand by them, especially if I thought their cause was noble. (Participant 13)

These participants are examples of the “realistic heroes” that exist in the world. They are the citizens who make a difference every day by standing up for their peers and openly speaking out against the things they disagree with. As Lawrence A. Lengbeyer explains, their selflessness is “uncalculated,” and they behave courageously by shutting “out of one’s cognizing certain considerations that, if taken fully seriously, would arouse fear.”⁹ Having grown up in a world of terrorism, war, and political unrest, Generation Z is familiar with fear. However, they are also familiar with change—for themselves and others.¹

The Reluctant Activist

Five out of the thirteen participants of this study were categorized as reluctant activists. The reluctant activist describes the participants who were first hesitant towards advocating for social-political issues. Later, their feelings of indifference or disinclination were overshadowed by their insurgency and realization for the need for social change. Typically, when a person’s attitude changes from that of indifference to that of concern, it is due to a shift in “self” or an overwhelming surge of emotion that then drives their activism.¹² As opposed to the confident, “born” activist, the reluctant activist is circumstantial or accidental. A series of unfortunate events would have drawn them to different perspectives, specifically those events that were deemed climactic for the individual. In a study done by Tracey Ollis entitled *The Accidental Activist: Learning, Embodiment and Action*, it was found that emotions like anger and frustration account for the change in purpose and desire for better circumstances, as did collectivity.¹² Essentially, the activists become subject to a “turning point” and then are drawn to revolutionary movements.

This type of activism is also found in Suzanne Collins' *The Hunger Games* trilogy, specifically in the second novel of the trilogy, *Catching Fire*. As opposed to the first novel, the prospect of revolution and societal disobedience becomes more prevalent in *Catching Fire*. While in *The Hunger Games* the rebellious attitude and insurgence of the protagonist Katniss Everdeen is shown, it is not until the second novel when the rebellious attitude and insurgence of the dystopian society is shown. *Catching Fire* was analyzed instead of *The Hunger Games* because the novel demonstrates the formation and progression of an uprising, and a young activist's response to a quickly changing world.¹³ In a way, the demonstrations, protests, and revolts that transpire in *Catching Fire* mirror the demonstrations, protests, and revolts that have transpired in the real world. Thus, the Katniss that exists in *Catching Fire* can then be compared the Generation Z young adults that exist in the real world.

A common theme found in the novel is the struggle between reluctance and willingness. Due to having already been awakened to the true horror of the world, Katniss longs for normalcy and wants no part in a rebellion.¹³ However, the possibility of a social change being brought forth secretly excites her. In the beginning of the novel, during a conversation with the antagonist President Snow, Katniss is "chilled and somewhat elated by the possibility" of an uprising.^{13(p21)} Yet, Katniss is still hesitant to take action, and she does not speak out against injustices of her world, although she knows they exist. It is until she witnesses the government's unfair and brutal treatment of a friend when the heroine's rage gives her the strength to defy.¹³ Her friend's sufferings, as well as the sufferings of those around her, lead her to accept and take part in the rebellion.¹³ She decides to stop running away from the issues of the world and to confront them. For instance, Katniss says, "At some point you have to stop running and turn around and face whoever wants you dead."^{13(p118)} She then decidedly tells her friend Gale, "I'm not going anywhere. I'm going to stay right here and cause all kinds of trouble."^{13(p119)}

Catching Fire was published in 2009 and the first film was produced in 2013. Ten out of the thirteen participants said they have seen the film and/or read the novel in their youth, including the four mentioned below. Overall, the participants' perceptions of the novel can be summarized as inspirational. For example, Participant 6 gave her thoughts on how she felt about the story:

I think that I was able to place myself in [Katniss'] shoes, and I kind of saw why she acted the way she did ... [I'm] encouraged. Because I feel the same anger that [the characters] feel when things happen to them.

The case of the reluctant activist is complex, as it may take a while for the activist to reach that emotionally-charged transitioning point. Often, they start reluctant for fear of the movement. This fear is often due to belief in the government's indifference or intervention. Also, the fear may be due to the intimidating and aggressive mood which often characterizes some demonstrations. In Suzanne Collin's *Catching Fire*, Katniss describes a horrendous scene following a speech she gives to a crowd of people.¹³ While she does not intend to encourage an act of civil disobedience, her audience "presses the three middle fingers of their left hand against their lips and extend them" in a public salute to Katniss, the "girl who defied the Capitol."^{13(p61)} The gesture, though simple, fills Katniss with fear, because she knows she has "elicited something dangerous."^{13(p62)} In the scene, turmoil and violence breaks out between the people and the Capitol's police officers, and the idea of an uprising appears more detrimental to Katniss, who at first feels "acts of dissent" should be defused instead of supported.^{13(p62)} Participant 1 of this study, who later found his mode of activism to be more small-scale than large-scale, describes a similar event and feeling. While observing an anti-Trump protest on campus, the participant talks about an incident that made him reluctant to join the movement:

On the rally [when] one of the Trumps was coming in ... my friend and I were walking past that event and he was wondering whether or not he wanted to go protest. But as we walked by, the [protesters] that were there were being very hateful and spiteful and calling us out for not immediately jumping in ... So, that's why I wouldn't want to get involved with any of that. A good cause could turn into a reason to be rowdy.

Other participants spoke on their dissatisfaction of the world as a reason for their resistance. However, their dissatisfaction also served as a catalyst. The "turning point" for more than half of the participants was the 2016 election of Donald Trump as the forty-fifth president of the United States. Specifically, the reluctant activists of this study spoke about their immense frustration with the event, and how they began to rethink their personal values in the years that followed. Like Katniss Everdeen, the participants believed their leader unfit, for his speech also gave the impression that "his primary concern is the welfare of the citizens ... when nothing could be further from the truth."^{13(p22)} For example, Participant 5 spoke on how his political platform changed in response to the election:

I grew up in a conservative, Christian household ... [After the election] I began to look at how society and the government kind of meshed in a more broad view and kind of understand how history led to some of the problems that we currently face today, as well as how our society has molded. That's when I became a bit more libertarian.

While government changes served as one type of catalyst for most participants, society and interpersonal relationships are contributing factors as well. When the “I” mentality turns into a “We” mentality, it is often due to the influence of peers. The emotions or enlightening that comes about towards social-political issues typically follows a conversation and/or an interaction with others who have more insurgent personalities, thus the individual is forced to contemplate differing viewpoints.¹⁴ Collective action is an absorbing force, and it succeeds in channeling and redirecting emotions of frustration to collectivity.¹⁴ This kind of redirection is also seen in *Catching Fire*. Influenced by the insurgent actions of her friend Gale, Katniss says, “Gale is right. If people have the courage, [a rebellion] could be an opportunity. He’s also right that, since I have set it in motion, I could do so much.”^{13(p123)} While explaining his rise to activism, Participant 10 talked about how he was drawn to collective action by his environment and society:

Before the semester, I wouldn’t have considered myself politically involved, [then] I was influenced by society ... [Now], I’m a part of YDSA, Young Democratic Socialists of America. I’ve [even] been a part of a committee that designed a survey to hand out to the community members gauging their well-being in the community and what policies they think will increase their well-being.

The emotions and testimonies of reluctant activists can be described as progressive and transformational. Similar to the fictitious character Katniss Everdeen from *The Hunger Games* trilogy, their insurgency “caught fire,” having been triggered by an eye-opening event and/or the influence of peers. While their turn towards the movement was gradual, it was a defining moment for each, summarized as an “awakening” for many. However, this process is not abnormal for Generation Z, who’s once disengaged and “lackluster” view of the country is now beginning to turn into something different.¹

The Instinctual Activist

The final type of activist that arose from this study was the instinctual activist. Unlike the previous types, instinctual activists are not activists due to personal “turning points” or selflessness towards a minoritized group. Instead, their insurgency is a trait that they are “born” with. The term defines a person whose acts of defiance are instinctual and impulsive, usually without reason or contemplation, and a result of curiosity and oppression. Four out of the thirteen participants in this study were categorized as instinctual activists because they moved recklessly, yet passionately, driven by motivation and a desire for something more.

An instinctual activist is also a term one could use to describe Thomas, the protagonist in James Dashner’s *The Maze Runner*. A common theme found in the novel is the redefining of identity and the transformation of fear into something useful. From the first chapter and onward, Thomas’s impulsive insurgence is a hybrid combination of fear and curiosity. He refuses to accept the normalcy of the boxed, maze-like world and seeks liberation immediately. His actions are reckless, yet instinctive, as an urge to take some sort of action overwhelms him. For example, after becoming enlightened by the horrid state of the world, Thomas immediately seeks liberation and change. When he learns about the tasks of the “Runners,” or people who run through the maze each day to find a way out, Thomas instantly feels like the job was made for him.¹⁵ Despite others’ warnings that the task is dangerous and unattractive, Thomas is drawn to it. Even when he sees the danger of the job himself, he remains rooted in his conviction, for it “called to him as much as hunger and thirst.”^{15(p46)} In a sense, he is driven by a motivational desire for freedom.

The Maze Runner was published in 2009, and the first film was produced in 2014. More than half of all participants said they have seen the film and/or read the novel in their youth, including two out of the four mentioned below. Overall, the participants’ perceptions of the novel can be summarized as positive and reflective. For example, Participant 12 gave his thoughts on how he felt about the story:

Sometimes you don’t catch it at first. As you engage in the story, you start to understand *this* is what they really mean. In the first movie, your viewpoint on their lives was one way, and as the story kept building on, you started to change your viewpoint. I feel like that same thing would be true if you brought that out to the real world because most people are ignorant to the things they just haven’t experienced themselves ... I really love [*The*] *Maze Runner*. That’s my favorite.

The traits of an instinctual activist can often be spotted at a young age. As explained by Darwin, instinct and impulse usually encompass the young, and the justifications for their actions may not even be known.¹⁶ However, that instinctual spirit may not fade away with time. Since instinct is often correlated with curiosity, those who remain curious later in their lives may also remain instinctive.¹⁵ Curiosity allows one to question the unknown. It was found in this study that curiosity also allows one to question societal norms and customs that may be opposed to a common, social belief. In *The Maze Runner*, Thomas was labeled a “rule breaker” because his curiosity caused heroic, yet rebellious tendencies.^{15(p152)} Similar to the instinctive activists of this study, Thomas “felt his fear whisked away like a swarm of gnats caught in the wind, replaced by an intense curiosity” for his world and

its subjective formalities.^{15(p2)} For example, Participant 11 spoke of her opinion on the controversy surrounding the pledge of allegiance:

People are finally realizing that the pledge of allegiance is foolish. When I was younger, I was like, why do we have to stand up and do this? I don't actually care about anything they're saying. They don't care about me [as an African American]. I stopped standing after a while. I used to get in trouble for not standing, but eventually they stopped requiring me ... I've always been defiant. I can't even explain it. It's just my natural tendency.

An instinctual activist may impulsively or instinctively advocate for a social-political movement that pertains specifically to their experiences. For example, while participants of this study pronounced that they cared about a wide range of issues, they advocated more so for personal offenses, like race, gender, and/or LGBTQ issues. They were quick to comment on their personal state of affairs, and they explained how their environments shape their personality. For instance, Participant 3 stated her opinion:

I've always been very defiant towards men wanting me to be subservient to them. That relates as a whole to me being defiant towards society, because society is the patriarchy ... [My response] is always to keep fighting. I don't want to be subservient in any part of my life.

The theme of "fighting" was a constant theme among the instinctual activists of this study, whose instinct to take action was almost as defined as their determination to create change in the world. In an article written by Caroline Myss, she explains the hidden truths of heroism. In its modern form, a hero is a person "who knows how to count on his own inner resources, his own instincts, his own fight-or-flight mechanism."¹⁷ Because they rely on the voice in their head, they are survivors, for they know their truth and act accordingly by using their instincts.¹⁷ In Dashner's *The Maze Runner*, Thomas exemplifies this type of heroism on several occasions.¹⁵ In justification for his and his friends' heroic acts, Thomas says, "Most people would've given up by now. But I think we're different. We couldn't accept that a problem can't be solved ... And we've kept fighting no matter how hopeless it's gotten."^{15(p305)} Likewise, Participant 12 of this study connects his instinctive behavior to activist/heroic behavior in the following statement:

I guess I have a fight response more than a flight response. Even if you run away, nothing changes. Ignoring your problems doesn't make the problem go away ... [For example], when I see wrongs being done in my view or hear about them I try to actively make a change. Whether it's small or a big change ... I don't think as a whole that [my generation just has] no power. I feel like we do. It's kind of like we're all superheroes, but we just haven't realized it yet.

The testimonies and statements of the instinctual activists can be deemed just as insurgent as their personalities and experiences. Coincidentally, all participants who were categorized as this type were of a minority group, causing their statements to be more radical than the rest. However, similar to the character from James Dashner's *The Maze Runner*, they entered into this world feet first, running towards liberation in an otherwise confusing and disgruntled world. Unlike the former types, the instinctual activist has a more personal motive. They are neither fighting for the sake of others nor undergoing a transformative process. Instead, equipped with curiosity and motivational desire, they fight for themselves, as it is their instinct to do so.

DISCUSSION

As there has been relatively little research pertaining to dystopian literature's relationship with activism, specifically in regards to Generation Z, this research aimed to shed light on the genre's implications. In accordance with the rise of dystopian literature within the last decade, the results of this study were to be expected. Since the hundred Generation Z participants who took the questionnaire were of those who grew up during a time of terrorism, financial crisis, media manipulation, and social-political unrest,^{1,7} the findings from this study suggest an interesting relationship between Generation Z members and the dystopian worlds and characters.

The interviewed participants described dystopian literature/literary media as synonymous to the real world. In regards to the Insight and Imagery Vividness subscales, the realization that the dark themes addressed in the fiction paralleled those in the real world dawned on the participants the more they engaged with the storyline. They began to visualize the world and its symbolism and described it as "too real" or a "reflection" of society. Participants also noted how the exposure to dystopian narratives lead to the discovery of previously unrecognized qualities within themselves. For example, Participant 12 expressed how Thomas in *The Maze Runner* made him realize that if he had a stance for something, then he had the ability to overcome any adversity that came in his way. Similarly, Participant 4 expressed how Katniss in *The Hunger Games* trilogy inspired her to speak out even if her life or

reputation were in danger. Thus, reflecting on the life of a fictional character allowed participants to reflect on their own lives. They recognized that insurgency was not a fictional trait but a real one.

The Empathy subscale was also justified by the responses found in the interviews. Participants were not only discovering qualities about themselves and their world that were previously unknown. They were also discovering a new identity, guided by the identities of the fictional characters in dystopian literature/literary media. The research by Justin Scholes and Jon Ostenson (2013) found that dystopian fiction's appeal to Generation Z young adults was due to the content's familiarity.¹⁸ The fictional characters who lived in a type of backward situation, environment, and/or world mirrored that of young Generation Z members, who were also in the stages of questioning both themselves and "the underlying values of a flawed society."¹⁸ Therefore, Generation Z members did not only understand the characters, but they identified with them. For example, Participant 6 addressed her relation to Katniss in *The Hunger Games* trilogy in that she also felt the same anger and frustration towards the government as the character did towards hers. Another participant, Participant 11, also expressed her relation to and admiration for Katniss, since the character rose from poverty to wealth in an inspirational way.

The Rejection of Literary values subscale scored a low mean average, with the responses ranging between "Not True At All" and "Slightly True." According to the questionnaire participants, the themes and storylines that exist in literature and the literary media are not generally ignored or deemed irrelevant, and twelve out of the thirteen interviewed participants expressed their reasoning as to why. The majority of the participants felt that it was important to see the themes in dystopian literature since they are so reminiscent of society. Since dystopian literature typically portrays what the future could be, the participants perceived the fiction as a warning. For example, Participant 4 stated that dystopian literature showed her how to be mindful of corrupt leaders, as the wrong leader could make a decision that negatively affects the world. Yet, interviewed participants also noted how dystopian literature is important in that it both projects and eases their fears for the world. Although the world may become dark and oppressive, it can also always be saved and revolutionized.

A Typology Theory

The findings from this study also suggest that the activism of Generation Z young adults relates to the activism found in the dystopian literature/literary media with which they grew up with. The theme of revolution is prevalent and popular in several YA dystopian novels and films, including in *Divergent*, *Catching Fire*, and *The Maze Runner*. Despite the harsh, oppressive world, the protagonist almost always prevails against the corrupt government, society, science, or technology that once abused them. Based on the direct content analysis of this study, the dystopian literature's protagonists' activism or insurgency was generally described as one of three: selfless, reluctant, or instinctive. Coincidentally, the activism of some Generation Z members may be described as the same. The interviews from the qualitative portion allowed a relationship to exist between the typology of activism found in dystopian literature/literary media and the typology of activism found in Generation Z young adults.

An outlier from the qualitative portion of this study allowed for another relationship to exist. While the majority of the thirteen interviewed participants of this study were able to be categorized in their activist type, one participant was not categorized at all. The specific participant, Participant 8, did not exhibit activist traits and did not hold any specific progressive or activist beliefs. Also, Participant 8 showed neither interest in nor previous exposure to dystopian literature/literary media. Based on what was discovered in the quantitative portion of this study, in regards to the Rejection of Literary Values subscale from the questionnaire, those who rejected literature or its importance were also shown to reject a connection with the characters or literary world. Therefore, since a connection did not exist between Participant 8 and dystopian literature/literary media, Participant 8 lacked the ability to become inspired by revolutionary-themed fiction. While it cannot be confirmed that Participant 8's lack of social-political justice was due to a disinterest in literature, a pattern was formed from the data. Whereas the interviewed participants who expressed an appreciation for dystopian literature/literary media represented an activist type, the interviewed participant who did not express an appreciation for dystopian literature/literary media did not contain an activist type.

CONCLUSIONS

Although this research puts forth an idea that ought not to be ignored, it is not definite that this research proves that Generation Z's activism has an influential relationship with dystopian literature and/or literary media. However, there is a kind of relationship that does exist, if not influential. The qualitative data attested that the activist traits of Generation Z young adults are similar to the activist traits of literary characters in the dystopian fictions *Divergent*, *Catching Fire*, and *The Maze Runner*. Also, the qualitative data, paired with the quantitative data, attested that dystopian literature is not a neglected source of inspiration. While there are many factors that could lead a Generation Z member to activism, like cultural experiences and personal phenomenon, this study means to insert the possibility of another factor. Still, there are several important limitations of this study to consider before such an insertion receives validation.

Limitations and Implications

One limitation of this study was the lack of literature around speculative fiction's influence on activism. While it is known that personal experiences can impact one's level of insurgence towards social-political issues, it is less known that literature/literary media can have an impact as well, especially in regards to a generation of people. Also, it is not likely that literature alone, regardless of the genre, can influence one's activism. It is more likely that one's background, culture, and environment would trigger or be the reasoning behind activism. For example, the minorities in this study were more likely to be influenced by their marginalization rather than their interest in insurgent-type literature. The racism, discrimination, and oppression experienced by those participants, as well as towards their people, would have motivated them to adopt an activist or insurgent behavior. Similarly, the revolutionary culture of the world, characterized by demonstrations, marches, and speeches made in advocating for civil rights, would have also influenced those participants to become activists, as they would have seen how change should be advocated.² Frankly, the revolutionary culture of the world would have also influenced non-marginalized groups of people, including some of the participants in this study, who attributed their insurgence to their desire to see change occur for the betterment of their oppressed peers. In this regard, speculative fiction or dystopian fiction may neither be one of the reasons nor the sole reason for influencing one's activism.

In addition, another limitation of this study is its dependence on self-reported data. The qualitative portion of this study leaned heavily on the responses from the interviewed participants, who may or may not have truthfully reported their activist experiences and literary beliefs. Similarly, the majority of the interviewed participants openly identified as either Democrat, Liberal, and/or Libertarian. Due to their political identities, it is likely that the participants self-selected themselves to be interviewed because they distinguished themselves as an activist. It is also likely that the participants agreed to take part in the study because they had an interest and/or were a fan of the dystopian genre in literature and literary media. Thus, their willingness to participate in the qualitative portion of this study may have contributed to their anti-conservative, progressive mindsets as well as their inclination towards speculative fiction. The limitation creates a bias in data collection from the interviews and doubts any claims of causality between reading about dystopian fiction and activism since not all Generation Z young adults consider themselves activists and/or attracted to the literary genre.

In regards to the quantitative portion of this study, the limitations existed in the sample size and questionnaire content. The sample size for this study was only a hundred participants. A larger sample size would give more accurate mean values, identify more outliers, and overall contribute to a higher confidence level. A larger sample size would also give a more general synopsis of Generation Z's literary perceptions as a whole generation, including those not categorized as young adults. It is an important limitation to note that the quantitative data, as well as the qualitative data, in this study was collected by young adults and college students, which is a mere subset of the Generation Z population. This subset of Generation Z does not in any way reflect Generation Z as a whole or the interests/activist traits of Generation Z as a whole.

The questionnaire content of this study served as a limitation in that it focused on Generation Z's perceptions of literature instead of primarily the dystopian genre. Additionally, the content could have done more to comply with the qualitative portion of this study, which inquired more so on the activism of Generation Z members. Perhaps, as a suggestion for future research, the survey instrument could include questions about Generation Z's dystopian literary perceptions in relation to their activist characteristics. Then, the research would be given a more direct answer in regards to literature's influential relationship with activism. The typology of activists found in the qualitative portion of this study would also benefit from that addition, as the questions relating to their activist characteristics would confirm the extent to which it was selfless, reluctant, or instinctive.

Final Notes

The main conclusion to be drawn from this research is dystopian literature's potential significance in encouraging/shaping the progressive mindsets of its audience. It is certain that Generation Z young adults and children have been surrounded by fictional dystopian storylines in literature and in the literary media for the past decade. Along with the contemporary favorites, classical narratives like *The Giver* and *The Handmaid's Tale* have been revived and publicized for their viewing, and the surge of others has not gone unnoticed. Although the dark themes tell stories of oppression, they also tell stories of revolution. They define resilience as the idea that change can be facilitated by anyone as long as there is a passion to drive it.

Consequently, instead of dismissing dystopian literature as an attractive entertainment medium, the public should consider the fiction as an inspirational tool. For whatever enlightens and influences Generation Z also enlightens and influences the future. Therefore, if dystopian novels full of revolutionary promise and change have truly captivated Generation Z young adults, then the world is already in better hands than imagined.

ACKNOWLEDGEMENTS

This project was supported by the TRIO McNair Scholars Program at the University of Georgia. The author thanks Dr. Elizabeth Kraft, Dr. Chris Pisarik, Dr. Mylene Culbreath, and all the participants of this study.

REFERENCES

1. Dorsey, J. (2016). *iGen's Political & Civic Outlook*. Retrieved from <http://genhq.com/wp-content/uploads/2016/02/iGen-Gen-Z-Political-Civic-Outlook-Research-White-Paper-c-2016-The-Center-for-Generational-Kinetics.pdf>
2. Pimentel, J. (2018, June 1). 15 young activists who are changing the world. *Complex*. Retrieved from <https://www.complex.com/life/young-activists-who-are-changing-the-world/jack-andraka>
3. Elliott, T., and Earl, J. (2018). Organizing the next generation: Youth engagement with activism inside and outside of organizations. *Social Media + Society*. <https://doi.org/10.1177/2056305117750722>
4. Lepore, J. (2017, May 29). A golden age for dystopian literature. *The New Yorker*. Retrieved from <https://www.newyorker.com/magazine/2017/06/05/a-golden-age-for-dystopian-fiction>
5. Hopkins, E. J., and Weisberg, D. S. (2017). The youngest readers' dilemma: A review of children's learning from fictional sources. *Developmental Review*, 43, 48–70. <https://doi.org/10.1016/j.dr.2016.11.001>
6. Nance, K. A. (2010). Reading human rights literature in undergraduate literature classes: Professorial desire, disciplinary culture, and the chances of cultivating compassion. *Journal of Human Rights*, 9(2), 161–174. <https://doi.org/10.1080/14754831003761670>
7. Ames, M. A. (2013). Engaging "apolitical" adolescents: Analyzing the popularity and educational potential of dystopian literature post-9/11. *Faculty Research & Creative Activity*, 11. http://thekeep.eiu.edu/eng_fac/11
8. Miall, D. S., and Kuiken, D. (1995). Aspects of literary response: a new questionnaire. *National Council of Teachers of English*, 29. Retrieved from <https://www.jstor.org/stable/40171422>
9. Lengbeyer, L. A. (2005). Selflessness & cognition. *Ethical Theory and Moral Practice*, 8(4), 411–435. <https://doi.org/10.1007/s10677-005-8838-0>
10. Keczer, Z., File, B., Orosz, G., & Zimbardo, P. G. (2016). Social representations of hero and everyday hero: a network study from representative samples. *PLoS ONE*, 11(8), 1–17. <https://doi.org/10.1371/journal.pone.0159354>
11. Roth, V. (2011). *Divergent*. Katherine Tegen Books.
12. Ollis, T. (2008). The accidental activist: Learning, embodiment and action. *Australian Journal of Adult Learning*, 48(2), 316–335.
13. Collins, S. (2009). *Catching Fire*. Scholastic Press.
14. Eyerman, R. (2005). How social movements move, in *Emotions and Social Movements* (Flam, H., and King, D., Eds.) 1st ed., 41–56, Routledge.
15. Dashner, J. (2008). *The Maze Runner*. Delacorte Press.
16. Burghardt, G. M., and Jr, R. W. B. (2018). Wallace Craig's appetites and aversions as constituents of instincts: A centennial appreciation. *Journal of Comparative Psychology*, 132(4), 361–372. <http://dx.doi.org/10.1037/com0000155>
17. Myss, C. (2014). What is a hero? *KidSpirit Magazine*, 80–82.
18. Scholes, J., and Ostenson, J. (2013). Understanding the appeal of dystopian young adult fiction. *The ALAN Review*, 40(2). <https://doi.org/10.21061/alan.v40i2.a.2>

ABOUT STUDENT AUTHOR

Aysha Jerald is currently an undergraduate junior pursuing an A.B. in English, a B.S.Ed. in English Education, and a certificate in New Media Studies at the University of Georgia. She plans on attending graduate school to obtain a Ph.D. in Comparative Literature, and she intends to pursue a research career in digital media and literary influence.

PRESS SUMMARY

Although primarily an entertainment media, dystopian literature and film have drawn the attention of many readers and film-goers, especially those within Generation Z. In fact, recent studies have shown that Generation Z is particularly interested in the dark, fictitious genre, as the themes that exist in the literary world also exist in the real world. This research studies that connection and its implications, specifically in regards to Generation Z's activism and activist characteristics. It is posited, from this study, that the insurgent and revolutionary traits of the youngest generation could be linked to dystopian literary impact and exposure. Since radical and progressive book series and movies like *Divergent*, *The Hunger Games*, *The Maze Runner*, and others have dominated the literary scene for the past ten years, the findings of this mixed method research study suggest that the fiction did more than entertain. It suggests that the fiction inspired and enlightened the reformist young adults of the present day.

