



STANFORD
Undergraduate Research Journal

VOLUME 15 | SPRING 2016
Stanford, CA 94305

STAFF

The mission of the Stanford Undergraduate Research Journal is to encourage, recognize, and reward intellectual activity beyond the classroom, while providing a forum for the exchange of research and ideas.

EXECUTIVE

Editors-in-Chief Allison Dods
Emily Alsentzer

Financial Officer Taide Ding

Outreach Officer Michelle Lu

PRODUCTION

Lead Lucy Li

Scott Meyer
Jennifer Parker

LYNX

Managing Editors Kate Nelson
Hannah Shilling

Writers and Editors Alan Aw
Aron Tesfai
Hanna Payne
Kyle Lin
Ankit Baghel
Amalia Saladrigas

EDITORIAL

Natural Sciences

Section Editors Aanchal Johri
Emily Pang

Associate Editors Jack Andraka
Annabel Chen
Melodyanne Cheng
Genevieve Selden

Humanities

Section Editor Shu Chen Ong

Associate Editors Callie Hoon
Cory Bhowmik
Dana Huh
Stanley Xie
Jaydeep Singh

Social Sciences

Section Editors Taide Ding
Laura Zhang

Associate Editors Lorel Sim
Julian Watrous
Ramya Balasingam
Jessica Hui
Reesab Pathak

Engineering

Section Editor Vishnu Shankar

Associate Editors Albert Chu
Nitya Mani
Shivani Baisiwala
Esha Maiti

Visit surj.stanford.edu for...

- Complete journal archive
- Information for authors
- Interest newsletter
- Articles from *The Lynx* featuring the humans behind research
- Opportunities to join SURJ

LETTER FROM THE EDITORS

Dear Reader,

We are excited to present the fifteenth volume of the Stanford Undergraduate Research Journal. This edition of SURJ, like its predecessors, reflects our entire staff's commitment to identifying and publishing exceptional undergraduate research from Stanford and other universities around the world. The sixteen articles published in this volume, chosen from a pool of well over 100 submissions, come from a variety of academic disciplines. We invite you to examine the role of mangrove forests in carbon sequestration with Daryll Carlson ('17), investigate the mechanics of lower limb prostheses with Maurice Chiang ('19), explore the role of imperfect choreography in Jacques Demy's 'The Young Girls of Rochefort' with Carlos Valladares ('18), and delve into the perception of acupuncture in Beijing with Sharon Wulfovich ('17).

Our organization has endeavored since 2001 to support undergraduate researchers and foster interest in research throughout the broader undergraduate community at Stanford. In keeping with these goals, this year saw the expansion of SURJ's companion research magazine, LYNX, which highlights the research at Stanford and the experiences of the people behind it. LYNX now has a full staff of its own and publishes articles online biweekly. This year SURJ has also incorporated faculty readers into our review process, in accordance with our mission to publish undergraduate research of the highest quality. Though our editing process continues to be directed and primarily run by undergraduates, we look forward to continuing to build our relationship with Stanford faculty and with other research-oriented organizations on campus.

We could not be more grateful to our dedicated staff for all of their contributions. In particular, we would like to recognize our 2015-2016 officers: Engineering section editor Vishnu Shankar ('18); Social Sciences section editors Laura Zhang ('18) and Taide Ding ('17); Humanities section editor Shu Chen Ong ('17); Natural Sciences section editors Emily Pang ('18) and Aanchal Johri ('18); Production lead Lucy Li ('18); and Outreach lead Michelle Lu ('19). Their leadership, and the enthusiasm, patience, and hard work of the entire staff, were crucial to the production of this journal.

This journal would not be possible without the additional support of the offices of Undergraduate Advising and Research, the Vice Provost for Undergraduate Education, and Student Activities and Leadership. We would also like to thank Dr. Shay Brawn of the Program in Writing and Rhetoric for her mentorship during the editing process, and Vivienne Fong, our Faculty Advisor, for her advice and direction as we continue to expand and improve the journal. Finally, we are thankful to all of the writers who submitted manuscripts for consideration in this year's volume of SURJ, and to you, our readers, for demonstrating your support of undergraduate research.

Please join us in congratulating the authors published in this edition of SURJ. It is our hope that you are inspired, impressed, and encouraged by their contributions to research—we certainly are.

Sincerely,

Allison Dods ('16) and **Emily Alsentzer** ('16)
Editors-in-Chief

TABLE OF CONTENTS

SOCIAL SCIENCE

8

Acupuncture in the Treatment of
Frozen Shoulder, Tennis Elbow, and
Carpal Tunnel Syndrome

Sharon Wulfovich and Sakti Srivastava

16

Disentangling the Trilemma:
Factors that Shape Policy Choice

Lauren Iannolo

22

Location Models in the Art Market:
Artists as Firms Visualized on a
Product Space

Ricardo R. Rodríguez-Padilla

29

New Media Technology and Social
Movement Framing in the WTO
Protest: A Case Study in Peaceful
Majoritarian Possibilities for 21st
Century Activism

Michael Lachanski

35

On an Empirical Analysis of the
Kidney Allocation System

Khizar Qureshi and Robert Langer

ENGINEERING

44

State Transition Selection from
Absolute Limb Orientation in a
Powered Transfemoral Prosthesis

Maurice Chang

52

Interstitial Metal Hydride Alloys:
A Systematic Literature Review of
Contemporary Hydrogen Storage
Applications

Emily E. Petersen

60

A Sociological Perspective on Joyce's Women in The Portrait of the Artist as a Young Man and the Irish Parallel
 Mariam Sulakian

66

"To Confront a Portrait:" Photography, Parts, and the Whole in the Poetry of Walt Whitman
 Alex Torres

71

New Atheism: Constructing a New Nonbelieving Identity
 Canton Winer

79

Dance and the Postmodern Sublime in Jacques Demy's *The Young Girls of Rochefort* (1967)
 Carlos Valladares

86

Difficult Masses: Femininity and the Female Body in Germaine de Staël's *Corinne, Or Italy (Corinne)* and Mary Wollstonecraft's *The Wrongs of Woman, or Maria (Maria)*
 Emma Blauciak

NATURAL SCIENCE

94

Are E-Cigarettes a Safe Alternative to Smoking?
 Hayley Tartell

100

Carbon Sequestration and Microbial Communities in Mangrove Ecosystems
 Daryll Carlson

105

Expanding the Chromatic Range of Galaxies for Weak Gravitational Lensing Simulations
 Bradley Emi, Jason Rhodes, Plazas Malagón, and Andrés Alejandro

114

The Combined Effects of Ocean Acidification and Sea Surface Temperature Rise on Larval Sea Hare (*Aplysia dactylomela*) Development
 Anja Brandon

120

Vigilance in hand-reared agile wallabies (*Macropus agilis*) after release
 Amalia Saladrigas

FROM PUPPIES TO PENDULUMS

MARSHALL KUYPERS & UNDERSTANDING COMPLEX SYSTEMS

by Alan Aw

When a group of thirsty puppies competes for a bowl of milk, each tries to gain the biggest share. Each puppy's objective is simple and clear. So how is it possible that they end up huddling close to each other and moving in a pinwheel?

The study of how simple rules can lead to complex emergent phenomena—such as thirsty puppies creating pinwheels—fascinates Marshall Kuypers, Co-President of the Stanford Complexity Group (SCG). “Surprisingly many phenomena, ranging from ant colony behavior to price bubbles in economics, can be explained from this viewpoint,” Marshall says.

This field of study, also known as complex systems, has a rich history spanning multiple disciplines. The Nobel Laureate and philosopher Friedrich Hayek dedicated much of his work in the 1950s-70s to the study of complex phenomena in psychology and biology. In the 1940s mathematicians and information scientists such as Norbert Wiener and John von Neumann observed complex systems at play in cybernetics research.

“A very nice example involving agent-based models is the work of Thomas Schelling, who won the Nobel Prize in Economics in 2005,” explained Marshall. While investigating racial segregation in populations, Schelling found that the

geographical segregation came about from each individual expressing just a very slight preference for living near people of the same ethnicity.

Despite its ubiquity, it was only in recent decades that complex systems emerged as an independent field, with the establishment of institutions such as the Santa Fe Institute and the New England Complex Systems Institute as well as journals like *Complex Systems*.

“When I applied for a PhD program in the US, universities told me that if you want a PhD in complex systems you can start by applying to an economics or biology graduate program. There was no stand-alone program for such an interdisciplinary field,” says Marshall.

“It's weird.”

According to Marshall, the “weird,” interdisciplinary nature of complex systems is one key reason why it has not reached out to as many intellectual communities as it should. More needs to be done to spread the ideas of complex systems. “This is one of our goals at the Stanford Complexity Group.”

The SCG, founded in 2009, was originally created to serve a small body of faculty and graduate students interested in the field. Today, it is run by a team of PhD students whose diverse interests range from population biology to modern thought and literature. “All of us [in the SCG] are fascinated by how complex systems thinking relates to our intellectual or research pursuits,” Marshall says. Marshall himself has worked on agent-based models in economics, applying mathematical tools from network theory to investigate emergent properties in consumer-producer markets.

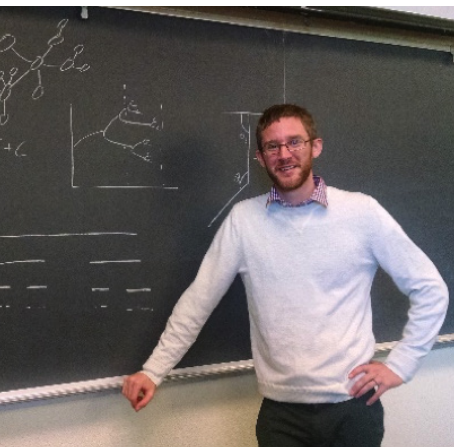
The SCG, Marshall elaborates, invites speakers from around the world to give talks on campus. It also offers an introductory seminar on complex systems every Winter Quarter, and makes three-minute video clips that communicate important ideas about complex systems in an accessible manner.

“The double pendulum is a standard example of a complex physical system. But if you Google it, you wouldn't be able to find a nice video explaining how it works,” explains Marshall. So the team at SCG decided to make one, with the help of a local artist. The video about the double pendulum can be found on the SCG website.

Apart from reaching out to the Stanford community and communicating complex systems to the public, the SCG has a larger vision of becoming the Bay Area hub for complex systems. “New Mexico has Santa Fe, the East Coast has the New England Complex Systems Institute. There's also one in the Midwest. Where are the West Coast's complex systems theorists?” quips Marshall.

For Marshall, complex systems is as under-noticed as it is emerging. Much is waiting to be discovered. Marshall, who also works at the Center for International Security and Cooperation (CISAC) and is pursuing a PhD with the Department of Management Science and Engineering (MS&E), hopes to find ways to marry his interest in complex systems with his current research on risk analysis.

“There seems to be little interaction between my current PhD work and my interest in complex systems that led me to pursue my PhD in the first place. But it's worth thinking about the potential connections that a relatively new field like complex systems could have.”



social science

Acupuncture in the Treatment of Frozen Shoulder, Tennis Elbow, and Carpal Tunnel Syndrome

A Literature Review, Survey, and Interview Based Study

Sharon Wulfovich and Sakti Srivastava MBBS, MS

Stanford University

Background: Frozen shoulder, tennis elbow and carpal tunnel syndrome are all prevalent conditions that lack a consensus over the most effective course of treatment.

Objectives: The aim of this study was to investigate the perception and effects of acupuncture on frozen shoulder, tennis elbow, and carpal tunnel syndrome in Beijing, China. It explored trends in age, gender, economic burdens and selection of western and/or Chinese medicine. This study provides an overview of current research on the effects of acupuncture on carpal tunnel syndrome, tennis elbow, and frozen shoulder and a brief overview on traditional Chinese medicine and the theory of acupuncture.

Methods: A total of 57 participants filled out questionnaires, and patient and doctor interviews were conducted in order to assess their views and experiences towards acupuncture and Chinese medicine.

Results: The survey and interview data indicated that participants and doctors view acupuncture to be effective and beneficial. In addition, after receiving treatment, the majority of participants indicated that their perception of acupuncture changed. The data indicates that the participants sixty years or older view Chinese medicine differently than the younger generation. Lastly, evidence of a health disparity and differing economic burdens between urban and rural citizens was found. Overall, most people would recommend and enjoyed receiving acupuncture.

Conclusion: Although further research is necessary, the survey data, interviews and current literature indicate that acupuncture can be used as a complementary treatment option for conditions including frozen shoulder, tennis elbow and carpal tunnel syndrome

KEYWORDS: Carpal Tunnel Syndrome, Frozen Shoulder, Tennis Elbow, Acupuncture, Conservative Treatment, Complementary Medicine, China

BACKGROUND INFORMATION ON ACUPUNCTURE

Traditional Chinese Medicine was established more than 3,000 years ago in China. Currently, it is used in most parts of the world and is considered a routine treatment in China, Korea, Japan, and Taiwan (Sierpina & Frenkel, 2005). Traditional Chinese Medicine treatment approaches include acupuncture, herbalism, moxibustion, massage, Tai Chi exercise, meditation and dietary changes (Sierpina & Frenkel, 2005).

Acupuncture is a way of preventing, diagnosing, and treating disease. It uses the insertion of needles at differing profundities and angles in various points throughout the body to produce an effect. While there is still a limited amount of research on acupuncture and its applications, theoretically acupuncture could cure or treat any illness that can be altered by a physiologic process. Acupuncture's long history in Asia and Western Europe has demonstrated its usefulness in improving human health (Armstrong, 1972). Its increasing acceptance and demand as

a treatment for chronic disorders and pain management has spurred an augmentation of scientific research on its validity (Langevin, Churchill, & Cipolla, 2001). Although there has been limited research on acupuncture and its mechanisms of action, its long utilization points to its potential usefulness.

UNDERLYING PHILOSOPHY OF ACUPUNCTURE

The principles of traditional Chinese medicine and the development of acupuncture are based on the oriental philosophy of Yin, Yang, and Qi (Lee, Lariccia, & Newberg, 2004). Yin and Yang refers to the theory that the universe is composed of five elements: wood, earth, fire, water, and metal. This oriental philosophy is based on the importance of maintaining a balance between the opposing Yin and Yang forces (Armstrong, 1972). All natural entities, including organs and senses, have been classified according to these forces. Therefore, traditional Chinese medicine's techniques aspire to find a balance between the Yin and Yang

forces in the body. Since disease, according to this philosophy, is caused by an unbalance between these opposing forces, by finding a balance, health can be achieved and disease can be treated or prevented (Lee et al., 2004). As a result of this belief, diseases, acupuncture points, and meridians are classified according to their characteristics with the Yin and Yang (Armstrong, 1972). Diseases can be explained and treated by determining which force is unbalanced. Similarly Qi, the energy of life, circulates through the meridians. Qi needs to be preserved and maintained in the right proportion for health to be maintained. The obstruction of the flow of Qi or an excess of Qi in certain regions of the body is another explanation for disease (Armstrong, 1972). These theories provide an explanation for disease as well as provide the principles on which acupuncture and other traditional Chinese medicine techniques were established.

HEALTHCARE IN CHINA

In 1949, the Chinese Communist Party rose to power in China. It created a healthcare system that was completely regulated by the government. But in 1978, China completely changed its healthcare system. Although the Chinese healthcare system has many issues, it is the only system in the world that incorporates both western medicine and traditional Chinese medicine in every treatment level. In China, traditional Chinese medicine (which includes acupuncture, massage, acupressure, moxibustion, herbal remedies etc.) is approximately 40% the healthcare delivered and is believed to be much higher in China's rural areas (Zheng & Hillier, 1995). This does not account for all of the self-prescribed Chinese medicine traditions that families commonly make for themselves (Hesketh & Zhu, 1997). This vast use of Chinese medicine illustrates the important value and tradition of its use in China. Both forms of medicine have their own departments of health, medical schools, research institutes, and hospitals. Furthermore, patients have the right to choose whether they receive western, Chinese medicine, or both treatments in approximately 95% of hospitals (Hesketh & Zhu, 1997). This combination allows for maximum effectiveness between the two discourses of medicine, as well as fosters the continuation of Chinese medicine traditions.

ACUPUNCTURE FOR FROZEN SHOULDER

With a prevalence of 2% to 3% across the globe, frozen shoulder or adhesive capsulitis is one of the most frequent explanations for shoulder injury and pain. Frozen shoulder usually consists of an extended loss of shoulder movement and pain, and is most prevalent in women between the ages of 40 and 70. Due to the range of injuries and symptoms, the treatment and prognosis differ between doctors and as a result there is still no clear treatment approach (Sun et al., 2001). Due to this lack of consensus, acupuncture as a possible treatment option has been explored.

Frozen shoulder is referred to as “shoulder at the age of 50 years” in traditional Chinese medicine. According to traditional Chinese medicine theory, frozen shoulder indicates a blockage of Qi and blood, deficiency of Yin, and is linked with fragility in the stomach and spleen. The blockage of Qi and blood results in pain and rigidity in the shoulder joints. Traditional Chinese medicine recommends that in order to unblock the Qi and blood, the rigid body part needs to be moved frequently. Through a combination

of physical exercise and acupuncture, a balance between the Yin and Yang can be achieved and health restored (Sun et al., 2001).

Historically, acupuncture has been used to treat frozen shoulder. Although further research is necessary to fully evaluate its efficiency, current research demonstrates its potential as a viable treatment option. In a single-blind randomized controlled trial by the University of Hong Kong, thirty-five participants with frozen shoulder were either assigned to an exercise group or an exercise and acupuncture group. Using the Constant Shoulder Assessment, shoulder function was evaluated at baseline, 6 weeks and 20 weeks. This study found that compared to the exercise group, the acupuncture and exercise group improved significantly more (Sun et al., 2001). This shows how the combination of both acupuncture and exercise, like prescribed in traditional Chinese medicine, offers an effective treatment method. Another study also found acupuncture to be beneficial in treating frozen shoulder. While this study did not use a control group, the large improvement in participants suggests that acupuncture can be a beneficial treatment. This study also showed that the quantity of treatments needed for improvement varies from participant to participant. This illustrates how responses to acupuncture treatment vary and as a result velocity of recovery can greatly differ (Tukmachi, 1999). Although there is not enough conclusive research to prove acupuncture can treat frozen shoulder, these findings point to the possibility of its efficiency.

ACUPUNCTURE FOR TENNIS ELBOW

Tennis elbow or lateral epicondylitis is prevalent in approximately 1-3% of the global population with a higher frequency in North America (Fink, Wolkenstein, Karst, & Gehrke, 2002; Trinh, Phillips, Ho, & Damsma, 2004). It affects both genders equally and generally occurs in people 40 years old or older (Johnson, Cadwallader, Scheffel, & Epperly, 2007). Although tennis elbow has a high prevalence, few treatments seem to be beneficial for the long-term. In the past 10 years, acupuncture has become increasingly used as treatment for the pain associated with tennis elbow (Trinh et al., 2004).

Tennis elbow is referred to as “elbow taxation, elbow pain and damaged sinews” in traditional Chinese medicine (Flaws & Sionneau, 2001). According to traditional Chinese medicine theory, a deficiency of blood, lack of Yin, and a lack of Qi causes tennis elbow. This deficiency mainly associated with fragility in the liver causes the tendons to become inflamed. Traditional Chinese medicine treatment options include herbal remedies, moxibustion, rest, and acupuncture that aim to restore balance and allow for a free flow of Qi, blood, and Yin (Flaws & Sionneau, 2001).

There have been many research studies on the effects of acupuncture on tennis elbow. Although further research is necessary, the National Institutes of Health asserts that due to promising research results, acupuncture can be considered a suitable treatment for tennis elbow (National Institutes of Health Consensus Conference., 1998). Although research results show promise, there is still mixed evidence to support acupuncture is effective in treating tennis elbow. Two systematic reviews and a meta-analysis found that acupuncture treatment provides only short-term pain relief that lasts from three days to two months (Bisset, Paungmali, Vicenzino, & Beller, 2005; Trinh et al., 2004; Trudel et al., 2004). Contrastingly, two systematic reviews, acknowledged the possibility of acupuncture's short-

term benefits, but determined that current research does not provide sufficient evidence to draw a conclusion (Assendelft et al., 2003; Green et al., 2002). In a placebo-controlled single blind clinical trial, 48 participants were either assigned to acupuncture treatment or to placebo acupuncture without any insertion of needles. Although this study didn't effectively follow-up with its participants, it concluded that traditional acupuncture is more effective at reducing pain than the control group. It also showed that the benefits of acupuncture for tennis elbow are primarily short-term (Molsberger & Hille, 1994).

ACUPUNCTURE FOR CARPAL TUNNEL SYNDROME

With an approximate prevalence of 1-3% across the globe, Carpal tunnel syndrome is the most frequent entrapment neuropathy (Atroshi et al., 1999). It affects people of all ages with a higher frequency in women (Bland, 2007). Carpal tunnel syndrome leads to pain, muscular dysfunction, paraesthesia and hypesthesia in the hand (Gerritsen et al., 2002; Palmer, Harris, & Coggon, 2006). According to the Bureau of Labor Statistics and U.S. Department of Health and Human Services, carpal tunnel syndrome is one of the most expensive and disabling upper-extremity disorders (Huisstede et al., 2010). Since the pathophysiological mechanisms for the pressure increase in the carpal tunnel are not fully understood, treatment and diagnosis are mostly clinically decided (Werner & Andary, 2002; Yao et al., 2012). Although there are many treatment options, none has been completely accepted as effective or ideal (Huisstede et al., 2010). Acupuncture, a common therapy in Asia for treating carpal tunnel syndrome, shows promise of being a possible treatment option (Yao et al., 2012).

According to traditional Chinese medicine, carpal tunnel syndrome is classified as "hand and finger tingling and numbness" and is associated with the liver (Flaws & Sionneau, 2001). A misbalanced Qi and blood deficiency primarily cause carpal tunnel syndrome. This blood deficiency and stagnation is the cause of the wrist pain. As a result, acupuncture and other herbal Chinese remedies can be used to "quicken the blood and dispel stasis" and as a result, reduce the pain and illness (Flaws & Sionneau, 2001).

Currently, there is very limited research on the effect of acupuncture on carpal tunnel syndrome. According to the subjective symptoms improvements reported, acupuncture has short-term benefits in treating carpal tunnel syndrome (Khosrawi, Moghtaderi, & Haghighat, 2012). A systematic review and meta-analysis concluded that although there is a possibility that acupuncture is beneficial for carpal tunnel syndrome, the lack of research does not provide sufficient evidence to support that claim (Sim et al., 2011). A randomized controlled trial that evaluated acupuncture compared to steroid treatment found that in the short-term, acupuncture is equally as effective as a low-dose steroid. This was measured through changes in nerve conduction

in participants with mild to moderate carpal tunnel syndrome (Yang et al., 2009). However, a study by Yao at the University of California, Davis Medical Center found different results; in a double-blind placebo controlled randomized trial, 41 participants were assigned to either an acupuncture or placebo acupuncture group for 6 weeks. Although both groups improved from baseline, no statistically significant difference in improvement between the control and experimental group was found (Yao et al., 2012). These studies provide mixed evidence; further research is necessary in order to fully evaluate the effectiveness of acupuncture on carpal tunnel syndrome.

METHODS

A survey was conducted on a sample of 57 participants who were either receiving acupuncture treatment from Beijing City Haidian District Shuangyushu Community Health Center or at Peking University Third Hospital. Surveys were administered in-person after a doctor from their respective center provided an introduction. All participants provided their written consent, and all questions were answered by the protocol director (after they were translated). All surveys were conducted in Chinese and therefore both the survey, consent form, and survey data were translated. This study, after being approved by Stanford University's Internal Review Board and granted Human Subject Approval, was conducted over a 2-week period in November-December 2015.

The survey inquired about the participant's diagnosis, age, gender, duration of symptoms, intensity of symptoms (ranked on a scale from zero to five), frequency of pain (categorized by either multiple times a day, once a day, 2-5 times a week, once a month, once a year, or could be written in), medications taken, other treatment

and frequency of acupuncture visits. It then asked open-ended questions regarding treatment, such as: Have you found acupuncture treatment to be beneficial? Do you believe your treatment was effective? Have you found acupuncture treatment to be an economic burden? Has your perception on acupuncture changed after receiving treatment? Has your perception on Chinese medicine changed after receiving treatment? Would you recommend acupuncture treatment to others? What made you choose acupuncture? Have you enjoyed the experience of receiving acupuncture? These questions were followed with a comment section that encouraged participants to share any other thoughts. All returned surveys were translated and then reviewed. They were sorted based into categories based on their diagnosis: carpal tunnel syndrome, frozen shoulder, tennis elbow, musculoskeletal disorders, neurological disorders and other. In total, 6 participant's responses were in the carpal tunnel category, 10 participants in the frozen shoulder category, 7 participants in the tennis elbow category, 8 participants in the musculoskeletal category, 11 participants in the neurological category and 15 participants in the other category.

In addition to the questionnaire component of the study,

“Acupuncture’s long history in Asia and Western Europe has demonstrated its usefulness in improving human health (Armstrong, 1972). Its increasing acceptance and demand as a treatment for chronic disorders and pain management has spurred an augmentation of scientific research on its validity.”

interviews were conducted with participants and acupuncture medics.

RESULTS AND DISCUSSION

Based on the data collected from the questionnaires, some trends were identified. In regards to age and gender, the mean age was 52 with a standard deviation of 17.2 and the majority of the participants (70.2%) were female (see Table 1). The small sample size of tennis elbow, carpal tunnel syndrome, and frozen shoulder did not allow for a significant trend to be established between age and gender. But overall, the higher female percentage and average age above 40 matches up with previous estimates. Many external factors could have influenced both the age and gender distribution

Category:	Mean Age (SD)	Female: n (%)	Male: n (%)
All Participants (n=57)	52 (17.2)	40 (70.2)	17 (29.8)
Carpal tunnel syndrome (n=6)	43 (23.4)	3 (50.0)	3 (50.0)
Tennis elbow (n=7)	51 (13.4)	4 (57.1)	3 (42.9)
Frozen shoulder (n=10)	52 (13.3)	6 (60.0)	4 (40.0)
Musculoskeletal conditions (n=8)	54 (14.2)	7 (87.5)	1 (12.5)
Neurological conditions (n=11)	52 (28.6)	8 (72.7)	3 (27.3)
Other (n=15)	55 (14.2)	12 (80.0)	3 (20.0)

Table 1. Age and gender distribution amongst research participants.

Characteristic	All Participants (n=57)	Carpal tunnel syndrome (n=6)	Tennis elbow (n=7)	Frozen shoulder (n=10)	Musculoskeletal conditions (n=8)	Neurological conditions (n=11)	Other (n=15)
Using Additional Treatment: n (%)	34 (59.6)	1 (16.7)	0 (0.0)	6 (60.0)	6 (75.0)	9 (81.8)	12 (80.0)
Additional Western Treatment	12 (21.0)	0 (0.0)	0 (0.0)	3 (30.0)	1 (12.5)	3 (27.3)	5 (33.3)
Additional Chinese Medicine	9 (15.8)	1 (16.7)	0 (0.0)	2 (20.0)	2 (25.0)	1 (9.0)	3 (20.0)
Additional Western and Chinese Medicine Treatment	13 (22.8)	0 (0.0)	0 (0.0)	1 (10.0)	3 (37.5)	5 (45.5)	4 (26.7)
Not Using Additional Treatment: n (%)	23 (40.4)	5 (83.3)	7 (100)	4 (40.0)	2 (25.0)	2 (18.2)	3 (20.0)

Table 2. Shows percentage of participants using additional treatment besides acupuncture. That percentage is then broken up to show what type of additional treatment they were using.

found. For example, time of data collection could have influenced gender distribution as well as a selection bias as females could be have been more willing to participate in this study. Although these factors could have influenced the data, overall, I observed a higher female percentage in clinics receiving acupuncture. This could be related to female's being more willing to accept Chinese medicine and use it as a treatment option.

The use of acupuncture as a sole treatment or complementary treatment option was assessed. The data shows that most participants took supplementary medication in addition to their acupuncture treatment. Most participants were either taking western and additional Chinese medicine or solely additional western medicine. Classifying participants by conditions, 16.7% of participants with carpal tunnel syndrome reported to be using another form of Chinese medicine along with acupuncture. Amongst acupuncture participants with frozen shoulder, 30% were taking Western treatment, 20.0% were using additional Chinese medicine techniques, and 10.0% were using both. Contrastingly, all participants with tennis elbow reported not taking any additional medication to acupuncture.

Chinese medicine has been around for more than 3,000 years in China. As the only healthcare system that incorporates both traditional Chinese medicine and western medicine, people in China view Chinese medicine in a different ways. The older generation, that grew up with predominately Chinese medicine tend to view it differently than the current generation. The data collected shows this trend, as a higher percentage of participants over the age of sixty indicated that they have always believed in Chinese medicine (Table 3). This, combined with the study's average participant age, points to the fact that older people tend to be more willing to get Chinese medicine treatment. This is further emphasized by interviews conducted where many of the older participants would remark how Chinese medicine is the only medical treatment that they perceive as effective. They were also all very excited that someone was studying it and repeated the importance of never doubting Chinese medicine as a treatment option. This strong view on the validity of Chinese medicine is not seen in the younger generation, where none of the participants between the ages of 18 to 30 indicated that they have always

Characteristic	Age: 18-30 (n= 6)	Age: 30-40 (n= 8)	Age: 40-60 (n= 19)	Age: 60+ (n= 17)
Using Additional Treatment: n (%)	4 (66.7)	6 (75.0)	12 (63.1)	11 (64.5)
Additional Western Treatment	0 (0.0)	2 (25.0)	7 (36.8)	1 (5.9)
Additional Chinese Medicine	0 (0.0)	4 (50.0)	2 (10.5)	5 (29.3)
Additional Western and Chinese Medicine Treatment	4 (66.7)	0 (0.0)	3 (15.8)	5 (29.3)
Not Using Additional Treatment: n (%)	2 (33.3)	2 (25.0)	7 (36.9)	6 (35.5)
Always believed in Chinese Medicine: n (%)	0 (0.0)	3 (37.5)	4 (21.0)	9 (52.9)
Have not always believed in Chinese Medicine: n (%)	6 (100)	5 (62.5)	15 (79.0)	8 (47.1)

Table 3. Shows the percentage of participants using additional treatment and their beliefs towards Chinese medicine.

believed in Chinese medicine. This does not mean that younger participants do not use traditional Chinese medicine; rather, these results imply that they do not see it as a sole treatment option and often view western medicine as more accurate.

In an interview with a doctor that conducts acupuncture, I asked whether they considered acupuncture to be an alternative form of treatment. They answered that they view acupuncture as complementary treatment that should be integrated with other treatments. This view highlights the Chinese healthcare system's perspective on effective treatment, a combination of western and Chinese medicine.

In regards to the benefits of acupuncture, most participants viewed acupuncture to be beneficial but a smaller percentage (although still majority) perceived acupuncture to be effective. An equal percentage of participants with carpal tunnel syndrome, tennis elbow, and other viewed acupuncture to be effective and beneficial (Table 4). Contrastingly, participants with frozen shoulder, musculoskeletal conditions, and neurological conditions perceived acupuncture to be beneficial but not as effective. This finding illustrates that there are other reasons besides effectiveness that people receive acupuncture treatment. For example, a lot of participants commented that acupuncture might not be as effective as Western medicine but has zero side effects and that acupuncture is a traditional form of medicine in China. This combined with the inherent relaxing value in acupuncture could explain the difference in the perception between effectiveness and beneficial. But overall, the majority of participants viewed it as both effective and beneficial, indicating how people are overall content with acupuncture as a treatment option.

In regards to the acupuncture practitioners' views on the effectiveness of acupuncture, during an interviews, they explained that the effectiveness of acupuncture varies by condition. They further explained that the effects of acupuncture can also vary by

Category:	Perceive Acupuncture to be Beneficial: n (%)	Perceive Acupuncture to be Effective: n (%)	Perception of Acupuncture Changed after Treatment: n (%)	Perception of Chinese Medicine Changed after Treatment: n (%)	Always Believed in Chinese Medicine: n (%)
All Participants (n=57)	54 (94.7)	51 (89.4)	41 (72.0)	31 (54.4)	14 (24.5)
Carpal tunnel syndrome (n=6)	5 (83.3)	5 (83.3)	5 (83.3)	2 (33.3)	2 (33.3)
Tennis elbow (n=7)	4 (57.0)	4 (57.0)	3 (42.9)	2 (28.6)	1 (14.3)
Frozen shoulder (n=10)	10 (100)	7 (70.0)	7 (70.0)	6 (60.0)	1 (10.0)
Musculoskeletal conditions (n=8)	8 (100)	6 (75.0)	5 (62.5)	3 (37.5)	1 (12.5)
Neurological conditions (n=11)	7 (63.6)	5 (45.5)	8 (72.7)	5 (45.5)	4 (36.4)
Other (n=15)	13 (86.7)	13 (86.7)	7 (46.7)	7 (46.7)	5 (33.3)

Table 4. Shows the distribution of beliefs in acupuncture and Chinese medicine

Category:	Acupuncture is not an Economic Burden: n (%)	Would Recommend Acupuncture: n (%)	Enjoyed Receiving Acupuncture: n (%)
All Participants (n=57)	49 (85.9)	53 (92.9)	43 (75.4)
Carpal tunnel syndrome (n=6)	6 (100)	5 (83.3)	5 (83.3)
Tennis elbow (n=7)	4 (57.0)	5 (71.4)	5 (71.4)
Frozen shoulder (n=10)	7 (70.0)	10 (100)	7 (70.0)
Musculoskeletal conditions (n=8)	7 (87.5)	8 (100)	6 (75.0)
Neurological conditions (n=11)	10 (90.9)	10 (90.9)	7 (63.6)
Other (n=15)	13 (86.7)	15 (100)	12 (80.0)

Table 5. Shows the percentage of participants who believe acupuncture is not an economic burden, would recommend acupuncture, and enjoyed receiving acupuncture. Age and gender distribution amongst research participants.

patient and can depend on the area applied. But after many years of conducting acupuncture, they witnessed that acupuncture is the most beneficial on facial paralysis, pain in limbs, sprain, insomnia, headaches and gynecological diseases.

Similarly, this study also collected data on how people's perception of acupuncture and Chinese medicine was influenced by treatment. The data indicates that the majority of participants' perception of acupuncture changed prior to treatment, but a smaller (but still majority) percentage of people's perception of Chinese medicine changed after treatment. A similar trend is also seen in all conditions.

In China's healthcare system, people pay by a fee for service method. Due to the shift from a public to private healthcare system, there is a growing disparity between the urban and rural citizens. Since this study was conducted in acupuncture clinics in Beijing, the population that took the survey is not representative of China's population as a whole. The data collected shows how in cities, for most people acupuncture is not an economic burden due to higher incomes and health insurance access (Table 5). However, some participants were not from the city and indicated that acupuncture did pose a significant economic risk. For example, one of the participants indicated that the reason that acupuncture is a financial burden for him is because he is a farmer in the rural areas of China. As a result, in order to get his treatment at that hospital he would have to pay it completely out of pocket. This illustrates the disparity between rural and urban areas, as rural citizens do not have access to the same benefits, have to travel for high quality care and have to pay for the expensive treatment completely on their own. This creates a disparity in health care and results in higher overall disparities.

In addition, a fee for service health system creates incentives for health professionals to misdiagnosis and provide incorrect but

more expensive treatments. This corruption in the health system has created mistrust in health professionals. In one interview, a participant commented that they only trust Chinese medicine, as western medics tend to overprescribe expensive and unnecessary treatments. The mistrust in doctors was further emphasized when I was shadowing a kidney stone removal surgery in a Chinese hospital, and during the surgery the doctors took many photos as well as collected the kidney stones in order to be able to prove to the patient that they actually received the surgery and that it was not a sham. This mistrust in the system and profit-driven incentives is a very big issue in China's health care as it creates a huge waste of resources, disparities in health, and can be dangerous to patients' lives.

LIMITATIONS

Limitations of the study include selection bias, limited acupuncture centers visited and other external factors. In regards to selection bias, there was no proper randomization. Participants' selection was based on the doctors' recommendation and if they were waiting for their treatment. In addition, participants were only recruited at two health centers in Beijing, China. Moreover, this study only recruited participants that were already receiving acupuncture treatment and as a result is unable to conclude anything about the general population's beliefs on acupuncture or Chinese medicine. Furthermore, all of these participants were receiving acupuncture treatment and as a result already had some thoughts on acupuncture in order to be willing to receive it as a treatment. Other external factors could have included people's willingness to participate in the study. Also acupuncture, just like any treatment, can range depending on the doctor. Since participants were recruited from a couple of clinics, the differences in doctor's treatment could have also potentially impacted the data. All of these factors could have affected the study outcome.

CONCLUSION AND RECOMMENDATIONS FOR FURTHER RESEARCH

The reviewed evidence suggests that participants and doctors view acupuncture to be effective and beneficial. In addition, after receiving treatment, the majority of participants indicated that their perception of acupuncture changed, although not as many said the same thing about Chinese medicine as a whole. The data indicates that the participants sixty years or older view Chinese medicine differently than the younger generation. Evidence of a health disparity and differing economic burdens between urban and rural citizens was found. Overall, most people would recommend and enjoyed receiving acupuncture.

Although there are several limitations, the survey data, interviews, and current literature indicates that acupuncture can be used as a complementary treatment option for conditions including frozen shoulder, tennis elbow and carpal tunnel syndrome.

More research in the field of acupuncture and Chinese medicine would help elucidate its validity. Due to China's unique history with Chinese medicine, large population, as well as current healthcare system, further research in China could help uncover more about the interactions between western and Chinese medicine. Furthermore, more research on the effects of acupuncture and other Chinese medicine remedies on carpal

tunnel syndrome, tennis elbow, and frozen shoulder could help elucidate the validity of acupuncture and Chinese medicine as a treatment option.

ACKNOWLEDGMENTS

I would like to acknowledge James Garth and Zhongling Shen for their invaluable help in translating throughout the research study. I would also like to acknowledge Guangyu Cao, Tang Yicun and Jialu Li for their help in translating the data. I am very grateful to Dr. Zejun Huo and Dr. Li Ba for allowing me to conduct my study in their clinics, introducing me to potential participants and answering all my questions. I would also like to acknowledge all the patients who agreed to participate in this study.

REFERENCES

- Anton, H. a. (1993). Frozen shoulder. *Canadian Family Physician*, 39, 1773–1778.
- Armstrong, M. (1972). Acupuncture. *American Journal of Nursing*, 72, 1582–1588.
- Assendelft, W., Green, S., Buchbinder, R., Struijs, P., Smidt, N., & Mellor, S. (2003). Tennis elbow. *BMJ*, 327, 329–330.
- Atroshi, I., Gummesson, C., Johnsson, R., Ornstein, E., Ranstam, J., & Rosén, I. (1999). Prevalence of carpal tunnel syndrome in a general population. *Jama*, 281, 153–158.
- Bisset, L., Paungmali, A., Vicenzino, B., & Beller, E. (2005). A systematic review and meta-analysis of clinical trials on physical interventions for lateral epicondylalgia. *British Journal of Sports Medicine*, 39, 411–422.
- Bland, J. (2007). Carpal tunnel syndrome. *BMJ*, 335, 343–346.
- Cohen, M. (2006). Exploring the other side of acupuncture. *Journal of Alternative and Complementary Medicine*, 12, 707–708.
- Eisenberg, D. M., Davis, R. B., Ettner, S. L., Appel, S., Wilkey, S., Van Rompay, M., & Kessler, R. C. (1998). Trends in alternative medicine use in the United States, 1990-1997: Results of a follow-up national survey. *JAMA: The Journal of the American Medical Association*, 280, 1569–1575.
- Finando, S., & Finando, D. (2011). Fascia and the mechanism of acupuncture. *Journal of Bodywork and Movement Therapies*, 15, 168–176.
- Fink, M., Wolkenstein, E., Karst, M., & Gehrke, A. (2002). Acupuncture in chronic epicondylitis: A randomized controlled trial. *Rheumatology*, 41, 205–209.
- Flaws, B., & Sionneau, P. (2001). *The Treatment of Modern Western Medical Diseases with Chinese Medicine: A Textbook & Clinical Manual*. Blue Poppy Press. Retrieved from <https://books.google.com/books?id=u2hkIdlxXgC>
- Gerritsen, A., de Krom, M., Struijs, M., Scholten, R., de Vet, H., & Bouter. (2002). Conservative treatment options for carpal tunnel syndrome: A systematic review of randomised controlled trials. *Journal of Neurology*, 249, 272–80.
- Green, S., Buchbinder, R., Barnsley, L., Hall, S., White, M., Smidt, N., & Assendelft, W. (2002). Acupuncture for lateral elbow pain. *Cochrane Database of Systematic Reviews*, 1, 1–16.
- Haake, M., Muller, H., Schade-Brittinger, C., Basler, H., & Schafer, H. (2007). German acupuncture trials for chronic low back pain. *Archives of Internal Medicine*, 167, 1892–1898.
- Hagberg, M., Morgenstern, H., & Kelsh, M. (1992). Impact of occupations and job tasks on the prevalence of carpal tunnel syndrome. *Scandinavian Journal of Work and Environmental Health*, 18, 337–345.
- Hesketh, T., & Zhu, W. (1997). Traditional Chinese medicine: One country, two systems. *BMJ*, 315, 115–117.
- Hsiao, W. C. (2014). Correcting past health policy mistakes. *Daedalus The Journal of the American Academy of Arts and Sciences*, 143, 53–68.
- Hudak, P. L., Cole, D. C., & Haines, A. T. (1996). Understanding prognosis to improve rehabilitation: The example of lateral elbow pain. *Archives of Physical Medicine and Rehabilitation*, 77, 586–93.
- Huisstede, B. M., Hoogvliet, P., Randsdorp, M. S., Glerum, S., van Middelkoop, M., & Koes, B. W. (2010). Carpal Tunnel Syndrome. Part I: effectiveness of nonsurgical treatments. *Archives of Physical Medicine and Rehabilitation*, 91, 981–1004.
- Johnson, G. W., Cadwallader, K., Scheffel, S. B., & Epperly, T. D. (2007). Treatment of lateral epicondylitis. *Am Fam Physician*, 76, 843–848.
- Khosrawi, S., Moghtaderi, A., & Haghghat, S. (2012). Acupuncture in treatment of carpal tunnel syndrome: A randomized controlled trial study. *Journal of Rehabilitation Medicine*, 17, 1–7.
- Langevin, H. M., Churchill, D. L., & Cipolla, M. J. (2001). Mechanical signaling through connective tissue: A mechanism for the therapeutic effect of acupuncture. *The Federation of American Societies for Experimental Biology*, 15, 2275–2282.
- Lee, B. Y., Lariccia, P. J., & Newberg, A. B. (2004). *Acupuncture in Theory and Practice Part 1: Theoretical Basis and Physiologic Effects*. *Clinical Perspectives in Complementary Medicine*, 11–18.
- Lewith, G. T. (1984). How effective is acupuncture in the management of pain? *The Journal of the Royal College of General Practitioners*, 34, 275–8.
- Lewith, G. T., & Machin, D. (1983). On the evaluation of the clinical effects of acupuncture. *Pain*, 16, 111–27.
- Moffet, H. H. (2009). Sham acupuncture may be as efficacious as true acupuncture: A systematic review of clinical trials. *The Journal of Alternative and Complementary Medicine*, 15, 213–216.
- Molsberger, A., & Hille, E. (1994). The analgesic effect of acupuncture in chronic tennis elbow pain. *British Journal of Rheumatology*, 33, 1162–1165.
- National Institutes of Health Consensus Conference. (1998). *Acupuncture*. *JAMA*, 280, 1518–1524.
- Palmer, K. T., Harris, E. C., & Coggon, D. (2006). Carpal tunnel syndrome and its relation to occupation: A systematic literature review. *Occupational Medicine*, 57, 57–66.
- Pomeranz, B., & Chiu, D. (1976). Naloxone blockage of acupuncture analgesia: Endorphin implicated. *Life Sciences*, 19, 1757–1762.
- Richardson, P. M. H., & Vincent, C. A. (1986). Acupuncture for the treatment of pain: a review of evaluative research. *Pain*, 24, 15–40.
- Scharf, H.-P., Mansmann, U., Streitberger, K., Witte, S., Krämer, J., Maier, C., ... Victor, N. (2006). Acupuncture and knee osteoarthritis: A three-armed randomized trial. *Annals of Internal Medicine*, 145, 12–20.
- Sierpina, V. S., & Frenkel, M. a. (2005). Acupuncture: A clinical review. *The Southern Medical Journal*, 98, 330–337.
- Sim, H., Shin, B.-C., Lee, M. S., Jung, A., Lee, H., & Ernst, E. (2011). Acupuncture for Carpal Tunnel Syndrome: A systematic review of randomized controlled trials. *The Journal of Pain*, 12, 307–314.
- Sun, K. O., Chan, K. C., Lo, S. L., & Fong, D. Y. T. (2001). Acupuncture for frozen shoulder. *Hong Kong Medical Journal*, 7, 381–391.
- Tatewaki, M., Harris, M., Uemura, K., Ueno, T., Hoshino, E., Shiotani, A., ... Takahashi, T. (2003). , 285, R862–R872.
- Trinh, K. V., Phillips, S.-D., Ho, E., & Damsma, K. (2004). Acupuncture for the alleviation of lateral epicondyle pain: a systematic review. *Rheumatology*, 43, 1085–90.
- Trudel, D., Duley, J., Zastrow, I., Kerr, E. W., Davidson, R., & MacDermid, J. C. (2004). Rehabilitation for patients with lateral epicondylitis: A systematic review. *Journal of Hand Therapy*, 17, 243–266.
- Tukmachi, E. S. (1999). Frozen shoulder: A comparison of western and traditional chinese approaches and a clinical study of its acupuncture treatment. *Acupuncture In Medicine*, 17, 9–21.
- van Rijn, R., Huisstede, B., Koes, B. W., & Burdorf, A. (2009). Review and the Tunnel Carpal a Systematic Review. *Scandinavian Journal of Work, Environment and Health*, 35, 19–36.
- Werner, R. a., & Andary, M. (2002). Carpal tunnel syndrome: Pathophysiology and clinical neurophysiology. *Clinical Neurophysiology*, 113, 1373–1381.
- Yang, C.-P., Hsieh, C.-L., Wang, N.-H., Li, T.-C., Hwang, K.-L., Yu, S.-C., & Chang, M.-H. (2009). Acupuncture in Patients With Carpal Tunnel Syndrome. *The Clinical Journal of Pain*, 25, 327–333.
- Yao, E., Gerritz, P. K., Henricson, E., Abresch, T., Kim, J., Han, J., ... Zhao, H. (2012). Randomized controlled trial comparing acupuncture with placebo acupuncture for the treatment of Carpal Tunnel Syndrome. *The American Academy of Physical Medicine and Rehabilitation*, 4, 367–373.
- Zheng, X., & Hillier, S. (1995). The reforms of the Chinese health care system: Country level changes: The Jiangxi study. *Social Science & Medicine*, 41, 1057–1063.

SHARON WULFOVICH

Sharon Wulfovich is a junior at Stanford University, majoring in Human Biology with a focus on epidemiology and preventative medicine. This fall, she studied abroad in Beijing, China where she studied mandarin and history at Peking University and conducted research on Chinese medicine. She currently conducts research at the Stanford Hospital's Emergency Room and Stanford Bariatric and Metabolic Interdisciplinary Clinic.



SAKTI SRIVASTAVA, MBBS, MS

Associate Professor, Surgery and Bioengineering (courtesy)
Dr. Sakti Srivastava trained as an orthopedic hand surgeon and is currently Associate Professor in Department of Surgery at Stanford University School of Medicine. He serves as Division Chief of Clinical Anatomy and Director of the newly established Digital MEDiC initiative. He has a research interest in educational technologies and surgical simulation. Dr. Srivastava's special interest is in integrative medicine and holistic healing. He teaches a number of courses related to complementary and alternative medicine and studies "medical mysteries".

Disentangling the Trilemma: Factors that Shape Policy Choice

Lauren Iannolo

Georgetown University

This paper uses panel data from forty-nine countries to examine factors that influence a country's position on the Mundell-Fleming Trilemma. It specifically investigates exports as a percentage of GDP, financial development, usage of foreign currency denominated debt, and the presence and frequency of financial crises. This paper is unique in that it characterizes a country's Trilemma policy based on its fundamental economic characteristics instead of examining the effects the policies of exchange rate stability, monetary independence, and capital openness have on the economy. Using data from the World Bank and existing literature, we find that there are trends in how countries choose their Trilemma policy. Countries that rely on exports tend to fix their exchange rate to keep prices stable and a financial crisis in the last five years tends to lead to a more freely floating exchange rate and greater capital controls. The results demonstrate that there is in fact a pattern by which countries choose their economic policies that is worth studying in greater detail.

INTRODUCTION

In a globalized world, it is more important than ever for a country to build international policy to fit its needs. Countries' needs vary greatly depending on a large number of economic factors unique to their demographics, geography, and history. A particularly important policy choice in international economics is a country's Trilemma decision. The Mundell-Fleming Trilemma states that a country cannot pursue the three policies of a fixed exchange rate, complete financial openness, and independent monetary policy at the same time [1]. Each of these policies does not only affect the country that chooses the policy, but it also has implications, through trade and international finance, for other nations, making it an important and interesting foreign policy tool.

This paper adds to existing literature by employing a new take on Trilemma policy research. Instead of examining the effects of a particular Trilemma policy stance, we try to pinpoint the factors that contribute to the policy in the first place. It could be the case that there is an optimal choice for a country based on its economic characteristics. We find that based on the specified economic factors of exports as a percentage of GDP, financial development, quantity of foreign currency denominated debt, and the presence and frequency of financial crises, countries seem to be recognizing their preferred Trilemma policy choice and acting upon it.

LITERATURE REVIEW

Existing literature has examined the effect on exports as a percentage of GDP and the level of financial development given a particular Trilemma policy stance. Previous papers tend to differentiate between the optimal choice of developed countries and developing countries [2,3,4]. Calvo and Reinhart determined that interest rate policy is becoming even more important to

countries as a means to stabilize their currency as well as control inflation. They found that variance in interest rates is greater when there is a history of greater inflation, which tends to be the case more often in developing countries [2]. These findings seem to indicate that developing countries would value a policy stance that allows them to fix their exchange rate and/or have an autonomous monetary policy. Other studies have shown that there is a correlation between a fixed exchange rate in developing countries and a favorable economic outcome. Levy-Yeyati and Sturzenegger find that fear of appreciation of currency in developing countries led to faster output and productivity growth over the long run, specifically impacting the long-run component of GDP. They cite China's reluctance to float its exchange rate as a strategy that is supposed to preserve the competitiveness of its exports [3]. Rogoff, Aghion, Bacchetta, and Ranciere's research supports this characterization by finding evidence that the more financially developed a country becomes, the better it will do with a more flexible exchange rate [4].

Existing literature has also examined the relationship between exchange rates, large amounts of foreign currency denominated debt, previous financial crises, and financial fragility [2,5,6,7,8]. Large amounts of foreign currency denominated debt are known as "original sin" because countries whose liabilities are denominated in foreign exchange are by definition unable to hedge against currency fluctuations [5]. They are therefore exposed to the risk of an unfavorable change in the exchange rate in addition to the usual risks taking on debt brings. For this reason, Goldstein and Turner argue that countries with a large amount of foreign currency denominated debt would want to avoid a depreciation of their currency because it can destroy the net worth of firms and initiate a wave of insolvencies, a financial crisis, and a steep fall in economic growth [6]. Additionally, resulting financial crises themselves can also impact the exchange rate policy of countries. Calvo and Reinhart also noted that after the financial crises of the 1990s, intermediate exchange rate regimes or the "middle

ground” were disappearing and that countries were being driven toward “corner solutions” of one extreme, fully fixed or fully floating, exchange rates. Examples they cite include the creation of the European Union, East Asian countries’ abandonment of their pegs, and South American countries’ dollarization [2].

THEORETICAL MODEL

This paper uses the assumptions of the Mundell-Fleming Trilemma, which states that a country cannot pursue the three policies of a fixed exchange rate, complete capital openness, and independent monetary policy at the same time [1]. It is important to note that the policies are linearly related to one another. See Figure 1 below for discussion of the Trilemma’s depiction and significance.

Model Specifications

The Trilemma is based on uncovered interest rate parity (UIP). UIP is given by the following equation in which, for example, the

$$(1 + i_s) = (1 + i_e) \frac{E_{s/\epsilon}^e}{E_{s/\epsilon}} \quad \text{Equation 1}$$

two currencies are dollars and euros [10]:

In this equation, i_s specifies the U.S. interest rate while i_e specifies the interest rate in the euro zone. $E_{s/\epsilon}$ represents the spot dollar/euro exchange rate and $E_{s/\epsilon}^e$ represents the expected future exchange rate. For example, $\frac{\Delta E_{s/\epsilon}^e}{E_{s/\epsilon}}$ would represent the expected dollar depreciation. The left side of the equation indicates the gross U.S. deposit dollar return and the right side of the equation indicates the gross euro deposit (expected) dollar return for a “risk neutral” investor [10]. This means that the investor does not care that the left side of the equation might be certain while the right side is not. To a “risk neutral” investor, the returns of investing in either side of the equation are equal and he or she is indifferent.

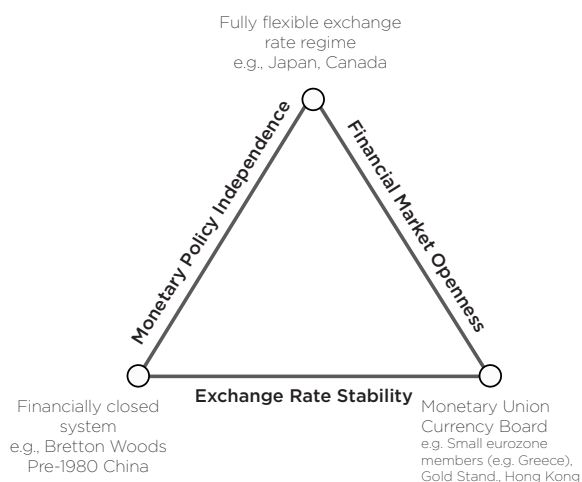


Figure 1. Mundell-Fleming Trilemma. The Trilemma model is typically depicted as a triangle. Countries do not have to fall at vertices alone. Countries can fall “inside” the triangle, meaning that they are pursuing some combination of the three policies instead of choosing two and executing them to their extremes. The vertices represent countries that have chosen policy extremes instead [9].

Model Implications

The Mundell-Fleming Trilemma holds true because pursuing all three of the policies at once would not work under UIP. UIP states that the depreciation or appreciation between two currencies will equal the interest rate differential between them, but this cannot be the case under a pegged exchange rate. Somehow the interest rates still have to equalize. This must mean that the pegging country cannot have control over its monetary policy (the way it controls its interest rate) and exchange rate unless it is instituting capital controls in the foreign exchange market, which means it is sacrificing financial openness [10].

Based on the model outlined above and conclusions drawn from the literature review, we should see a pattern in how countries choose their Trilemma policies. A country that is less financially developed might benefit from a fixed exchange rate because it would reduce risk of asset fluctuation in the eyes of foreign investors. The mercantilist view states that countries with high levels of exports would want to keep their currencies devalued so to keep their exports relatively cheap to potential buyers [3]. If a country has a history of inflation crises, it might value autonomous monetary policy more since through adjusting the domestic interest rate, the central bank can attempt to fight inflation [2]. If a country has more foreign currency denominated debt, depreciation in the home currency would increase the debt burden, so in this case, the country might value being able to prevent large depreciations at once [6].

DATA AND DESCRIPTIVE STATISTICS

The data sample is made up of yearly country specific data from 1993 through 2008 for 49 countries. This paper relies on economic data collected from The World Bank and the IMF, as well as sources explained subsequently.

Trilemma Policy Choice

The three Trilemma policy choices are measured using Aizenman, Chinn, and Ito’s indices, the exchange rate stability index, the monetary independence index, and the degree of financial integration index [11]. The exchange rate stability index is defined as the inverse of the standard deviation of the monthly rate of currency depreciation and a value closer to 1 means an exchange rate is more fixed. The monetary independence index is the correlation of a country’s interest rate with a base country’s interest rate and a value closer to 1 means that the country has more monetary independence. The degree of financial openness is measured with the Chinn-Ito capital openness index (KAOPEN), which is based off information regarding restrictions found in the IMF’s Annual Report on Exchange Arrangements and Exchange Restrictions. A KAOPEN value closer to 1 means that the country’s economy is more open. The indices are linearly related to one another to mimic the linearity of the Trilemma choices in theory [11]. For example, all indices cannot increase at once, if one increases some combination of the other two must decrease. See table 1 for a summary of Trilemma index statistics.

Foreign Currency Denominated Debt and Original Sin

Foreign currency denominated debt and original sin will be measured using a combination of Eichengreen, Hausmann and Panizza’s original sin index (OSIN1) and Reinhart’s data for the

YEARS		EXCHANGE RATE STABILITY INDEX, MEAN				
	Developed	Developing	LATAM	EU	SE Asia	
1993-1999	0.536	0.494	0.578	0.568	0.485	
2000-2008	0.667	0.481	0.573	0.738	0.479	
Overall	0.610	0.487	0.575	0.664	0.481	

YEARS		MONETARY INDEPENDENCE INDEX, MEAN				
	Developed	Developing	LATAM	EU	SE Asia	
1993-1999	0.289	0.481	0.500	0.277	0.439	
2000-2008	0.142	0.427	0.445	0.125	0.360	
Overall	0.206	0.450	0.470	0.192	0.395	

YEARS		KAOPEN (CAPITAL CONTROLS INDEX), MEAN				
	Developed	Developing	LATAM	EU	SE Asia	
1993-1999	0.920	0.448	0.513	0.846	0.542	
2000-2008	0.962	0.542	0.681	0.950	0.522	
Overall	0.943	0.503	0.607	0.903	0.531	

Table 1. Trilemma index summary statistics. LATAM stands for Latin America, EU stands for European Union, SE Asia stands for Southeast Asia. On average, developing countries float their interest rate more than developed countries. Many of the developed countries belong to the EU and therefore have their currencies fixed with the other member nations, earning them a 1 on the exchange rate stability index. Developing countries also value monetary independence and capital controls more than developed countries. The difference in the use of capital controls by developed and developing countries is especially stark, with developing countries implementing them much more often. There are clearly differences between the tendencies of Latin America, the European Union, and Southeast Asia. The indices also vary over time. For example, the EU became much more open with more stable exchange rates at the cost of monetary policy independence.

Region	No. of Countries	OSIN1	externdebt	exportspercen	creditpercen
		1993-2008	1993-2008	1993-2008	1993-2008
North America	3	0.663	31.082	22.905	103.473
Central America	4	0.997	45.687	40.954	38.112
South America	7	0.981	41.876	24.129	32.876
Africa	4	0.971	38.654	41.104	73.869
Western Europe	16	0.635	78.692	35.939	98.619
Eastern Europe	4	0.998	49.856	31.396	25.967
Central Asia	5	0.888	21.710	24.516	90.273
South-East Asia	5	0.960	70.546	85.795	81.755
Pacific/Oceania	2	0.874	61.071	23.118	99.530
Developed	24	0.697	67.619	38.574	103.050
Developing	26	0.967	43.953	36.095	48.318

Table 2. Summary Statistics of Selected Key Variables by Region. Table represents arithmetic mean of variables by region.

external debt to GDP ratio [12, 13]. Eichengreen, Hausmann and Panizza define OSIN1 for a country as one minus the fraction of debt denominated in a domestic currency divided by total debt issued by the country [12]. Countries that have a higher fraction of their debt denominated in a foreign currency will have higher index values. To take into account that although a country might have all its debt denominated in a foreign currency, the size of the debt could be insignificant, we then multiply the OSIN1 index by the external debt ratio. The external debt ratio gathered from Reinhart's data is total gross external debt divided by GDP [13]. We use this newly created measure of foreign currency denominated debt to measure a country's burden of external debt.

Types and Frequency of Financial Crises

The types of financial crises we examine are currency crises, inflation crises, external debt crises, and banking crises since these most closely affect a country's economic foreign policy choice. We obtain the data from the "This Time is Different" spreadsheets located on Carmen Reinhart's website [14].

Level of Financial Development

This paper borrows from Rogoff, Aghion, Bacchetta, and Ranciere who use the credit to GDP ratio as a measure of a country's financial development [4]. The greater a country's access to credit, the more financially developed it is considered.

Instruments for Export Magnitude

Free trade agreement data was collected from the WTO's Regional Trade Agreements Information System [15]. Only bilateral free trade agreements that covered goods which entered into force before or during 1993-2008 were examined and counted towards the total. The seaport data was collected by country from Searates.com by counting the locations listed for a given country [16].

EMPIRICAL STRATEGY

Endogeneity, Need for Instruments, and the 2SLS Model

In the case of my model, a simple Ordinary Least Squares (OLS) regression may include endogeneity. For example, it is possible that the degree to which the exchange rate is fixed or floating may affect exports as a percentage of GDP if the theory is in fact true that depreciated exchange rates can contribute to an increase in the current account and financial growth. Since the other two Trilemma indices are linearly related to the exchange rate index, it may be an issue with them as well. To attempt to control for this, I use a Two-Stage Least Squares (2SLS) model with instrumental variables. The three instruments I have chosen are the number of bilateral trade agreements a country is a party to in a given year, whether or not the country is a coastal country, and the number of ports that a country has. Below are the 2SLS specifications:

$$Y_{it} = \alpha_0 + \beta_1(externdebt)_{it} + \beta_2(crisislast5)_{it} + \beta_3(exportspercen)_{it} + \beta_4(creditpercen)_{it} + X_{it}\beta_5 + \sum \lambda_i(year) + \epsilon_{it}$$

Equation 2

$$Z_{it} = \phi_0 + \delta_1(freetrade)_{it} + \delta_2(numports)_{it} + \delta_3(coastal)_{it} + X_{it}\delta_4 + \sum \lambda_i(year) + \epsilon_{2it}$$

Equation 3

Where *i* indicates a country in year *t*, *Y* is one of the three Trilemma indices, *externdebt* measures the amount of original sin/foreign currency denominated debt a country has, *crisislast5* is

a binary variable indicating whether the country experienced a currency, external debt and/or banking crisis within 5 years prior to the given year. *Exportspercen* is exports as a percentage of GDP and *creditpercen* is domestic credit as a percentage of GDP. $\beta_1, \beta_2, \beta_3, \dots, \beta_5$ estimate the effects of the above variables on the Trilemma indices. *X* is a vector of control variables that vary by country, including the GDP per capita and inflation level, and ϵ_{it} is the error term. To control for the fact that there could be separate yearly trends in the data, I add in year fixed effects.

In equation (2) where I employ the instruments, *Z* is *exportspercen*, *freetrade* is the number of bilateral free trade agreements a given country is engaged in in a given year, *numports* indicates the number of ports and container terminals that a country has, and *coastal* is a binary variable indicating whether or not the country rests on a major body of water (an ocean or a sea). The f-statistic for the joint significance of the instruments was 58.84, which is significant at the 1% level.

Supplemental Regression for Further Detail

On another note, the importance of the East Asian crises of the late 1990's emerged during the literature review phase. The crises began in July of 1997 in Thailand when a run on the Thai baht turned into currency crises that spread around East Asia and had ripple effects throughout the rest of the world [17]. These "Asian miracle" countries that experienced rapid growth throughout the 1990's were suddenly in trouble. A major cause of the crises was the fact that many of the Asian countries relied on foreign currency denominated debt [17]. For this reason, we include a regression that looks to see if there is a significant difference in Trilemma

Variables	(1) Exchange Rate Stability Index	(2) Monetary Independence Index	(3) KAOPEN
Exports/ GDP	0.00367*** (0.001) [0.008]	-0.00257*** (0.001) [0.001]	0.00115 (0.001) [0.313]
External Debt	-0.00151*** (0.000) [0.001]	0.00096*** (0.000) [0.000]	-0.00075*** (0.000) [0.010]
Crisis in the Last 5 Years	-0.04985* (0.026) [0.053]	-0.00105 (0.016) [0.949]	-0.03422 (0.023) [0.144]
Domestic Credit/GDP	-0.00043 (0.000) [0.176]	-0.00017 (0.000) [0.327]	0.00000 (0.000) [0.993]
Year FE	YES	YES	YES
Observations	777	772	772
Number of Countries	49	49	49
R-squared	0.04777	0.20872	0.40946
Robust standard errors in parentheses, p-values in brackets *** p< 0.01, ** p<0.05, * p<0.1			

Table 3. 2SLS Model Statistics. The data in this table are significant and suggest it is possible to comment on the Trilemma stance a country may choose given the economic characteristics of a country.

policy specifically regarding foreign currency denominated debt. In other words, was the East Asian Crisis a watershed moment that changed the policies pursued by countries with large amounts of foreign currency denominated debt? If the answer is yes, we should see the slope change on the external debt term after the crisis.

RESULTS

Please refer to Table 3 throughout this section for specific data points.

Exchange Rate Stability Index Coefficients

We can see that on average, if a country increased its foreign currency denominated debt to GDP, the exchange rate stability index would decrease. This means that countries that are highly indebted in foreign currencies would tend to float their currency more than their counterparts. Earlier, we discussed that Goldstein and Turner hypothesized that countries would want to prevent rapid accelerations in currency mismatch and this is consistent with those findings. Floating exchange rates tend to have smaller, more frequent changes rather than a large adjustment made by the government at once as is the case with fixed exchange rates.

The effect of external debt is not especially large on the exchange rate stability index, but the regression also only explains about 5% of the variation in the exchange rate stability index, so it's significance to the 1% level is notable. We can safely assume that external debt is a contributing factor to the Trilemma policy choice and that it is negatively correlated with exchange rate stability.

The variable we instrumented for, exports as a percentage of GDP, is significant at the 1% level in the exchange rate stability index regression. An increase in exports is positively correlated with an increase in the index and a more fixed exchange rate confirming what we discussed previously. This effect is notable in both magnitude and significance, making it an important determinant of Trilemma policy. Similarly, countries with higher domestic credit as a percentage of GDP (the variable that measured financial development) are more likely to float their exchange rates. When we consider these results together with the results of the exports/GDP variable, it shows that countries do recognize which policies theory dictates spur growth given their own fundamentals (financial development, reliance on exports) and pursue them. However, with a higher p-value, the effect of

financial development on the exchange rate is not as persuasive as the effect of exports/GDP.

Our crisis in the last 5 years coefficient is highly significant, showing us that if a country experienced a crisis in the last 5 years, it is more likely to float its exchange rate all else equal. As mentioned earlier, many economies experienced financial turmoil during this period. For example, Asian economies were forced to drop their pegs under pressure from a myriad of problems. It seems that countries did undergo changes in their Trilemma policy as a result of contemporary crises.

KAOPEN (Capital Openness Index) Coefficients

The coefficient on the external debt term indicates that on average, countries with higher amounts of foreign currency denominated debt tend to place more emphasis on a Trilemma policy that includes capital controls. Currency mismatch, as previously discussed, makes countries vulnerable. By instituting capital controls, countries can try to manage some of the volatility associated with outside investment. Capital controls allow the government and economic institutions to prevent currency speculation, which could cause the currency to be vulnerable to overly rapid fluctuations. Since our results illustrated the tendency to let the exchange rate float, it makes sense that the government of a country highly indebted in a foreign currency would want to maintain all the protection it can in other ways in the event of a crisis.

The coefficient on the crisis in the last five years term indicates that countries that have experienced a crisis recently are more reluctant to open up their economy and more likely to institute capital controls. Although capital controls are a contentious issue in the related literature, economists have on occasion endorsed them. For example, Paul Krugman wrote to the Prime Minister of Malaysia in an open letter printed in Fortune in 1998: "Currency controls are a risky, stopgap measure, but some gaps desperately need to be stopped" [18].

Monetary Policy Independence Index Coefficients

The coefficients between the monetary independence index and exports as a percentage of GDP and domestic credit as a percentage of GDP are supported by the linear relationship of the Trilemma indices. It follows that in countries with high levels of exports, since there is an emphasis on pursuing both a more fixed exchange rate as well as an open economy, that there must be a trade-off among the policies. Clearly these countries have sacrificed monetary independence for the policies of exchange rate stability and fewer capital controls.

Supplemental Results Coefficients

The supplemental results for the capital openness (KAOPEN) index are the most noteworthy. Before the East Asian crisis, countries with higher amounts of external debt chose Trilemma policies that left their economies more open. However, after the crisis, these countries tended to institute more capital controls (lower levels of the index) the more debt they had, which can be seen in Figure 2 (below). The data points are the averages for individual countries before and after the crisis and are colored based on the pre-1998 or post-1998 average. It is clear from Figure 2 that the slope of the trend line decreased significantly, demonstrating the post-crisis emphasis on capital controls. This drastic change in policy after the East Asian crisis reflects the

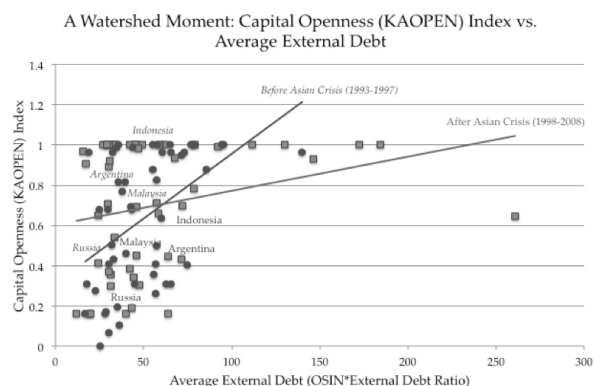


Figure 2. Relationship between the KAOPEN Index and Average External Debt before and after the East Asian Crisis.

advice of Paul Krugman discussed earlier. Specific countries are labeled that experienced some of the largest changes in policy, most notably, the recipient of Krugman's letter-- Malaysia. The East Asian crisis was in fact a watershed moment that changed many opinions on capital controls and necessitated their increased introduction into economic policy.

CONCLUSION

The results show that there are trends in the way countries choose their economic foreign policy and that these trends are based on their fundamentals. It is clear that export-heavy countries are making decisions about their exchange rate with their export dependence in mind, just as it is clear that the less financially developed are making Trilemma decisions with their long run growth rate in mind. This paper also supports the work of others, which suggests that there is a clear difference in the effects of policies on developing and developed countries, especially those that have experienced a financial crisis. It suggests that not only is there a difference in the effects of policy, but there also should be a difference in strategy if countries wish to get ahead.

The results of this paper could be furthered if the variables examined were looked at over a longer period of time and even more recently, after the Great Recession, to see if they still hold true. It is important to continue studying the strategies behind economic policies and their effects. Those who understand the tools will be better equipped to use them in the future.

ACKNOWLEDGEMENTS

This paper could not have been possible without the help of Barry Eichengreen, Ricardo Hausmann, and especially Ugo Panizza who personally supplied their data for the original sin indices, as well as other control variables, which were crucial to this thesis. I also owe a huge thank you to my senior thesis advisor, Bob Cumby for making the researching and writing process such a fun and educational experience.

REFERENCES

1. Bearce D. Monetary Convergence Hypothesis. In: Mansfield E, Martin L,

eds. Monetary Divergence. Ann Arbor: University of Michigan, 2007:13-17.

2. Calvo G, Reinhart C. Fear of Floating. *Quarterly Journal of Economics* 2002; 117:379-408.
3. Levy-Yeyati E, Sturzenegger F. Fear of Appreciation. *World Bank Policy Research Working Paper* 4387 2007.
4. Aghion P et al. Exchange Rate Volatility and Productivity Growth: The Role of Financial Development. *Journal of Monetary Economics* 2009; 52:494-513.
5. Eichengreen B Hausmann R. Exchange Rates and Financial Fragility. *NBER Working Paper* 7418 1999.
6. Goldstein M Turner P. Controlling Currency Mismatches in Emerging Markets. Washington, DC: Institute for International Economics, 2004.
7. Saibene G Scicouri S. Effects of Currency Devaluations on the Economic Growth in Developing Countries: The Role of Foreign Currency Denominated Debt. *International Atlantic Economic Society* 2012; 40:190-209.
8. Reinhart C. This Time is Different Chartbook: Country Histories on Debt, Default, and Financial Crises. *NBER Working Paper* 15851 2010.
9. Aizenman J. "The Impossible Trinity (aka The Policy Trilemma)," unpublished.
10. Feenstra R Taylor A. *International Economics* 2nd Ed. New York:Worth Publishers, 2011.
11. Aizenman J Chinn MD Ito H. The "Impossible Trinity" Hypothesis in an Era of Global Imbalances: Measurement and Testing. *Review of International Economics* 2013; 21:447-458.
12. Eichengreen B Hausman R Panizza U. The Pain of Original Sin. *Kennedy School of Government, Harvard University* 2003.
13. Reinhart C, Rogoff K. This Time is Different Chartbook: Country Histories on Debt, Default, and Financial Crises. *National Bureau of Economic Research* 2010.
14. Reinhart C. "Dates for Banking Crises, Currency Crashes, Sovereign Domestic or External Default (or Restructuring), Inflation Crises, and Stock Market Crashes (Varieties)." Internet: <http://www.carmenreinhart.com/data/browse-by-topic/topics/7/>, [Feb. 25, 2015].
15. "WTO Regional Trade Agreements Information System." Internet: <http://rtais.wto.org/UI/PublicSearchByCrResult.aspx>, [Feb. 25, 2015].
16. "World Sea Ports." Internet: <http://www.searates.com/maritime>, [Feb. 25, 2015].
17. Moreno R. "What Caused East Asia's Financial Crisis?" *Federal Reserve Bank of San Francisco*, 1998. Internet: <http://www.frbsf.org/economic-research/publications/economic-letter/1998/august/what-caused-east-asia-financial-crisis/>, [Feb. 25, 2015].
18. Neely C. An Introduction to Capital Controls. *Federal Reserve Bank of St. Louis Review*, 1999:1.



LAUREN IANNOLO

In May 2015, Lauren Iannolo graduated magna cum laude from Georgetown University with a major in Economics and a minor in Computer Science. She originally became interested in economics after learning about the Federal Reserve's role in response to the 2008 Financial Crisis during Principles of Macroeconomics her freshman year. On campus, she was involved with Smart Woman Securities, an organization that educates young women in financial literacy, which gave her the opportunity to dine with Mr. Warren Buffett in Omaha, NE during her junior year. She has interned for the Department of the Treasury, the Office of Administration and the Council of Economic Advisers at the White House, and is very interested in public policy as it relates to economics and finance. Since graduation, she has been working full-time as a consultant in the International Arbitration area of Navigant Consulting in Washington, DC.

Location Models in the Art Market: Artists as Firms Visualized on a Product Space

Ricardo R. Rodríguez-Padilla

Columbia University

While economists have contributed significantly to our understanding of the labor supply decision of modern-day artists, they have paid comparatively limited attention to the latter's product design and differentiation choices. Thus, we lack an empirical approach modeling how artists—like firms in a more “traditional” market—choose to design their products in terms of style and content, given the offerings of the other artists they “compete” with. While such a model might at first seem trivial or idiosyncratic, the factors that motivate production of art by artists should be of interest from both a social and economic perspective, and of relevance to direct and indirect participants of what is a salient, multibillion dollar global market (TEFAF Art Market Report, March 9, 2015).

This paper will analyze artistic product design and differentiation choice by visualizing artists—the central “firms” of the art market who produce the essential “product” of the market itself—as a set of offerings on a product space. I find that it is possible to create a location model for a set of artists that could allow us to study maximal vs. minimal product differentiation in the art market. Thus, we could begin probing whether an artist who just entered the market would find it more profitable to create art similar in style and/or content to other, established and successful artists—minimal product differentiation—or as unique and original as possible—in the language of location models, maximally differentiated. Such a model could also have implications for understanding the process of innovation in art, how art movements evolve, and how artists become successful, among other topics. The focus will be on the visual arts and in particular on painting, though of course most of the theories and models may be applicable to other mediums as well.

AN ECONOMIC LOOK AT ARTISTIC PRODUCTION

In considering artistic production, most economists have focused on labor supply—what drives an artist to enter the art market and supply hours of artistic labor. They have focused on the effect of arts and non-arts wages (the wages artists may find in careers and part-time jobs that will complement their earnings), creativity, earnings outlooks compared to other careers, and nonmonetary rewards [1]–[3]. (Casacuberta C, Gandelman N, preliminary draft of “The Artist Labor Supply Model” 2006). Yet, comparatively scant attention has been paid to the factors driving their choice of what to produce in the time they have allocated to artistic production—if they have decided to enter the market at all.

William Bryant and David Throsby look at a creative optimization decision, in which an artist chooses the amount of “creative effort” to invest in a particular work [4]. This choice is conditioned on the artist's expectation of the reaction to the work in two main domains: the “market” and the “world of ideas”, or “artworld.” They assume the market only cares about the economic value of the product, while the artworld is only concerned with its cultural value. The artist, naturally, is assumed to care about both. Like in earlier work by Throsby [5], this model is realistic in that it constrains an artist to sustain a level of consumption. Thus, financial considerations have a role to play in

her choice in that she needs to produce for the market as well as the artworld.

Other models have artists working part-time in both the artistic world and non-artistic world [5]. Others further complement the non-arts earnings with earnings not directly related to an artist's labor activities in a given period; instead, the artist also earns income from royalties from past work (Casacuberta C, Gandelman N, preliminary draft of “The Artist Labor Supply Model” 2006). Artists of course must earn these non-arts market earnings to financially support their activities in the world of art, a realm where they gain utility rather than *dis*utility from work [3].

In the creative optimization decision by Bryant and Throsby, the artist instead works *only* in the arts, and chooses a proportion $\alpha_i \in [0,1]$ of a trait T called “creativity” to invest in every individual art piece. Works with $\alpha_i = 0$ are called “commercial”, while works with $\alpha_i = 1$ are called “creative.” Their specifications imply that, given an artist's utility is increasing in both the cultural and economic valuations of her given painting, and conditioning on a creative preference parameter $\lambda_i \in [0,1]$ that weights each of these valuations, some creative effort above 0 will maximize the overall value of the painting and the artist's utility. The artist will thus invest some creative effort in each painting she produces, as long as her utility is increasing in both the cultural and economic valuations of the given paintings.

Given the particularities of artistic earnings (in that a portion

of artists' earnings may come, as some economists have detailed, from sources unrelated to their labor efforts in a given period, such as grants, subsidies, and royalties), it is possible that in the art market, product design is unlinked from a goal of profit maximization. Indeed, given an earnings outlook that implies a small relationship between an artist's choice of product and her earnings (in the context of Bryant and Throsby, an artist expects her choice of α_i to have no bearing on the economic value of the work and on her own earnings, where we may see v_c equal to a constant), an artist may be entirely free to produce exactly what she desires. Thus, free of economic concerns, the artist is free to set $\lambda_i=1$ and thus $\alpha_i^*=1$, such that—unlike for most other industries—financial concerns may not be the main crux behind product design choice in the art market.

To address these considerations, Bryant and Throsby hint at complicating the model by endogenizing the creative preference parameter λ_i : “Although it is beyond our present scope to do so here, we can at least suggest one or two comparative static results based on informal speculations about an artist's preference formation [...] These and other conjectures could be translated into statements about how λ_i changes as a result of feedback effects, and the consequent impacts on creative choices in equilibrium could be worked out.” [4]

In short, an extremely important factor to consider when modeling an artistic product design choice is what actually motivates an artist to enter the art market, and what kind of factors play into her utility function. In particular, attention has to be paid to the weighting of both financial and nonmonetary rewards, and how these rewards are determined by the artist's choice of product [6].

Moreover, attention should certainly be paid to a factor largely absent from the existing literature: the social aspects of the market itself, and the interpersonal relationships between artists. Bryant and Throsby simply mention that in their model, the artist is a member of a non-empty set of artists. Yet, most models like theirs ignore the role of competition and of influence in the art market structure, and the role they play in product design.

THE ARTISTIC PRODUCT SPACE: A LOCATION MODEL

For the development of a location model and an artistic product space, we can think about individual artists as Singer described: bundles of Lancastrian characteristics of ‘decorativeness’ and ‘intellectual appeal’ [7]. Artists are thus locations on a one- or multi-dimensional space defined by a set of artistic characteristics that determine an artist's style. The location model will thus have artists competing and influencing each other in the product space to define their work, producing a general “Picasso product” or “Rembrandt product”, etc.

Further dimensions can be added to the product space by thinking of how much of the overall space each artist occupies (measured by how diverse her style is, or how eminent she is), and by imbuing each artist with a measure of how closely related she is to other artists in the space. We can indeed define this as an artist's “mass”, or influence in product design. While this applies to other markets and industries as well (many firms might look to others for ideas), this is perhaps most true for the art market, where artists are deliberately grouped together into movements based on the similarities of their styles.

In turn, considering the issue of stylistic similarities between artists lends itself to the most critical implication of the location model: minimal versus maximal product differentiation. Sensitive to how models are specified, the optimal strategy for some markets can be either the former or the latter. In the art market case, we can consider artists as being motivated by either choosing a completely new, unoccupied space in the product space far from other artists as something truly new and innovative, or choosing a well-established, crowded location as a way to benefit from association with the artists already located there both at present and in the past, and innovating and expanding from there.

This concept of interplay of artists both across and within generations is at the center of Katherine Giuffre's “Sandpiles of Success.” She writes: “Present-day status is based on a position within a web of ties and also has embedded within it the history of past positions... Career ladders in the art world are not so much ladders as they are sandpiles. The movement of actors within the field changes the shape of the field.” [8] We can thus view artists as forming into dynamic clusters based on proximity to each other on the product space. These clusters can grow, shrink, and move together around the space or separate based on the influence and the cohesion of the artists in the cluster, and the artists in the product space itself. We can further think about the *Zeitgeist* as the cluster of artists that dominates the space, be it by the number of artists in a cluster, space occupied in the product space, total influence held on the product design of other artists, etc. By creating the product space in different time periods, we can track and visualize the development of art movements and careers of individual artists, as a network of artistic connections and associations, and identification with product-defining characteristics. We can see how movements or *Zeitgeists* have been born, died, and evolved, linking a continuum of artistic careers over history to trace the development of artistic styles over time.

CONSTRUCTING THE PRODUCT SPACE

A formal approach for defining a product space with endogenously-defined locations will depart from the utility/ profit function that each artist must maximize. This function in turn takes as basic parameters the location of other artists in the space, the cost of production of an art piece, and the transportation costs “consumers” face from “travelling” to a location other than their respective ideal products’. In the particular case of the art market, we can consider the utility function for an artist who has just entered a market in which other past and currently active artists have already chosen their locations (i.e. defined their styles by occupying a portion of the space). In turn, this new artist would thus consider the monetary and nonmonetary rewards associated with each portion of the product space, such as the expected earnings associated with a particular product or style, and the probability of success given the observed eminence of artists located close to or on that location of the product space. As Singer detailed, “The artist's decision depends not only on his human capital endowments but also on his attitudes towards risk. He has to balance high potential rewards, attainable with relatively low probability, against low potential rewards, attainable with high probability.” [7]

To construct the product space of artists as a new entrant would observe it, we must define a measure of artists'

relationships with each other. For the purposes of this paper, Google ngrams—which report and plot the yearly percentage of all words in a database of more than 5 million digitized books that correspond to a particular search query—were used to find the correlation between the mention of artists’ names in books with the mention of other artists and with the movement with which they are most usually associated. This “culturomic” approach comes with particular advantages and disadvantages, which will be covered later in the paper.

The assumption behind the use of ngrams is that an artist will be mentioned more in books in a given year when artists that are similar to her in style are mentioned more that year as well. Thus, if in a given year the mention of “Claude Monet” in English language books is higher than the previous year, we would expect the mention of “Édouard Manet” to be higher as well. That high or low correlation will be defined as the degree to which these two artists are stylistically related. A product space would then be constructed by plotting the correlations among artists with each other.

Once the frequencies/ mentions of artists are obtained, a correlation matrix is calculated that will show the correlation between each artist on the product space. For the application shown in this paper, 74 artists were selected ranging mainly from Impressionism to Pop Art, though some Renaissance and Baroque painters were included. While the theory of the location model is not sensitive to the particular choice of products to be included in the product space, its usefulness as a model for product design choice relies on the assumption that products in the space will be an accurate representation of the universe of competing products that a firm—in this case, an artist—would observe within its market. Thus, the goal for this particular set of artists was to capture notable individuals from representative movements throughout the history of art that would account for the diversity of artistic styles an artist might conceivably be influenced by. However, the product space is also useful for analyzing product differentiation at more granular levels, such as studying only the artists in a particular medium, period in time, or geographic profile. Data was obtained for mentions in books starting in 1870 and ending in 2008.

The average correlation of each artist with every other artist

was calculated, and a measure called mass for every artist i was devised that was simply:

$$= \frac{1}{(1 - \frac{1}{n} \sum_j \rho_{i,j})} \quad \text{Equation 1}$$

or the inverse of 1 minus the average correlation of an artist with all other artists included in the product space.

Finally, artists’ correlations with the different artists on each axis are plotted to define the product space. For the purposes of this paper, the size of each marker on the plot was designed to reflect the average percent of mentions in each year for every artist. This measure is thus independent of other artists’ mentions and implies that more popular artists that are mentioned more often in books will have a larger presence on the product space. The filling of each marker is in turn a reflection of each artist’s mass. While ideally the axes on the product space would be more representative of product characteristics (like level of abstraction, or commercial inclination), for practical purposes the axes were defined to be the correlation with the specific artist (and in turn, that artist’s style) on the axis.

This visualization allows us to see Guffrie’s “sandpiles” in practice: every new artist in the space changes the mass for everyone. The artists in the product space inherently define it, and no artist can be located on the space without a relation to another artist.

Because of this relationship, we can define then the artist’s “mass” as his gravitational pull and his influence in product design. Of special importance is the distinction that this is not necessarily correlated with influence in terms of achieved eminence, reputation, or renown. Instead, artists with a high “mass”—because we assumed from the start that correlations imply stylistic connections—are those artists that are overall more closely related to the styles and products of the other artists in the product space. Thus, we presume that someone with a very high mass is one who has had a major influence in how artists design their products.

Thus, it makes sense that popular artists who are monumental figures in the pantheon of art history—such as Raphael, Michelangelo, Rembrandt—have a low “mass” if they are uncorrelated with the type of work that the other artists in

MOST INFLUENTIAL			LEAST INFLUENTIAL		
Ranking	Name	Mass	Ranking	Name	Mass
1	Edgar Degas	3.9986	65	Peter Paul Rubens	1.8518
2	Pablo Picasso	3.9673	66	Diego Velázquez	1.5610
3	Wassily Kandinsky	3.9671	67	Michelangelo	1.5158
4	Joan Miró	3.9049	68	El Greco	1.3454
5	Piet Mondrian	3.8848	69	Albrecht Dürer	1.1825
6	Fernand Leger	3.8752	70	Anthony van Dyck	1.0092
7	Édouard Manet	3.7874	71	Leonardo da Vinci	0.9473
8	Henri Matisse	3.7797	72	Rembrandt van Rijn	0.8625
9	Alfred Sisley	3.7434	73	Jan van Eyck	0.8216
10	Marcel Duchamp	3.6815	74	Sandro Botticelli	0.6676

Table 1. Rankings of Product Design Influence

	Dependent Variable: Mass	Dependent Variable: Average
lifespan	0.0017*** (1.56×10^{-5})	-0.0007*** (7.85×10^{-7})
separation	-0.0038*** (5.48×10^{-7})	-0.0013*** (7.14×10^{-8})
constant	3.2475*** (0.0934)	0.8053*** (0.0055)
Adjusted R²	0.9536	0.9281

Table 2. Regression Results

the product space are producing. Especially because these artists are far separated from us in terms of time, it is unlikely that many artists are currently painting in the styles of Michelangelo or Rembrandt. On the other hand, artists are likely to appear more closely related and produce works more stylistically comparable to someone like the 20th century abstract painter Wassily Kandinsky, who occupies position no. 3 in the list of artists with the highest mass in the product space of 74 artists constructed in this paper. He follows only Edgar Degas and, unsurprisingly, Pablo Picasso. These three artists are joined by the likes of Piet Mondrian, Édouard Manet, and Henri Matisse in the top 10 most influential artists in product design of the 74 included in this paper. Table 1 presents the top 10 most influential artists in this product space, and 10 *least* influential artists, which as ventured above, are particularly influential figures in the history of art, but given their temporal separation from modern-day artists, are unlikely to be directly influencing the latter in terms of product design.

Note that for Da Vinci, Rembrandt, van Eyck, and Botticelli, their mass below 1 implies their average correlation with the other artists in the product space was actually *negative*, implying that a higher interest in the other artists in the space (as measured by the mention in books) was correlated with a decrease in interest in these artists. Once again, the assumption is that this reveals information on the stylistic connection between these artists and their products.

Giving the artists in the “least influential” column the benefit of the doubt, perhaps our bias toward including more recent artists in the product space is affecting these results, and had more early artists been included (specially more in the period after the Renaissance and before Impressionism), their average correlations would have been higher. This reiterates the sensitivity of this approach to the actual artists included in the product space, since by definition these average correlations and the resulting visualization of the product space depend on the artists included. From the 74 artists included, 26 died more than 100 years ago (of which 13 died more than 300 years ago), 44 died within the last 100 years, and four (Wayne Thiebaud, Jasper Johns, Bridget Riley, and Frank Stella) are still alive.

As is evident then, this approach may have serious implications both in terms of how a new artist chooses her location in this universe of artists (with consideration as to the success she may gain by choosing the same location as Picasso and his cluster of

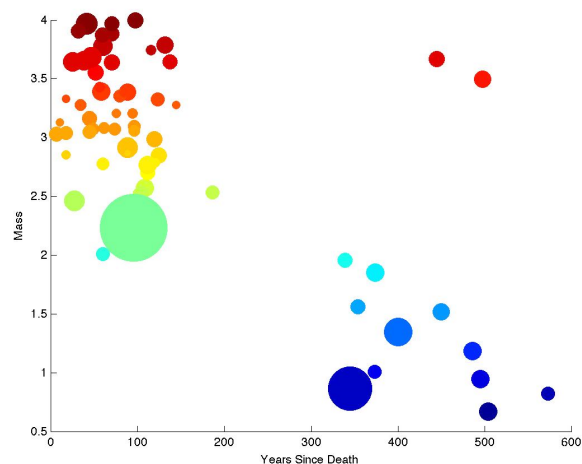


Figure 1. Separation versus Mass

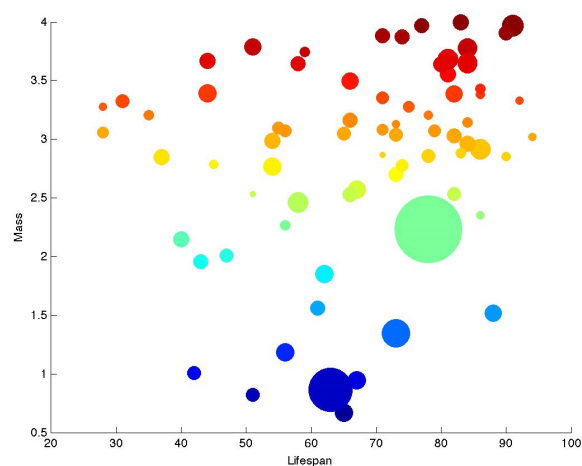


Figure 2. Lifespan versus Mass

artists versus choosing the one for Andy Warhol, for example, or an entirely unoccupied location), and a very strong implication for how we measure the influence of artists on other artists when it comes to product design. In particular, eminence in art history does not necessarily translate to eminence and influence in art production. Two potential hypotheses arise to describe the observed phenomena.

The first is related to the years since an artist’s death, and seems to provide an explanation for why some of the biggest names in art are considered here as the least influential in terms of product design. That is, the longer since an artist has died, the less of an impact she will have on present and future artists’ styles and product design decisions.

A simple OLS regression of mass and of average correlation on a constant, lifespan, and separation (defined as years since the death of the artist) seems to confirm this hypothesis: the more separation in time from an artist, the less of a mass she has, and therefore the less influential she is in determining today’s product design. Figure 1 displays this relationship. Note that the two outliers in top-right corner of the plot are Pieter Bruegel and Hieronymus Bosch, whose styles are rather similar and, in the eyes of a modern viewer, potentially familiar to the works we might expect from more contemporary artists (explaining, perhaps, their high mass).

The second hypothesis is the intuition that perhaps longer-

lived artists (like Joan Miró, or Pablo Picasso) will have higher mass on the assumption that because they lived longer, they had the chance to explore different styles of art and thus nurtured the kinds of stylistic and historical associations that will produce very high correlations with the other artists in the product space. For this particular product space (which requires us assuming that we actually could have Picasso, Rembrandt, and Michelangelo, all together in one product space), this potential hypothesis does not seem supported at first glance. Looking at a plot of lifespan on mass (See Figure 2), we cannot observe much of a significant relationship. Note that the large green marker corresponds to Pierre-Auguste Renoir.

Moreover, while the first regression in Table 2 seems to find a small, but positive and significant effect of an artist's lifespan (assuming it correlates closely with length of career) on her mass, this may simply result from our transformation of average correlation into the measure $mass = \frac{1}{1 - average}$, as the second regression offers up a significant, small but *negative* relationship between an artist's lifespan and her average correlation with all other artists. Given that *mass* approaches infinity as *average* approaches 1, it's likely that the effect between lifespan and an artist's influence is actually slightly negative, as in the second regression.

Another point of interest is the correlation of artists with general art movements. In particular, if we apply the same concept of mass to these movements, then we can rank art movements themselves in terms of their influence on product design. Using this approach, it appears German Expressionism is the most influential art movement in terms of product design in this paper's product space, followed by Surrealism and Abstract Expressionism.

Figures 3 and 4 finally present a sample product space, with Pablo Picasso and Vincent Van Gogh as the defining artists for Figure 3, and Picasso, Van Gogh, and Claude Monet for Figure 4. Since in this visualization of the product space the axes represent each individual artist's correlation with the artist on that axis, each point on Figure 3 can be interpreted as the unique location of an artist relative to how similar he is in style to both Pablo Picasso (the x-axis) and Vincent Van Gogh (the y-axis), where similarity is defined by the correlation between the artist's mentions in books and the mentions of Picasso and Van Gogh, respectively:

CRITIQUING AND EVALUATING THE MODEL

A succinct evaluation of the model will identify that—with some exceptions—the correlation between movements and their corresponding artists tends to be very high. Moreover, at the artist level, we observe high correlations between artists that we know are closely related in style and movement. A particularly salient example is that of Jasper Johns, Robert Rauschenberg, and Roy Lichtenstein, where the correlation between Johns and Rauschenberg is 0.9933, and between Johns and Lichtenstein is 0.9792. Another salient example is that of Claude Monet and Édouard Manet, whose ngram in Figure 5 (plotted as well with Alfred Sisley) demonstrates how closely together the mentions of artists of related styles can vary.

Nevertheless, however striking some of these plots may appear, the use of ngrams carries with it certain issues, which ideally would be corrected by the use of more sophisticated approaches to measure the correlation between artists.

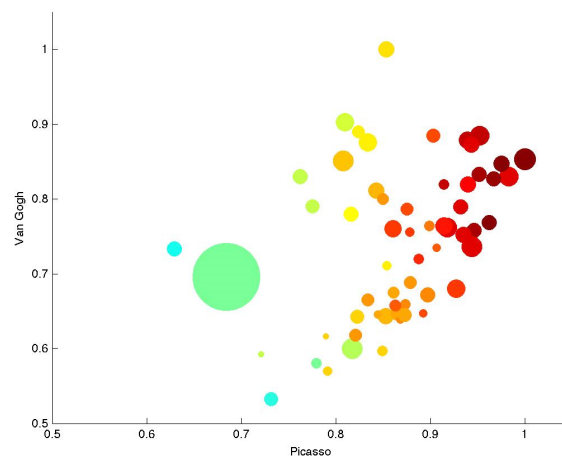


Figure 3. Product Space (Two-Dimensional)

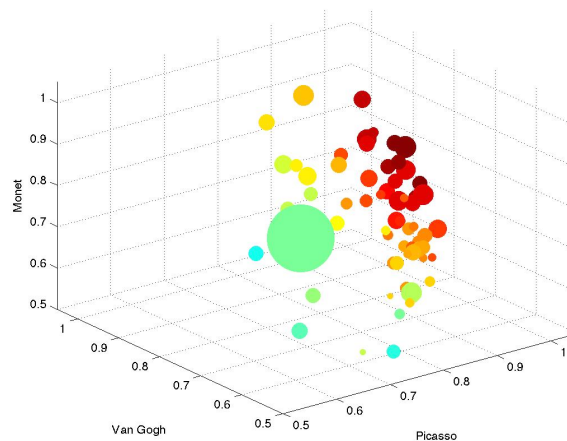


Figure 4. Product Space (Three-Dimensional)

The first category of issues relates to the specifications of the search itself, in that one must first decide on both the specific search query to use for each artist, and the sample of books to search through. The former choice becomes particularly important for hyphenated names, aliases, mononyms, alternative spellings, and a diverse range of other special circumstances where our selected search query could result in an inaccurate representation of an artist's true mentions in books. While Google remedies some of these concerns by allowing both case-insensitive search queries and the grouping together of alternative spellings, care must nevertheless be given to the selection of search queries. For example, the Renaissance painter Raphael was excluded from the sample product space in this paper, as it is unclear that any of the search terms "Raphael", "Raphael Sanzio da Urbino", "Raphael Urbino", or any other alternative version of his name would result in an accurate reading of his mentions in books, given the rare use of the latter two phrases to refer to the artist, and the potential for many false positives when searching for the former. Regarding the selection of the sample of books to conduct the search on, at the moment Google only allows searches on a single language at a time. For more regional artists or those who would appear more frequently in particular languages, there is a risk of underrepresenting their mentions in books.

The second category of issues is when to begin counting the mentions of artists. For example, the different time intervals

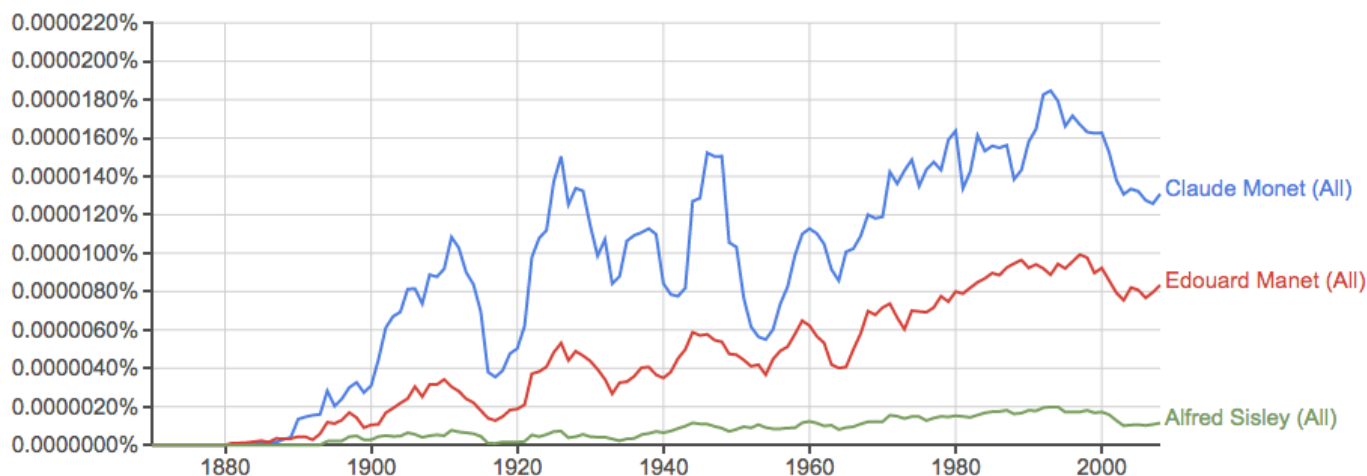


Figure 5. ngram: Claude Monet, Édouard Manet, Alfred Sisley

in which individual artists began being mentioned in books (for example, many artists' mentions in books is 0 in 1890, when Monet was being mentioned already) may of course bias the correlations. Yet, for our purposes this is acceptable, since it naturally separates out members of art movements by lowering the correlations of artists who were born or achieved eminence much later than the artists of earlier movements. However, if an artist revives an old style, then perhaps this might be an issue. Then, the correlation would be lower than it should be, since the assumption for this model is that the correlation will be able to capture the degree of similarity in styles between artists.

The third category of issues relates to sample selection, as lesser-known artists (who are the large bulk of the market) may not be eminent enough to be mentioned frequently in books and, as such, may not even show up in the product space at all. Moreover, when they do, it may reveal more about an artist's capacity to be interesting enough to be mentioned in a book than her stylistic similarity to other artists.

Final considerations are the possible critique that this approach will only reinforce art movements as they are currently defined. Artists will thus be close to other artists in the product space precisely because they've already been historically associated to the artists they're close to. However, the approach is not particularly aimed at redefining art movements, but at empirically showing how they may appear on a product space and influence a new artist's choice of location, since the assumption is that art movements have been historically defined precisely because of the stylistic similarity of the artists included in them.

Lastly, omitted variables might be driving correlations between artists, as opposed to stylistic similarities. For example, popular artists may be more highly correlated with the state of the art market than less popular artists. We could thus imagine three artists X, Y, and Z existing in the product space, such that X and Y are highly similar in style and Z is dissimilar. However, both X and Z are very popular artists, while Y has not yet achieved a comparable level of eminence. Because X and Z are popular artists, regardless of their dissimilarities in style they move closely with the cycles of the art market and with the interest in mentioning popular artists in books, while artist Y is less affected by this omitted variable (the art "business cycle"). In turn, the correlation resulting from mentions of these artists in books will incorrectly attribute a high degree of stylistic similarity to artists X and Z, and a low degree of similarity between X and Y.

Moreover, we can imagine a scenario where a popular artist will have a high correlation with the other artists in the product space not because of her stylistic similarity with them, but because of the general ability of popular artists to stoke increased interest in and attention to the arts and, in turn, to lesser-known artists. As such, correlation with other artists would actually be capturing an artist's "spillover" onto other artists, i.e. her ability to incentivize an increased interest in the arts, and not stylistic similarities. As such, an explanatory model for changes in the frequency of an artist's mentions in books that includes both systematic factors (i.e. total art sales, or less directly though correlated with the art market, GDP and GDP growth) and the mention of other artists might be useful.

CONCLUSION

While economists have looked at topics like the amount of creativity an artist may choose to use on the production of a given painting, and the relationship between the economic and the cultural considerations of an artist with regard to her work, little has been written about how artists may choose to design their style and their brand based on the probability of success—both culturally and financially—associated with that style, the definition of "innovation" in the art market and how it takes place, and the relationships between artists and how they contribute to each other's products, or artworks. My attempt was to visualize a product space of artists and their styles as a framework that would facilitate thinking about each of these issues.

While these first steps in attempting to visualize an artistic product space suffer from the potential setbacks laid out above, this approach may nevertheless be useful as a framework for thinking about product design and differentiation in the art market, and as an opportunity to test other hypotheses on creativity, achieved eminence, artistic production, and the relationships among artists. For example, given Giuffrè's hypothesis on interpersonal relationships and success, we could regress an artist's achieved eminence on the correlation with other artists (in other words, how much of an artist's success can we attribute to her being a unique artist versus being part of a very crowded portion of the product space?). We could also think about how this correlation with other artists affects an artist's career earnings, and thus draw some potential conclusions on the issue of maximal versus minimal product differentiation in the art market. This

could facilitate our understanding of the factors driving artistic production, and in turn advance the body of knowledge around a global market of both great economic and cultural value.

REFERENCES

1. Alper N, Wassall G. Artists' Careers and their Labor Markets. In: Ginsburg VA, Throsby D, eds. Handbook of the Economics of Art and Culture, Volume 1. Amsterdam: Elsevier, 2006:813-864.
2. Menger P. Artistic Labor Markets and Careers. *Annu Rev Sociology* 1999; 25:541-574.
3. Schneider L, Steiner L. The happy artist: an empirical application of the work-preference model. *J Cultural Economics* 2013; 37:225-246.
4. Bryant W, Throsby D. Creativity and the Behavior of Artists. In: Ginsburg VA, Throsby D, eds. Handbook of the Economics of Art and Culture, Volume 1. Amsterdam: Elsevier, 2006:507-529.
5. Throsby D. A Work-Preference Model of Artist Behaviour. In: Peacock AT, Rizzo I, eds. Cultural Economics and Cultural Policies. Dordrecht: Springer Netherlands, 1994:69-80.
6. Frey B. Not Just for the Money: An Economic Theory of Personal Motivation. Cheltenham: Edward Elgar, 1997.
7. Singer L. Supply Decisions of Professional Artists. *Am Economic Rev* 1981; 71(2):341-346.
8. Giuffre K. Sandpiles of Opportunity: Success in the Art World. *Social Forces* 1999; 77(3):815-832.



RICARDO R. RODRÍGUEZ-PADILLA

Ricardo Rodriguez-Padilla is a recent graduate of Columbia University, where he studied Financial Economics and Business Management. Originally from San Juan, Puerto Rico, Ricardo currently resides in New York City while working for an economic consulting firm and preparing to apply to graduate programs in Economics. He is interested in exploring topics in Industrial Organization as they relate to more idiosyncratic markets, like arts and culture. He hopes to pursue a career in economic research after graduate school.

New Media Technology and Social Movement Framing in the WTO Protest: A Case Study in Peaceful Majoritarian Possibilities for 21st Century Activism

Michael Lachanski

Princeton University

We hypothesize that mass decentralized media (e.g. camcorders, smartphone video capabilities, and ubiquitous surveillance) has increased the benefits to nonviolent protest, an option only available to activists, while increasing both activists' and authorities' accountability for violent actions. Majoritarian activists, whose policy preferences are favored by a majority of media consumers, are uniquely poised to benefit from these changes in the communications landscape. Taking the 1999 WTO protests in Seattle as a case study, we investigate how different anarchist groups utilized the aforementioned power asymmetry inherent in the new media technology to establish favorable or unfavorable narratives in mainstream media. Because both nonviolent obstruction and property destruction were utilized by different anarchist factions in Seattle, we can use the WTO protests to analyze how each type of protest was covered in the media and received by the public. As hypothesized, close study suggests that both mass media coverage of and public response to violent tactics on the part of anarchists and authorities was uniformly negative. In contrast, the rest of the anti-WTO movement, benefitting from police brutality and the nonviolent obstructionist tactics, received sympathetic mass media coverage from across the political spectrum and majority support for its policy proposals post-protest.

I. INTRODUCTION

One month before the end of the last millennium, a disparate group of 50,000 individuals in organizations ranging from the AFL-CIO labor union to the “Black Bloc” Eugene, Oregon-based anarchist movement took to the streets to protest the World Trade Organization (WTO) ministerial conference in Seattle. As recounted in the book *The Battle of Seattle*, disparate groups with ideologies ranging from anarchism to environmentalism were united by an opposition to the WTO, which they believed to be emblematic of – in their words – “global capitalism”, “deregulation”, “environmental degradation”, and “unaccountability”. Although initially peaceful, the situation began to veer out of control as the anarchist breakaway group Black Bloc smashed windows, set fires, and battled with police throughout the failed conference. [1] Nonetheless, in spite of American aversion to anarchist ideologies and a hostile mainstream media (MSM hereafter), polls showed that a majority of Americans sympathized with the protesters. [2]

For the first time, activists used their websites to publicize images and video of police brutality against protesters as well as to reflect the peaceful nature of the majority of protesters themselves. The Seattle protest illuminates how majoritarian movements that relying exclusively on nonviolent resistance are uniquely positioned to utilize new media technologies, i.e. ubiquitous video surveillance and social media, to receive positive portrayals from the traditional media and consequently increase

their standing in the public sphere. On the other hand, new media technologies not only decrease the direct effectiveness of violent factions, but also discredit the causes that violent factions support. Early 20th century peace activist Richard Gregg provided a psychological and scientific explanation for why nonviolent resistance against violent attackers is the most effective way to challenge the authority of the attackers. [3] By taking the effects of new media technology into account, we can extend Gregg's theory to help explain why the American people supported the anti-WTO movement. Even people unaffiliated with the movement and unlikely to visit activist websites were impacted by sympathetic imagery in MSM sources like Time magazine and the evening news. In part this was because hundreds of anti-WTO activists brought digital camcorders and provided their footage to the mainstream media. [4] While some elements of the traditional media were quick to conflate the entire anti-WTO movement with the violent Black Bloc minority, activists armed with camcorders captured the fundamentally nonviolent nature of the protests, clarified the events of the WTO protest, and won the sympathy of the American people.

As the 21st century media landscape undergoes radical changes, there has been a spate of interest by academics in how new media technology like Twitter, blogs, and Facebook decentralize media control away from traditional media authorities, and transfer it to activists. [5] While this technical revolution in new media has been well-documented, and the positive effects of new media on activist organizing have been well-established,

there have been few case studies on how the interaction of new media technology and traditional media outlets utilized by the MSM has changed the effectiveness of different tactics. Globally, traditional media outlets like television and newspapers are still the dominant news medium, making an understanding of how activists can use these mediums to spread their narratives. [6] By analyzing which tactics are most effective when used in tandem with new media technology, we can predict which types of social movement activism are most likely to achieve success. For this first pass analysis, we partition the menu of activist tactics into just two types: non-violent resistance and violent resistance. Non-violent resistance does not result in the direct destruction of physical property or life, while all other activities are grouped into violent resistance. We will see that this crude distinction is sufficient to understand how new media technologies generate differential MSM responses to different forms of activism. The simple observations we make in our case study concerning media-activist interaction can nonetheless provide helpful guidance for directing the limited resources of social movements.

II. GREGG AND NONVIOLENT ACTIVISM

We briefly review the peace studies literature as it pertains to social movement tactics and analyses of media-social movement interaction. While nonviolent members of the WTO protest movement cited Martin Luther King Jr. as an inspiration for their tactics, much of the theory behind King's nonviolent civil disobedience came from Richard Gregg's book, *The Power of Nonviolence*. [7] In his book Gregg explained why nonviolent tactics are the most effective way to win the attention of potentially sympathetic non-activist spectators. Gregg argued that protesters who use nonviolent tactics, especially in the face of violence, will win the empathy of spectators because spectators will identify with the bravery of the protesters. Mass media plays a paramount role in Gregg's theory because it is needed to create "a global audience" that will raise "the threat of bad publicity" for violent opponents. [3] Gregg's theory posits that nonviolent activism can help create a particular frame or interpretation of the nonviolent protesters as heroes and violent authorities as villains. Gamson and Wolfsfeld, in their seminal work on media technology and social movements, suggest that if activists want the media to create pro-movement frames, they should devote more resources to interacting with the media and appoint specific representatives to the media. [8] New media technology not only augments these options for nonviolent social movements, but creates new ways for activists to win the sympathy of journalists and even bypass the traditional media altogether. We find that both forces help explain the success of the anti-WTO movement.

III. HISTORICAL REVIEW

Contrary to their MSM portrayal, the protests in Seattle were not a spontaneous uprising against the WTO, but were planned by activists in the Direct Action Network (DAN) several months in advance. [9] While each DAN participant had different demands (ranging from stricter environmental and labor regulations, favored by participating labor unions, to the abolition of the WTO, favored by anarchists), DAN's immediate goal was to end the secrecy of the WTO negotiations by winning a seat at the conference for disenfranchised stakeholders like unions

and environmental groups. [17] Members of DAN trained and coordinated labor unions, environmental groups, anarchists, and a hodgepodge of other anti-WTO groups to protest nonviolently by forming human chains and blocking key intersections that WTO delegates needed to get through to travel from their hotels to the conference center. [12] These nonviolent actions ensured that the WTO would be shut down unless the local authorities resorted to violence. If the authorities used mace and tear gas, the camcorder-armed protesters would (and in fact, did) videotape the violence perpetrated against them with the intent of broadcasting it. [12] These tactics guaranteed that either the WTO conference would fail because the delegates could not get to the conference building or that the activists would win the conscious sympathy of the American people. George Lakey, an anti-WTO media representative, echoes Gregg when he explains the media-generated benefits of non-violence:

If they allow us to go ahead and do what we intend to do, we accomplish something worthwhile and related to our issue. If they repress us they put themselves in a bad light...If the protesters are stopped violently, the public is educated and new allies can be won...Drama does what nothing else can do: it arouses the attention of otherwise occupied parts of the citizenry, it educates them on a gut level, it motivates them to find ways of acting that make sense in their terms, and it even attracts many of them into the movement itself. [7]

Lakey makes it clear that both nonviolent tactics and media presence are necessary for the anti-WTO movement to succeed. If a nonviolent protest is interrupted by police brutality and no one records it, it does not make an impact. Within the framework of Gamson and Wolfsfeld, activists in movements dedicated to nonviolence still face the challenge of attracting media attention that will frame the movements' actions appropriately.

IV. DISCUSSION

The activists leading the anti-WTO movement were consciously aware of the power that nonviolent tactics had when combined with guaranteed media coverage. For instance, during the WTO Protests, the Independent Media Center, an anti-WTO outfit provided over 450 activists with cameras. [12] One activist even went so far as to say that, "EVERYONE seemed to have a cell phone or was taping the events with their video camera." [13] The level of conscious media control granted by new technology strengthens Gregg's arguments in four ways: (1) nonviolent protests have access to more powerful dynamic visual mediums (videos) rather than simply photos and texts as in Gregg's time, (2) new media technology allows activists to create media favorable to their cause, (3) new media technology allows activists to circumvent the MSM and transmit their messages directly to the public, and (4) new media technology, by creating new ties between journalists and activists, provides another channel through which nonviolent movements attract favorable media coverage..

First, modern visual media creates new opportunities for social movements to create empathy in the average viewer by forcing viewers to adopt the perspective of a member of that social group. Our extension of Gregg's argument is straightforward:

if static photos and text arouse mass support for nonviolent resistance (especially relative to violent resistance), then surely dynamic mediums like video and audio will be even more effective at doing so. This technique is most vividly demonstrated during the numerous scenes of police brutality in the iconic film *This is What Democracy Looks Like*, an anti-globalization film released by Indy Media. [14] The film synthesizes the footage of hundreds of different activists into a pastiche that shows how the otherwise unprovoked Seattle police beat and maced the human chain of protesters so that the WTO delegates could get to the convention center. Viewers almost cannot help but empathize with the bruised and teary-eyed protesters. By creating their own alternative media like DVDs, audios, and websites, nonviolent social movement activists helped create a pro-movement frame.

The second way that new media technology empowers social movements is by changing the balance of power between journalists and protesters. Although social movements once depended on the traditional media to provide coverage of their protests, new media technology allowed activists to videotape the exploits of their social movement. [8] By selectively giving their recordings to the traditional media, activists can influence how the media will frame their movement. Most Americans who felt sympathy for beaten and bloodied WTO protesters were not exposed to these images through anarchist or leftist web pages or magazines, but through the pages of the moderate *Time* magazine and *USA Today*. [15] This is because the scenes in *This Is What Democracy Looks Like* were available to the public and the traditional media before they were compiled into a film. [16] Leftist traditional media outlets were even more inclined to spread pro-movement frames. *Mother Jones* magazine, a leftist traditional media publication, was quick to post protester-provided videos of the police brutality onto its website. [16] The frame of peaceful activists versus an overzealous police force is a trope in American politics, and members of the anti-WTO movement were glad to utilize the MSM to spread this frame.

While both of these complementary strategies used new media technology, they are plausibly the result of what Gamson and Wolfsfeld described as a willingness to devote more resources (i.e. camcorders and activist-journalists) to a social movement's media coverage. However, we can identify new framing techniques that uniquely resulted from the social movement's monopoly on new media technology. [8] If the traditional media fails to respond to the social movement's media, then that social movement will dominate the framing of the issue via its monopoly on new media technology. Even worse for traditional media, if it fails to cover the movement in sufficient depth, they run the risk of being bypassed by activist controlled media, e.g. *Indymedia.org*, which received far more visitors during the WTO Protest than the MSM's *cnn.com*. [17] For example, the unsympathetic libertarian-right *National Review* (favorably quoting the *Washington Post*) questioned why the traditional media did not give live coverage to cover the "one of the largest acts of civil disobedience in history." [18] At the

same time, it is not uncommon for activist media to become the story through traditional media reviews of activist media designed for mass consumption, like the film *Battle in Seattle*. [19] Not only did the film itself invoke the typecast imagery of nonviolent resistance, but its release forced the traditional media to cover the events of the protest in reviews of the film. In some cases, the traditional media even covered activist media as an example of the growing influence of new media technology on the movement. [16] Activists' control over new media technology makes them extremely useful to less technology-savvy journalists who may depend on them for content in the new mediums. Because they are dependent on the activists' new media technology for their journalism, traditional media journalists become less likely to dismiss nonviolent protests.

We believe that the most important impact of the new media technology is to turn activists into citizen-journalists themselves. This facilitates the creation of social networks between traditional media journalists and activists. Activists

who possess control over new communication mediums makes friends with and collaborators out of journalists. The shared experiences of traditional journalists and new media activists (like police beatings and pepper sprayings) further contribute to the propagation of pro-social movement frames in the traditional media. [20] Nonviolent resistance is essential to establishing rapport between journalists and social movement members. MSM journalists who were not activists were struck by

the fact that the vast majority of activists, contrary to the journalists' expectations, were peaceful. [21][22] Given the massive positive effects generated by establishing a social movement-dominated new media presence, it should be no surprise that in post-Seattle anti-WTO protests police forces first destroyed the video cameras of activists before they committed acts of brutality. [23] While police and paramilitary authorities have always had the option to stop social movements via violence, the presence of video cameras raises the cost of doing so. As camera wielding protesters and traditional journalists spend more time in close proximity with each other, police attacks on members of social movements become more likely to injure journalists in the traditional media. It is difficult to imagine these journalists remaining unsupportive of the protesters' cause if they are injured by police.

We believe none of these channels applies to violent protestors. If media effects increase mass sympathy for nonviolent protestors and decrease mass sympathy for violent protestors then more powerful media effects should simply amplify those effects. Neither favorable imagery nor the direct transmission of violent activities by protestors to the public is likely to alter this calculus. MSM journalists are unlikely to fraternize in large numbers with violent protestors.

Nevertheless, several media theorists, most famously Slavoj Žižek, argue against Gregg: these thinkers believe that violence can be an effective tactic to win the support of the media consumers. Anti-WTO forces destroyed property to create a spectacle

Majoritarian movements relying exclusively on nonviolent resistance are uniquely positioned to utilize new media technologies, i.e. ubiquitous video surveillance and social media, to receive positive portrayals from the traditional media and consequently increase their standing in the public sphere.

that would naturally attract media attention to their cause. [8] Dressed in dark clothing and ski masks, several hundred Eugene, Oregon Anarchists moving in a Black Bloc formation destroyed the windows and vandalized Nike, Starbucks, and hundreds of other stores. [25] New media technology provided violent anarchists with some of the same framing options available to the nonviolent wing of the anti-WTO movement. For example, Black Bloc media activists provided 60 Minutes reporters with reels of footage from their activities. [25] The speed of the Black Bloc strike force combined with the televised media's failure to establish any presence on-the-ground ensured that only members of the movement themselves would be able to capture their exploits in any detail. Thus, they were able to define themselves and selectively share their exploits with any members of the media who were willing to cooperate. This self-framing shows the power of digital media to multiply the message of the any movement. But did the frames propagated by militants resonate with the viewing public?

Anarcho-primitivist scholar John Zerzan argues that violent protests are the only kind of protests that can even attract attention, let alone win the sympathy of the American people. [24] He writes that, "Decades of non-militant, write-your-congressperson advocacy has led only to an accelerating assault on nature," and cites, "Breaking business windows and skirmishing with police," as examples of successful protests. [26] Time magazine grudgingly seemed to concur, stating that "From now on, every objection [to the WTO] will be illuminated by the fires of last week." [27] Anarchist filmmaker Tim Lewis interviewed several anarchists who revealed their plan was to destroy property so that can "break the spell" that consumer goods have placed on people. [26] This strategy's effectiveness appears limited only to the direct victims of property destruction. New media technology enabled these violent spectacles to be broadcast to millions of Americans on national television.

The evidence that violent protests repulsed Americans is so overwhelming that even protesters supporting the use of violent spectacles admit that, "Since polls show strong support from Americans on most antiglobalization issues, it seems likely that the relatively small numbers at US protests since Seattle can be partly explained by the alienation of some potential supporters due to fears of violence." [28] The presence of new media technologies ensured that violent tactics would backfire because thousands of Americans were frightened away from the anti-WTO movement for every few victims of the anarchist movement.

Gregg, a victim of violent protests himself, highlighted more subtle weaknesses of violent tactics, noting that, "One reason many strikes fail is because both employers and employees...are ensnared in the same net of ideas and valuations, those of money and violence." [3] He further noted that, "Combatants [who are nominally opposed to each other] implicitly consent to a common set of moral values, despite their apparent opposition." [3] In other words, even if they attract media attention violent combatants cede the moral high ground they need to win widespread public support. Since new media technology reliably increases traditional media coverage, Gregg's theory implies that movements should rely exclusively on nonviolent tactics. If Gregg's critiques of violent tactics are correct, then the anti-WTO protesters won the support of the American people despite the use of violent tactics, not because of them.

Ample evidence supports Gregg's critiques. Most newspaper

coverage was quick to draw parallels between the actions of the Seattle police force and the anarchist Black Bloc. The equivocation of the two groups by the media in no way reflected the proportion or quality of the violence at the WTO. State forces numbered in the hundreds and used tear gas, pepper spray and physical violence indiscriminately. The Black Bloc numbered in the tens (likely less than thirty individuals) and smashed several storefront windows. [29]

Even armed with new media technology, the violent wing of the anti-WTO movement could not create frames that won widespread acceptance by the American media. In contrast to the examples of reporters writing sympathetic portrayals of the peaceful protesters above, we were unable to find a single sympathetic portrayal of the Black Bloc in the traditional media. When the Black Bloc mimicked the strategy of the nonviolent protesters and provided footage from their exploits to 60 Minutes, 60 Minutes repudiated their actions. [26] Nonviolent anarchist factions, aware of the chilling effect that images of burning dumpsters and slashed tires had on the wider public, publicly shunned the Black Bloc. Anarchist media representative David Solnit writes that, "Eugene activists [the Black Bloc]... allowed 60 Minutes and other corporate media outlets to do extensive features on them as 'Eugene Anarchists,' as one way to help the public dismiss what had happened in Seattle." [2] While this statement probably mischaracterizes the Eugene Anarchists' intentions, it reflects a widely-held perception of the Black Bloc's impact on the larger anti-WTO movement. New media technology marginalized violent activism by visually exposing their tactics (intended for a limited, local audience) to the wider public who cannot distinguish between the violent tactics of police authorities and the violent spectacles of social movements.

To Gregg's criticisms and the empirical evidence against violent tactics, a third critique can be drawn from the work of Gamson and Wolfsfeld: violence poisons the well by allowing some media entities to frame the entire movement in terms of the most egregious of the violent spectacles. [8] Anti-WTO journalist and organizer Eddie Yuen summarizes Lakey's arguments against violent tactics, stating that, "More militant activities always trump less militant ones...potentially endangering the nonviolent majority without their consent...Corporate media can always be counted on to emphasize 'violent' actions, however minor, to the exclusion of all other forms of protest." [28] While previously cited news coverage of the WTO protests exposes this summary as an exaggeration it certainly reflects the coverage of many media outlets that framed the entire anti-WTO movement as violent. For example, both the right-leaning, authoritarian New York Post and the liberal-leaning New York Times failed to distinguish between violent and non-violent opposition to the WTO during and even after the protest. [29] Because one violent faction's actions allow the media to portray the entire movement as violent, violent tactics impose a huge cost to any movement that even tolerates them. Juxtaposing the terrifying spectacles of one violent faction with the entire anti-WTO movement gives the authorities permanent leverage over movement activists. David Solnit reports that:

Police agencies produced a video...[that displayed] images of activists at the WTO in Seattle breaking corporate chain store windows...[for] the Los Angeles City Council just before a vote on funding a massive

police presence and new riot gear to counter the demonstrations. The Council was scared, the funding measure passed. [2]

In the age of mutual mass surveillance, existing MSM power structures will reliably utilize violent spectacles to characterize entire movements by their most violent elements; they will use violent spectacles to mobilize violent countermeasures. Because activists are in a position of weakness with regards to the use of force, they are unlikely to advance their causes or attract public approval by widely broadcasting physical assaults or property destruction.

CONCLUSIONS

The development of mass decentralized media like camcorders, smartphones, and the internet has only strengthened Gregg's critique of violent spectacles in the United States. Not only does violent spectacle attract negative attention from the media, but they may trigger revulsion in an uninformed but otherwise movement-sympathetic viewer. New media technologies, emphasizing the visual aspects of protests, give authorities and the traditional media the option to discredit entire social movements by highlighting violent actions by a minority of protesters. As documented in this case, they may scare the public into granting more powers to the authorities opposing the social movement. Just as new media technologies weaken the case for violence, they vindicate Gregg's support for the exclusive use of nonviolent resistance. Nonviolent resistance inspires camaraderie with the viewing public and wins the support of the traditional media by creating opportunities for journalists and activists to share experiences of authoritarian violence. In future work, we are interested in exploring how these aspects of Gregg's arguments bear on the Arab Spring, in which social media propagation of images of violence played a critical role.[30]

Activists of the future must take Gregg's advice to heart. Nonviolent tactics are the only tactics that can arouse the sympathy of the global citizen by distinguishing the activists who practice such tactics from the authorities who rely on violent coercion. The combination of media coverage and nonviolent activism are necessary ingredients for success. If there is no media coverage and the protesters are nonviolent, then the authorities can simply use force to end the protest. If there is media coverage and the protests are violent, then the general public will not only fail to sympathize with the social movement, but may even view the authorities as a positive stabilizing force. Movement discipline to ensure that activists only use nonviolent tactics will prove many future social movements' chief challenge. While the anti-WTO social movement may mark Seattle as an example of the successful use of new media technology and mostly nonviolent tactics to win over the American people, one wonders how much more successful the anti-WTO movement could have been without the violent property destruction perpetrated by the Black Bloc.

REFERENCES

1. D. Postman, "Group Rejects Others' Pleas Of 'No Violence' Black-Clad Anarchists Target Cars, Windows," *Seattle Times*, 1999 [Online]. Available: <http://community.seattletimes.nwsource.com/archive/?date=19991201&slug=2998734>. [Accessed: 01– Dec- 2015]
2. D. Solnit, "The Battle of the Story of the Battle in Seattle," in *The Battle of the Story of the Battle in Seattle*, edited, 5-55. Baltimore: AK Press, 2009.
3. J.K. Kosek, "Richard Gregg, Mohandas Gandhi and the Strategy of Nonviolence," *Journal of American History* 91, no. 4 (2005): 1318-1348, <http://www.jstor.org>.
4. T. Karon, "The Battle in Seattle: A Challenge to Politics as Usual," *Time*, December 3, 1999, <http://www.time.com/time/arts/article/0,8599,35449,00.html>.
5. A. Berman, "Iran's Twitter Revolution," *The Nation*, June 15, 2009. <http://www.thenation.com/blogs/notion/443634>.
6. Pew Research Center, "Internet Overtakes Newspapers as News Outlet," Pew Research Center Publications, <http://pewresearch.org/pubs/1066/internet-overtakes-newspapers-as-news-source>.
7. G. Lakey, "Mass Action Since Seattle: Seven Ways to Make Our Protests More Powerful," in *The Battle of Seattle: the New Challenge of Capitalist Globalization*, eds. Eddie Yuen, George Katsiaficas, and Daniel Burton Rose (New York: Soft Skull Press, 2001) 166, 225-228.
8. W. Gamson and G. Wolfsfeld, "Movements and Media as Interacting Systems," *American Academy of Political and Social Science* 528, no. 1 (1993): 114-125, <http://www.jstor.org/stable/1047795>. 114-125.
9. S. Guilloud, "Spark, Fire and Burning Coals: An Organizer's History of Seattle," in *The Battle of Seattle: the New Challenge of Capitalist Globalization*, eds. Eddie Yuen, George Katsiaficas, and Daniel Burton Rose (New York: Soft Skull Press, 2001) 225-228.
10. J. Gonzalez and C. Siemaszko, "Protest Blocks Trade Meet: Cops Give Up Fight; Parley Postponed," *New York Daily News*, December 1, 1999, http://www.nydailynews.com/archives/news/1999/12/01/1999-12-01_protest_blocks_trade_meet_co.html.
11. Big Noise Films, <http://www.bignoisefilms.com/films/features/100-whatdemocracylookslike>.
12. M. Hosler, Notes on "The Battle in Seattle" by Someone Who Was There, <http://www.whisperedmedia.org/mhosler.html>.
13. *This Is What Democracy Looks Like*, DVD, directed by IndyMedia (2000; Canada: Big Noise Films, 2003).
14. T. Overman, "WTO: A Primer," *Time*, December 1, 1999, <http://www.time.com/time/daily/special/photo/wto/7.html>.
15. S. V. Medis, "Digital Activism in Seattle," *USA Today*, <http://www.usatoday.com/tech/columnist/ccsam011.htm>
16. http://members.tripod.com/amused_2/seattle.html.
17. J. Perlstein, "An Experiment in Media Democracy," in *The Battle of Seattle: the New Challenge of Capitalist Globalization*, eds. Eddie Yuen, George Katsiaficas, and Daniel Burton Rose (New York: Soft Skull Press, 2001) 335.
18. J. Goldberg, "Clueless in Seattle," *National Review*, December 1, 1999, <http://article.nationalreview.com/266713/clueless-in-seattle/jonah-goldberg>.
19. Rotten Tomatoes: *Battle in Seattle*, http://www.rottentomatoes.com/m/battle_in_seattle.
20. M. Coatney, "D.C. Police Crack Down Early and Hard on Protesters," *Time*, April 16, 2000, <http://www.time.com/time/nation/article/0,8599,43194,00.html>.
21. *Seattle Post-Intelligencer*, "Demonstrators Colorful, Peaceful and Purposeful," December 1, 1999, <http://www.seattlepi.com/local/note01.shtml>
22. S. Sunde, "It's Protest City and Get Ready for the Big One Today," *Seattle Post-Intelligencer*, November 30, 1999, <http://www.seattlepi.com/local/prot30.shtml>.
23. J.S. Juris, "The New Digital Media and Activist Networking within Anti-Corporate Globalization Movements," *Annals of the American Academy of Social and Political Science* 597, no.1 (2005): 189-208, <http://www.jstor.org>.
24. J. Zerzan, "A New Movement." In *The Battle of Seattle: the New Challenge of Capitalist Globalization*, edited by Eddie Yuen, George Katsiaficas, and Daniel Burton Rose, 339-341. New York: Soft Skull Press, 2001.
25. S. Sunde, "Chaos Closes Downtown," *Seattle Post-Intelligencer*, December 1, 1999, <http://www.seattlepi.com/local/prot01.shtml>.
26. *Breaking the Spell*. VHS. Directed by Tim Lewis. 1999; Eugene, OR: CrimethInc, 1999.
27. Lacayo, Richard, Adam Zagorin, Steven Frank, Margot Hornblower, and Jay Branegan. "Rage Against the Machine." *Time*, December 13, 1999. <http://www.time.com/time/printout/0,8816,992808,00.html>.

28. E. Yuen, "Introduction." In *The Battle of Seattle: the New Challenge of Capitalist Globalization*, edited by Eddie Yuen, George Katsiaficas, and Daniel Burton Rose, 12-14. New York: Soft Skull Press, 2001.
29. D.K. Li. "Trade-Talk Riot's A Gas!; Foes Take a Page from the 1960s." *New York Post*, 1 December 1999. http://www.nypost.com/p/news/trade_talk_riot_gas_foes_take_page_Ecs68PkaCrnphd94Qi78VL.
30. "Arab Uprisings." *BBC*, 20 December 2013. <http://www.bbc.com/news/world-middle-east-12813859>.

ACKNOWLEDGEMENTS

I would like to thank Ariane Liazos, Stephen Jones, and Nikhil Chervu for their helpful criticisms of my essay. I would also like to thank Princeton University's Writing Center for their help with multiple rewrites of my paper.



MICHAEL LACHANSKI

Michael Lachanski is currently a graduate student at Princeton University's Woodrow Wilson School of Public and International Affairs. In the long-run, he is pursuing a career at the intersection of government and academia. His research interests include econometrics, finance, human capital policy, mass surveillance, computational social science, and conflict studies.

On an Empirical Analysis of the Kidney Allocation System

Khizar Qureshi¹, Robert Langer

Massachusetts Institute of Technology

The implementation of the new Kidney Allocation System (KAS) has had significant effects on both the offer and acceptance rates of deceased donor kidneys to transplant candidates with high-calculated panel reactive antibodies (CPRA) scores. Linear hypothesis testing on unequal variance between pre-KAS and post-KAS found t-Statistic values of 22.9579 for % offers to high CPRA (99-100) candidates and t-Statistic values of 17.45092 for percent acceptances by high CPRA (99-100) candidates. A simulation of pre-KAS acceptances under the post-KAS system found that high (low) CPRA candidates have an acceptance rate of 1.219 % (63.619) %, compared to 5.429% (1.279)% and 2.207% (1.385)% for pre-KAS and post-KAS actual acceptance rates. This is a significant indication of (i) lowered acceptance by high CPRA candidates based on a higher expected offer arrival rate and (ii) higher acceptance by low CPRA candidates based on a lower expected offer arrival rate.

INTRODUCTION

Patients diagnosed with End-Stage Renal Disease (ESRD) much prefer kidney transplantation over default treatments such as dialysis. While organs become available either through living or deceased donors, there are many more people in need of kidneys than there are available. Specifically, the waitlist currently holds 121,187 individuals². The United Network for Organ Sharing (UNOS) organizes deceased donor organ distribution within the United States through means of priority scoring rules and a waiting list. When a kidney arrives, it is offered to a compatible patient from the waiting list, who may then choose to accept or reject. However, because of organ shortage, several patients, particularly those with an unfavorable position on the waiting list, accept sub-optimal organs. Inversely, patients with waiting list priority often reject organs, with expectations of higher-quality kidneys arriving. This incentive asymmetry, along with ischemic time, (the time between donation and transplant), provide for inefficiencies in organ allocation. To address the issue, researchers have proposed various policies, scoring rules, and methods to minimize the number of wasted organs and maximize patients' life years from transplants. William Bryant and David Throsby look at a creative optimization decision, in which an artist chooses the amount of "creative effort" to invest in a particular work [4]. This choice is conditioned on the artist's expectation of the reaction to the work in two main domains: the "market" and the "world of ideas", or "artworld." They assume the market only cares about the economic value of the product, while the artworld is only concerned with its cultural value. The artist, naturally, is assumed to care about both. Like in earlier work by Throsby [5], this model is realistic in that it constrains an artist to sustain a level of consumption. Thus, financial considerations have a role to play in her choice in that she needs to produce for the market as well as the artworld.

The remainder of this paper contains the following. We review the candidate characteristics considered in kidney allocation. We then provide an overview of the simulation procedure. The primary contribution is the simulated acceptance probability and offer rates across candidate covariates. We use linear hypothesis testing to highlight the significance of the effect of the new scoring system on highly sensitive candidates.

Characteristics

CPRA (Calculated Panel Reactive antibody) is the percentage of donors expected to have Human Leukocyte Antigens (HLA) listed as unacceptable for a candidate on the waiting list [3]. If no unacceptable antigens are entered, CPRA defaults to zero. CPRA scores at or above \$80 %\$ receive 4 points. Any lower score receives zero points. This is a step function. CPRA can also be interpreted as a probability of a positive mismatch or the compliment of the probability of negative mismatch where p_i is mismatch probability for characteristic i .

$$CPRA = 1 - [1 - \sum_i p_i]^2 \quad \text{Equation 1}$$

The CPRA is calculated based on two sets of data:

1. HLA frequencies derived from HLA phenotypes of deceased kidney donors (01/2007 - 12/2008)
2. Ethnic frequencies derived from deceased kidney donors (01/2007 - 12/2008)

Calculation

Step 1: List all unacceptable antigens including equivalencies

Step 2: For each ethnicity, i :

1. Sum all the 1 locus haplotype frequencies (A, B, DR, DQ, C)
2. Sum all the 2 locus haplotype frequencies (AB, ADR, ADQ,

- AC, BDR, BDQ, BC, DRDQ, DRC, DQC)
- Sum all the 3 locus haplotype frequencies (ABDR, ABDQ, ABC, ADRDQ, ADRC, ADQC, BDRDQ, BDRC, BDQC, DRDQC)
 - Sum all 4 locus haplotype frequencies (ABDRDQ, ABDRC, ABDQC, ADRDQC, BDRDQC)
 - Sum all 5 locus haplotype frequencies (ABDRDQC)

$$CPRA = 1 - [1 - S_1 + S_2 - S_3 + S_4 - S_5]^2 \quad \text{Equation 2}$$

Below is an example of the calculation above, scaled by 100

	White	African American	Hispanic	Asian
A1	0.15713	0.05478	0.06246	0.03175
A3	0.13881	0.08041	0.07403	0.02579
B35	0.08472	0.07584	0.16722	0.08929
DR11	0.09416	0.1264	0.07369	0.07143
DQ3	0.125	0.04959	0.15723	0.17206
DQ7	0.14433	0.15552	0.13165	0.13968
Total (S1)	0.74415	0.54254	0.66628	0.53

	White	African Am.	Hisp.	Asian
S1	0.744	0.543	0.666	0.53
S2	0.231	0.135	0.198	0.147
S3	0.042	0.013	0.031	0.019
S4	0.004	0	0.002	0
P(-)	0.202	0.337	0.252	0.358
P(+)	0.798	0.663	0.748	0.642

	White	African Am.	Hisp.	Asian
CPRA	0.798	0.663	0.748	0.642
Ethnic weight	0.689	0.146	0.142	0.023

Table 1, 2, & 3.

Step 3: The final CPRA is a weighted sum of calculated CPRA and the ethnic weight

$$CPRA = \sum_i w_i CPRA_i \quad \text{Equation 3}$$

We see that the CPRA is 83.024.

Establishing Criteria for Listing Unacceptable Antigens

Centers determine their own criteria for unacceptable antigens.

The weights provide increased benefit to racial/ethnic minority candidates who are sensitized to HLA antigens that are relatively common in the deceased donor population. The probability of a positive crossmatch is a more accurate measure of sensitization and benefits minorities. [6]

Kidney Criteria

In addition, kidney quality is compartmentalized into ECD (expanded criteria) [4] and SCD (standard criteria). The former refers to kidney donors of at least 60 years of age and those of at least 50 years of age with two out of three special conditions, which include hypertension, terminal serum creatinine, or death from cerebrovascular accident. The latter refers to all non-deceased donors who do not belong to the ECD category. The trade-off for receiving a lower quality kidney is a reduced expected waiting time. The issue, however, is that often ECD kidneys have higher realized graft survival than SCD kidneys.

KIDNEY-PANCREAS SIMULATED ALLOCATION MODEL (KPSAM)

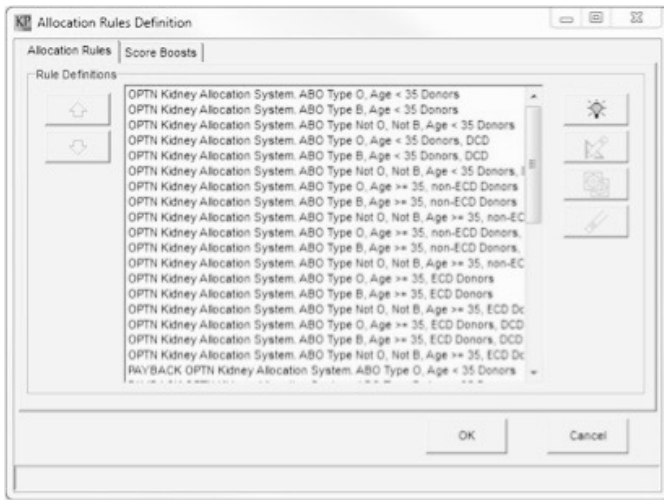
The purpose of KPSAM is to simulate the allocation of kidneys and/or pancreases to waitlisted transplant candidates. It provides an empirical outcome, given a perturbation in the allocation policy. The input data is typically based on Organ Procurement and Transplantation Network (OPTN) waiting lists and organs, and the program itself was developed by the Scientific Registry of Transplant Recipients (SRTR) [2]. It uses a variety of allocation rules, often accommodating special candidates³, alongside candidate/organ-specific characteristics to uniquely identify the allocation to a set of recipients under each rule or constraint. To reflect uncertainty of success and life expectancy, the software introduces a stochastic component through means of repeated simulation. After several iterations of applying these rules to candidates and organs, it is possible to define and determine the conditional distribution of outcomes.

The data that UNOS and OPTN provides includes not only patient observables such as blood type, unacceptable antigens, and geographic location, but also relevant information on the donor, as a measure for compatibility, and records of waitlisted patients' decisions to accept/reject organs offered.

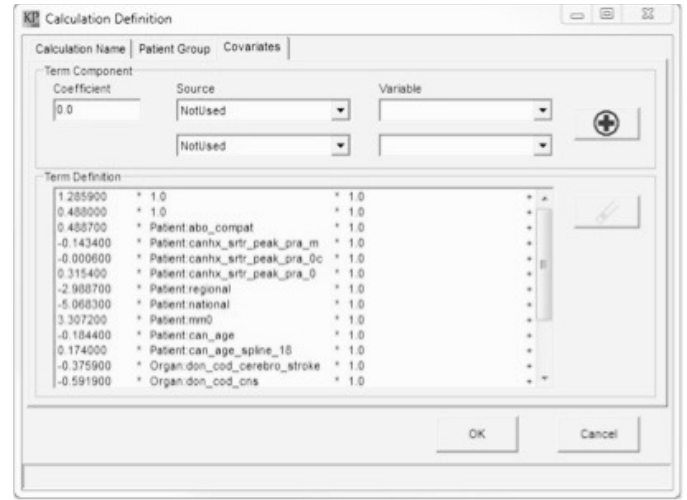
Kidneys from deceased donors are offered to waitlisted candidates according to a priority score. The priority score depends on the following donor-patient specific characteristics

1. Blood-type
2. Tissue-type compatibility
3. Duration on waiting list
4. Duration on dialysis
5. Age
6. Proximity to transplant center
7. Outstanding medical conditions

Ultimately, the likelihood that the patient is offered a donor organ depends on his/her priority score and a set of allocation rules that determines the weight of each factor contributing to the score. One rule may prioritize a geographic advantage, such as proximity to transplant center, to minimize ischemia time, while another may prioritize tissue-type compatibility, to maximize likelihood of transplant success. It follows that the implementation and existence of various rules results in varying expectations of



(a) Allocation Rules



(b) Patient Covariates

Figure 1: The figures above show the interface of the KPSAM Software. Figure (a) shows sets of rules that prioritize patients based on age, blood-type, and status. Figure (b) shows sets of patient characteristics used as inputs for scoring such as age, location, and medical history

patients, which then affects the individual propensity to accept or reject an organ.

Given a set of data containing donor-patient specific characteristics, patient specific characteristics, various allocation rules, and historical decisions, we simulate possible outcomes. Below is an example from the KPSAM software, in which a set of rules is entered for candidates with a particular set of covariates (characteristics).

However, the covariates represented in the simulation above are only a subset of criteria used to evaluate a kidney pairing. For each standard criteria donor, recipient candidates are prioritized through means of a decision tree. The most sensitive criteria for evaluation, which include geographic location, age, past donor status, and antigen compatibility, are sorted in Figure 2.

Overview of Probability Model

For each patient in set X with characteristic β for which data is available, and an organ is offered, such a simulation is performed. The resulting value, $X\beta$, will be transformed using an inverse logit transformation.⁴

$$p_i = \frac{\exp(X\beta)}{1 + \exp(X\beta)}, p_i \in [0, 1], \forall i \tag{Equation 4}$$

The below calculation provides the probability that a patient accepts or rejects an organ that is offered. In particular, a random number (n) between 0 and 1 will be generated. For all patients for which $p_i < n$, the organ will be accepted, and for all patients for which $p_i \geq n$, the organ will be rejected. If the entire set of patients refuse a particular organ, it will be labeled as “discarded”.

$$\text{Accept/Reject: } \begin{cases} 1 : p_i < n \\ 0 : p_i \geq n \end{cases} \tag{Equation 5}$$

Significance of Calculation

The calculation above is significant because the simulation helps determine under which allocation rules, for particular sets of patient characteristics, the probability of organ acceptance maximized. Such an output helps in understanding patient-decision mechanics and allocation rule performance. To minimize the number of organs discarded, a considerable objective when facing organ shortage, the allocation rule that maximizes overall probability of acceptance is used.

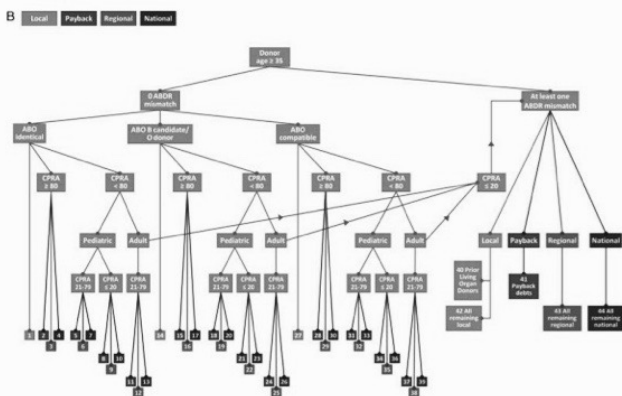


Figure 2: The figure above is an example of a decision tree for kidney allocation from standard donors. Notice that upon evaluation of blood-type, age, location, and antigen compatibility are considered. [5]

% Offers by Category	pre-KAS	post-KAS	t-Statistic
CPRA 0	0.706	0.766	-12.919
CPRA 1-79	0.192	0.225	8.367
CPRA 80-94	0.028	0.031	3.724
CPRA 95-98	0.007	0.008	1.476
CPRA 99-100	0.004	0.026	22.958

Table 4: The table above shows the (i) offer rate (ii) t-Statistic across CPRA category for pre-KAS and post-KAS

% Accept by Category	pre-KAS	post-KAS	t-Statistic
CPRA 0	0.013	0.014	-7.1
CPRA 1-79	0.019	0.022	-15.351
CPRA 80-94	0.053	0.041	19.567
CPRA 95-98	0.057	0.046	12.586
CPRA 99-100	0.056	0.022	17.451

Table 5: The table above shows the (i) acceptance rate (ii) t-Statistic across CPRA category for pre-KAS and post-KAS (Left)

Accept Rate by CPRA	pre-KAS	post-KAS	pre-KAS on post-KAS
CPRA 0	1.28%	1.39%	63.62%
CPRA 1-79	1.90%	2.24%	23.26%
CPRA 80-94	5.29%	4.13%	9.39%
CPRA 95-98	5.76%	4.64%	2.97%
CPRA 99-100	5.43%	2.21%	1.22%

Table 6: The table above shows the comparison of acceptance rates across CPRA categories between (i) pre-KAS, (ii) post-KAS, (iii) Simulated pre-KAS under post-KAS rules (Right)

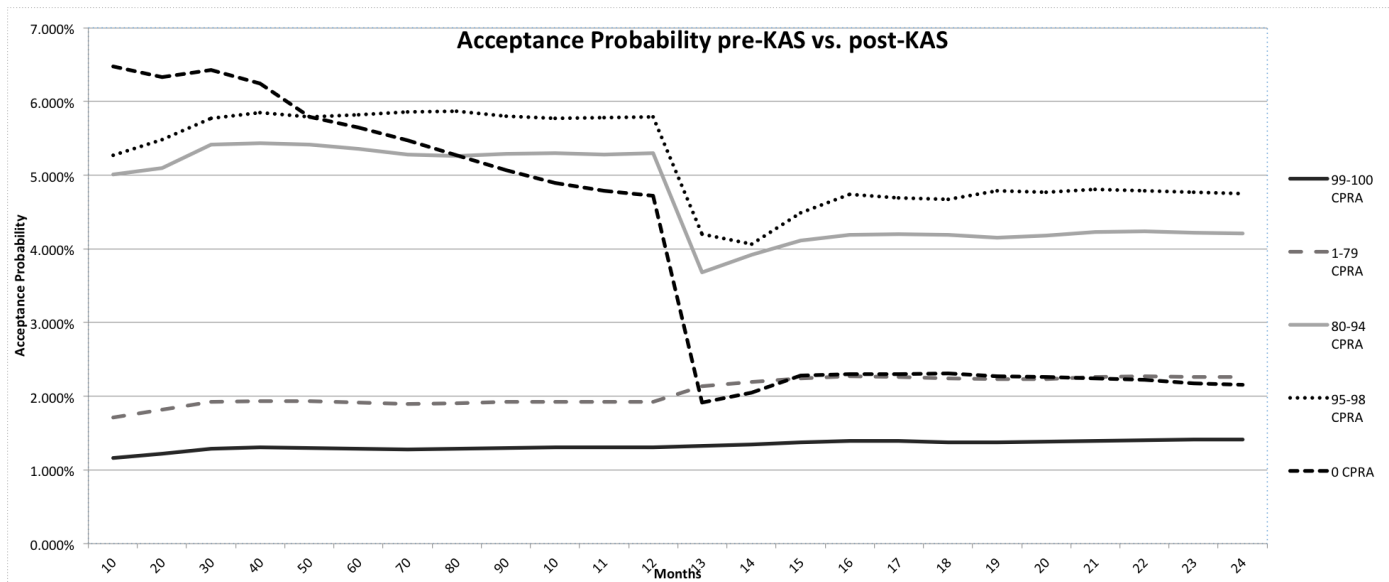


Figure 3: The plot shows the acceptance probability pre-KAS (0-12 months) and post-KAS (12-24 months) across CPRA levels

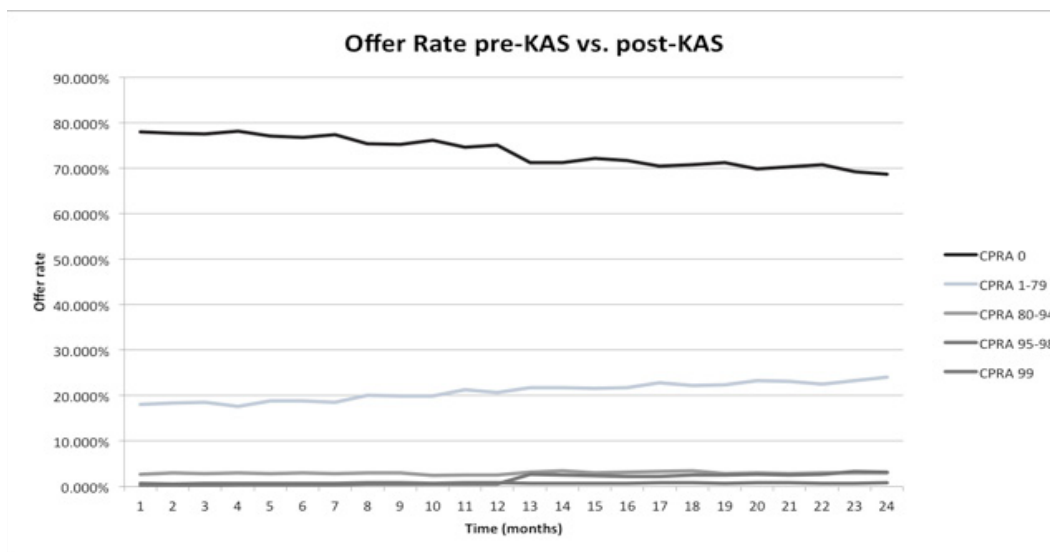


Figure 4: The plot shows the offer rates pre-KAS (0-12 months) and post-KAS (12-24 months) across CPRA levels

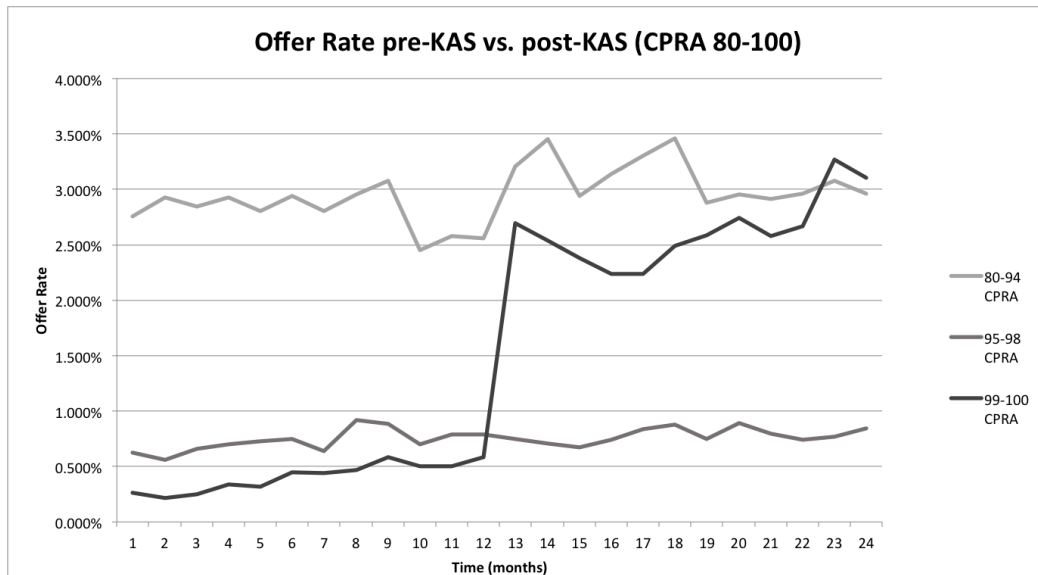


Figure 5: The plot above shows the offer rates pre-KAS (0-12 months) and post-KAS (12-24 months) for mid-high CPRA levels

Mean Top Offers (pre-KAS)	
Patient Age	50.1496
Donor Age	40.3997
CPRA	18.3741

Table 7: The table above shows the mean characteristics for each subset of top offers per organ prior to the introduction of KAS

Mean Top Offers (post-KAS)	
Patient Age	48.3328
Donor Age	39.8053
CPRA	32.8605

Table 8: The table above shows the mean characteristics for each subset of top offers per organ following the introduction of KAS

Decision (000's)	0	1-79	80-94	95-98	99-100	Total
Accept	555.86	138.25	19.51	5.01	2.87	721.51
Decline	7.41	2.72	1.09	0.35	0.13	11.71
Frequency	0.768	0.192	0.028	0.007	0.004	

Table 9: The table above shows the pre-KAS absolute count and frequency of candidate decisions by CPRA level

Decision (000's)	0	1-79	80-94	95-98	99-100	Total
Accept	463.25	146.57	19.69	4.93	17.04	651.48
Decline	6.73	3.4	0.923	0.271	0.395	11.72
Frequency	0.709	0.226	0.032	0.008	0.026	

Table 10: The table above shows the post-KAS absolute count and frequency of candidate decisions by CPRA level

RESULTS

Offer Rate by CPRA Category

Here we consider the acceptance rate by candidates and offer rate to candidates across CPRA categories. See Table 4.

Acceptance Rate by CPRA Category

See Tables 5 & 6.

Time Series of Acceptance Probability and Offer Rates

Here we study acceptance probability and Offer Rate of deceased donor kidneys pre-KAS in months 1-12 and post-KAS in months 12-24 across CPRA levels. See Figures 3, 4, & 5.

Covariates for Top Offers

Here we consider the covariates of the first ten candidates to which each donor kidney was offered following procurement under the two policies. See Tables 7, 8, 9, & 10.

DISCUSSION

Here we discuss key results pertaining to the offer rates to and acceptance rates by transplant candidates.

CPRA

The introduction of the KAS meant that patients with high CPRA levels would receive more points, shortening the queue for

higher quality organs. Candidates with a measurably higher CPRA rate (99-100) experienced a higher offer rate post-KAS (2.63 % vs. pre-KAS (0.41 %). The t-Statistic for unequal variance was 22.958.

Conditional on the new information, candidates with high CPRA levels developed an increased expectation for donor kidney arrival, and increased threshold for acceptance. Candidates with low CPRA levels (0) developed a decreased expectation for kidney arrival, and decreased threshold for acceptance. Candidates with measurably higher CPRA rates (99-100) had a lower acceptance rate post-KAS (2.207 %) vs. pre-KAS (5.429 %). Candidates with a measurably lower CPRA rate had a lower acceptance rate post-KAS (2.207 %) vs. pre-KAS (5.429 %). A simulation of acceptance rates of low CPRA pre-KAS, following post-KAS rules had a far higher acceptance rate (63.619 %), reflecting the instantaneous dispersion of information. Similarly, a simulation of acceptance rates of high CPRA pre-KAS, following post-KAS rules had a far lower acceptance rate (1.219 %), reflecting the instantaneous dispersion of information. While un-realistic, the latter simulation reflects the immediate effect of a policy, while the former study reflects the gradual effect.

Another consideration is the set of candidate characteristics for those receiving the top 10 offers for each kidney. The average CPRA level of a candidate receiving a top 10 offer increased from 18.3741 to 32.8605 (Table 7, 8). While the mean patient and donor age remain unchanged, this increase in mean CPRA indicates that post-KAS is able to allocate arrivals to those with high sensitivity nearly independent of patient and donor age.

CONCLUSION

Candidates with higher sensitivity to panel reactive antibodies (PRA) are compatible with fewer donor kidneys, and the introduction of the Kidney Allocation System (KAS) was meant to address this concern. The KAS point system allocates 202 points to high calculated panel reactive antibodies (CPRA) (99-100) candidates, 0 points to low-CPRA (0) candidates, and an

intermediate amount of points, as defined by a convex function, to mid-CPRA (1-97) candidates. The lower acceptance rate by high CPRA candidates is indicative of increased expectation of donor arrival. Such a result may have implication on the strategies that candidates would formulate to receive the kidneys, including (i) geographic, (ii) racial, and (iii) duration on dialysis. Registration in multiple centers, dishonesty about race, and increased dialysis would all increase offer rates. Such possibilities motivate further study. We also recognize the potential for depth in econometric analysis. The main contribution of this study, however, was to highlight the immediate effect of KAS on those with high PRA Sensitivity.

REFERENCES

1. Miles C. Update from the United Network for Organ Sharing. OPTN 2015
2. Rao PS., Akinlolu O. The Alphabet Soup of Kidney Transplantation: SCD, DCD, ECD-Fundamentals for the Practicing Nephrologist. *Clinical Jour. of the Am. Soc. of Nephrology* 2009; 1827-1831
3. Cho YW. Expanded Criteria Donors. *Clin. Transplants* 1998; 421-436}
4. Sung RS, Guidinger MK, Lake CD, McBride MA. Impact of the Expanded Criteria Donor Allocation System on the use of Expanded Criteria Donor Kidneys. *Transplantation* 2005; 79: 1257-1261
5. Cox, DR. Partial likelihood. *Biometrika* 1975, 62 269-276
6. Nyberg SI, Matas AJ, Kremers WK, et al. Improved Scoring System to Access Adult Donors for Cadaver Renal Transplantation. *Am. Jour. for Transplantations* 2003; 3:715

FOOTNOTES

1. Support from Paul E. Gray (1954) Fellowship. David H. Koch Institute for Integrative Cancer Research, Langer Lab, generously provided data on PRA sensitivity. Itai Ashlagi, Associate Professor, Nikhil Agarwal, Assistant Professor, Daniel Waldinger, Ph.D. candidate, provided generous feedback. A separate part of the study was completed with the Department of Economics, Massachusetts Institute of Technology
2. According to the National Kidney Foundation, as of 04/03/2016
3. Special candidates include past donors, diabetics, and newborns
4. The inverse logit transformation outputs an odds-probability based on the inner product of characteristics and their regressors

KHIZAR QURESHI

Khizar is an undergraduate at the Massachusetts Institute of Technology studying Mathematics and Chemical Engineering. His primary interests include statistics, econometrics, and theoretical computer science. He hopes to enter a PhD program relevant to those fields.



ROBERT LANGER

Robert Langer is the David H. Koch Institute for Integrated Cancer Research. He holds a BS in Chemical Engineering from Cornell University and a Sc.D. in Chemical Engineering from the Massachusetts Institute of Technology. He is the most cited engineer in history, and has won countably many awards.

SCIENTIFIC WHIMSY

JAKE HILLARD & RESEARCH AS EDUCATION

by Ankit Baghel

“There are two layers: I’m doing something that is intellectually challenging and interesting that no one has ever done before, and also I get to hang out with some really cool people.”

These are the two factors that guide Jake Hillard’s research. For Jake, research is about learning outside of the classroom. And unlike many young students, Jake’s attraction to research didn’t stem from a single field of study, but from the professors conducting the research. The driving force for his beginning research was his conversations with a professor at the University of Utah during Jake’s time in high school. “The best thing in the world is when you can get into a room with someone at the end of the day having talked about ten different projects that are interesting and fun.” Since his interest does not lie in a single topic, Jake has undertaken multiple projects both inside and outside the lab. Outside of the lab, his interests have inspired small projects ranging from unboiling an egg to building a one-wheel self-balancing skateboard using a DIY guide he found online.

Jake’s biggest project took place during his Freshman year at Stanford when he joined the Stanford Space Initiative’s (SSI) and Los Alamos National Lab joint project FINDER (Field Integrated Network Driven Entity Recognition), a project that focuses on creating a network to facilitate cooperation between groups of satellites. Jake used Distributed Sensor Networks, networks of sensor nodes that have equal control over the flow of information between them, to collect and exchange large amounts of data. His goal was to track human movement through a field of nodes. Using Raspberry Pi’s cheap, credit card-sized computers as nodes, he wrote software that

“gave them a way to see the world, a way to see friends, and a way with which to exchange information.” Two nodes can share information over a distance of 20 yards, so five nodes chained together can “leapfrog the information” and track movement over 100 yards. Jake recalls that while testing these nodes, “it was really exciting when I’d walk through the field, and each node would notice as I passed within range and indicate that it saw me.”



The applications of Distributed Sensor networks extend to all types of flow analysis involving large amounts of data. They were originally used on a large scale by the Department of Defense to locate the sources of radio transmissions through reverse triangulation. What sets FINDER apart is that it can be created using off the shelf materials and on a much smaller scale. It gives people more functionality and promotes learning by allowing them to build a do-it-yourself distributed sensor network.

Now an Electrical Engineering major in his sophomore year, Jake continues to promote education through his work in the Kovacs Lab on frequency layering in complex systems. He studies systems through oscillators: mechanical or electrical devices that operate through repeated fluctuations in energy between two states. For a simple example, think of the pendulum in an old clock. Electric oscillators vibrate when an electric current is applied to them. The frequency of this vibration depends on the material used to make the oscillator and on how the oscillator is made. Normally, one would expect two identical oscillators to have completely identical frequencies. However, oscillators are only accurate within a few parts per million. This may seem to be a small inaccuracy, but the effect is compounded when multiple oscillators are used in a system. These discrepancies can compound enough to render a system ineffective, especially when working with analog control systems, systems that convert information into signals of varying amplitudes (eg. clocks with minute and hour hands) as opposed to digital systems that use binary systems, or on/off switches, to store information (eg. digital clocks). It is because of these discrepancies that electronics like televisions and computers have moved away from analog circuits to digital circuits that are easier to implement and have reduced error.

In order to describe frequency error, Jake has built a circuit board containing several oscillators. It demonstrates how the addition or multiplication of “identical” frequencies results in layering. When adding frequencies, this layering produces beat frequencies (due to small differences in frequency), and when multiplying frequencies, it produces harmonics. The main role of the circuit board is that of an education tool to help people understand the basic concepts of embedded systems and why frequency error is an important “source of failure.” Research as an educational tool is important to Jake. He believes that, “if you just learn from school, a lot of people’s knowledge is fragile.” A concept may make sense in a given example that is taught in class, but once that example is slightly altered, all perceived understanding dissolves. To Jake, research, is “a way to apply ideas and explore the world.”

engineering

State Transition Selection from Absolute Limb Orientation in a Powered Transfemoral Prosthesis

Maurice Chiang

Stanford University

In order for prostheses to effectively restore healthy human locomotion, they must work in conjunction with the user's intact neuromuscular system. Given the potential benefits of active prostheses, it is necessary to expand research on a robotic leg interface and control system that provides appropriate coordination and intent recognition for locomotion. Consequently, this work investigates the use of an alternative biomechanical metric to improve transitions in current state-based control systems. By analyzing healthy human subject biomechanical function and extracting the orientations of the lower limb segments relative to gravity, we developed an improved condition to instigate the state transition from stance into push-off and swing for a powered knee and ankle on various slopes. This condition, which consists of a threshold on the thigh angle orientation, is shown in this work - both in healthy subjects and in an amputee subject - to be less sensitive to slope variations. Additionally, the thigh angle derivative is of notably greater magnitude than that of the previous condition (based upon ankle angle), which makes the threshold less sensitive to tuning. Because powered prostheses have only begun to emerge in the last decade, current literature lacks in-depth research into appropriate and reliable conditions for coordinating state transitions in a powered prosthesis with the motion of the user. Analysis of alternative conditions may improve the performance of prosthetic devices in reproducing healthy human walking biomechanics.

INTRODUCTION

With over 12 transfemoral amputations performed per 100,000 people in the United States each year (i.e., 30,000 transfemoral amputations performed yearly) [1], [2], progress in lower limb prostheses is of great importance. However, current products are predominately energetically passive, with prosthetic feet made of carbon-fiber leaf springs and prosthetic knees composed of rotational dampers. While these devices provide limited levels of resistance to thigh and shank movement [3], they do not provide actuation, or energy, to the knee and ankle. Because there is no net power at lower limb joints, they fail to emulate normal locomotion thus requiring the amputee to compensate for the loss of push-off at the ankle and knee joints by exerting up to four times as much hip torque and power during late stance [5], thus decreasing the efficiency of walking and potentially causing long-term musculoskeletal degeneration. Transfemoral amputees experience up to 60% greater metabolic cost [4] as compared to healthy gait. Additionally, prosthetic knees that behave as dampers are unable to provide static torque, thus inhibiting the amputee from achieving stability when the knee is not pushed against its hyper-extensive hard stop (which is often the case when standing on uneven ground) [6].

Given the musculoskeletal system's inherent complexity, the passive prosthesis does not provide an appropriate substitution for healthy biomechanics. However, recent technological advances in active prosthetics attempt to avoid these problems. With growing research in improving the power density of brushless DC motors,

increasing the energy density of batteries, and reducing the size and power consumption of microcontrollers and inertial measurement units (IMUs), new powered prosthetic devices demonstrate strong potential in better reproducing healthy kinematics, reducing pathologies in the residual limb and other intact joints, and decreasing the metabolic cost of locomotion [7]-[11].

Falling remains a chronic problem among lower-limb amputees; among those tracked in a 2001 survey, around half of the participants experienced at least one fall within a one month period [12]. For a prosthetic intervention, working in concert with a user's neuromuscular system is imperative in replicating healthy biomechanics. Prostheses, with power generation capabilities in line with biomechanically healthy function, could not only reduce the need for exertion of the amputee during gait, but could also alleviate amputee instability. The functionality of the powered system is dependent upon how well the prosthesis' control system works in concert with the user. Unlike the passive prosthesis, which receives its power exclusively from the user - and thus inherently moves in biomechanical coordination with the user - the active prosthesis requires a reliable control interface in order to prevent movement in discord with the user's intent.

This problem in active lower limb prostheses is currently a topic of limited research. Although real-time intent recognition research has been executed [13], [14], current research is lacking in push-off condition analysis. Current cooperative control frameworks that use physical sensor interfaces have made progress in effectively controlling the robotic prosthesis [7]-[9],

[15], [16], however, much work needs to be done to normalize control framework algorithms in conjunction with the intent of the central nervous system. As evident in Figure 1, plotted from healthy human gait collected by Winter and McIntosh et. al. [5], [17], the ankle angle and knee angle recorded over varying levels of slope are far from invariant. Rather, these conditions experience different behaviors at different gradients. Both the ankle and the knee angle exhibit significant variance in angle at push-off (approximately 40% of stride). Despite increasing research on powered prostheses, sloped walking research has been limited [18]. One algorithm for active prostheses utilizes the ankle angle as the push-off condition, with push-off angles varying from 8 to 13 degrees [8]. Humans, who naturally adapt to gradient change, experience significant adjustment in body movement to account for incline variation [19]. In uphill walking, increased knee flexion and ankle dorsiflexion during swing and initial ground contact allow for sufficient foot clearance and proper foot positioning [18], [20]. Additionally, in downhill walking, knee flexion also increases during load-bearing, push-off, and early swing phases; the ankle experiences exaggerated dorsiflexion from late stance to mid-swing, allowing for proper foot placement [18]. Thus, to accurately reproduce healthy biomechanical function, a control program must accurately be able to time push-off in a powered prosthetic device at numerous slopes.

Through analysis of healthy subject data over various slopes, corroborated by McIntosh et al., [17], this study explores absolute angles in the lower limb segments, in pursuit of a decreased level of variance in gait pattern over sloped ambulation, a more sensitive and robust push-off condition, and a more intuitive control framework. Given the lack of exploration into absolute angles as potential conditions for state-transition selection, we hypothesized that the thigh angle would provide significantly decreased angle variation and greater sensitivity, which would serve as a more effective condition for reproducing healthy biomechanical function in an active prosthesis control algorithm. By implementing this, we measured the effectiveness of the alternative thigh angle condition in an amputee subject. This result is important in growing understanding of slope and variations in human gait as well as improving the functionality and control framework of a powered prosthesis.

METHODS

Healthy Subject Data Collection and Analysis

Subject information

Healthy subjects aged 18-27 (one male, two females) took part in the study – the significant number of strides per patient yielded consistent results representative of transfemoral amputees. Each subject signed an IRB consent form and consented to the use of their personal information in association with the study. The Vanderbilt IRB approved the study.

Walking evaluation and recording

The subjects walked on an adjustable walkway (12', $\pm 0.1'$) at -10° , -5° , 0° , 5° , and 10° ($\pm 0.5^\circ$) without assistance. Each subject was instructed to don an OptiTrack motion capture suit equipped with motion capture markers. The participants were allowed to practice walking at the various degrees of incline at a fixed cadence of 105 steps per minute. At the predetermined cadence, all individuals were instructed to complete 20 strides at each slope grade. To eliminate the potential for confounding gait variations due to exhaustion, subjects were instructed to rest after each grade was completed.

Data were acquired using 12 OptiTrack motion-capture cameras provided by NaturalPoint, Inc., and were recorded in an Arena, a motion capture software environment, at 100 Hz. The motion-capture cameras were dispersed around a room in a fashion that maximized reliability and precision. Reflective markers were placed on central anatomical landmarks, in accordance with procedures outlined by the Arena operation manual, which roughly correspond to the Helen Hayes marker set [21]. After each ascent and descent of the adjustable walkway, a new sample file was taken.

Data Analysis

Twenty strides were selected for analysis from each grade level. The strides were then normalized over one gait cycle and averaged, with heel strike denoting a new stride. The OptiTrack system used the 34 markers on the subject to fit a skeletal model to the subject, from which Euler angles were derived and saved in a BVH file.

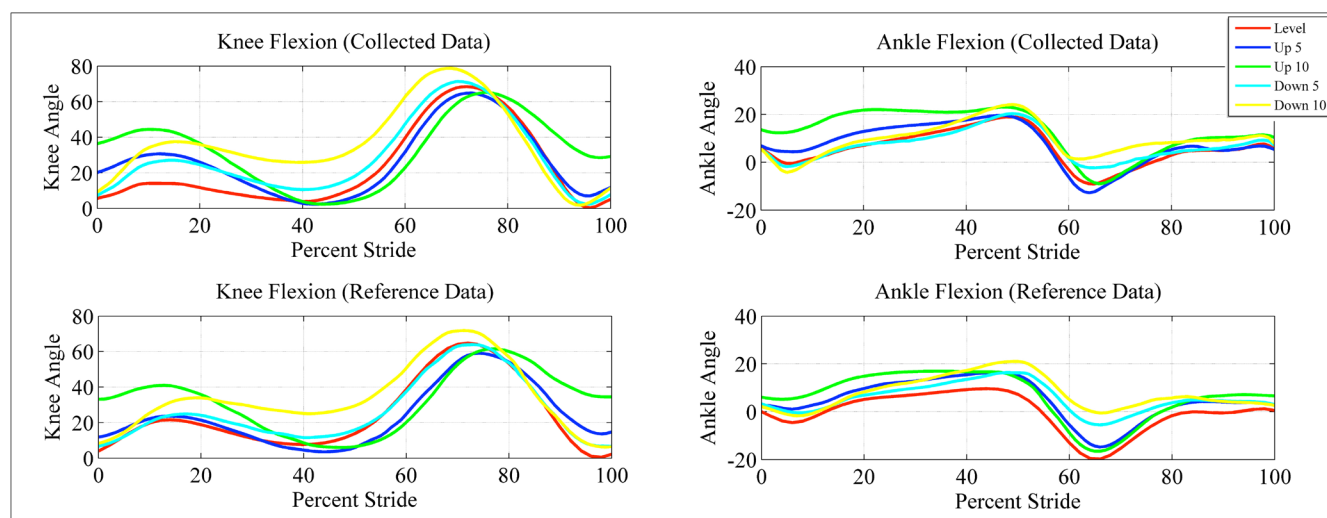


Figure 1: A graph of ankle and knee flexion of experimental data versus data from existing literature.

These data were imported into MATLAB and used to calculate internal and absolute angles. These segment and joint angles were calculated based on the conventions displayed in Figure 2. The figure indicates the direction for positive and negative angles. Thigh angle was calculated with respect to the gravity vector, with a neutral position equivalent to 0° of displacement. The ankle and knee angles were calculated with respect to other limb segments. The knee angle was calculated with respect to the thigh and the shank segments and the ankle angle was calculated with respect to the neutral standing position, with the shank perpendicular to the foot.

Points at the start and end of the gait cycle were selected (heel strike to heel strike), in accordance with an methodology [5], from each of the 20 gait cycles at a given incline; the selected strides were then averaged across the three healthy subjects. The data were analyzed in MATLAB and then plotted against data from existing literature ([5], [17]). The angles in the collected data

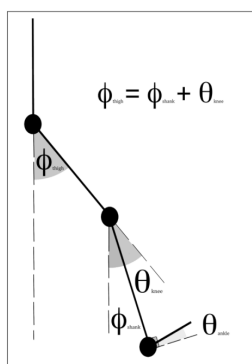


Figure 2: A representation of angel conventions in the leg – ϕ denotes an absolute angle, and θ denotes an internal angle.

Factor	Subject
Height	1.93 m
Weight	79 kg
Age	25 yrs
Cause of Amputation	trauma
Years Post-Independent Ambulation	7
Amputation Side	right
Prosthesis Configured Weight	509 kg
Prosthesis Configured Height	54.0 cm
Prosthesis Foot Size	28 cm

Table 1: Mechanical and Electrical Characteristics of the Powered Prosthesis

demonstrated strong agreement with the angles in the existing literature.

Upon finding that strong agreement (within one standard deviation) did indeed exist between the existing literature and the collected ankle and knee data, the thigh angle data was then extracted from the collected healthy human gait. A derivative of the thigh data was taken, the results were analyzed for the robustness of the condition at 40% of stride, and then they were compared with the angle and derivative of the existing ankle angle condition.

Prosthesis Design

Physical Structure and Actuation

The powered prosthesis, as pictured in Figure 3 (page 8), was designed with a powered knee and ankle unit, an adjustable pylon, a Lithium-ion battery, a load cell, and an embedded system to control and power both joints. The knee joint was actuated by a brushless Maxon EC30 motor and a three-stage speed reduction transmission. The ankle joint was actuated by a brushless Maxon EC60 motor and a three-stage speed reduction transmission. The specifications of the joint torques and range of motion are detailed in Table 1. The actuator output at the ankle was aided by a carbon-fiber spring in the foot, which helped bias the ankle torque toward plantarflexion, reflecting the fact that humans have proportionally larger gastrocnemii than tibiales anterior. The prosthesis was designed to be capable of handling up to 110kg of weight, and to fit all adult males and 90% of the adult female population.

Sensing and Control

Orientation sensing was conducted through a 6-axis IMU located in the shank, and several other contactless sensors were also used. Absolute and incremental magnetic encoders measured the angular positions and velocities of the knee and ankle joints; filters reduced high-frequency noise and accounted for drift and signal artifacts. Axial load in the shank was measured by

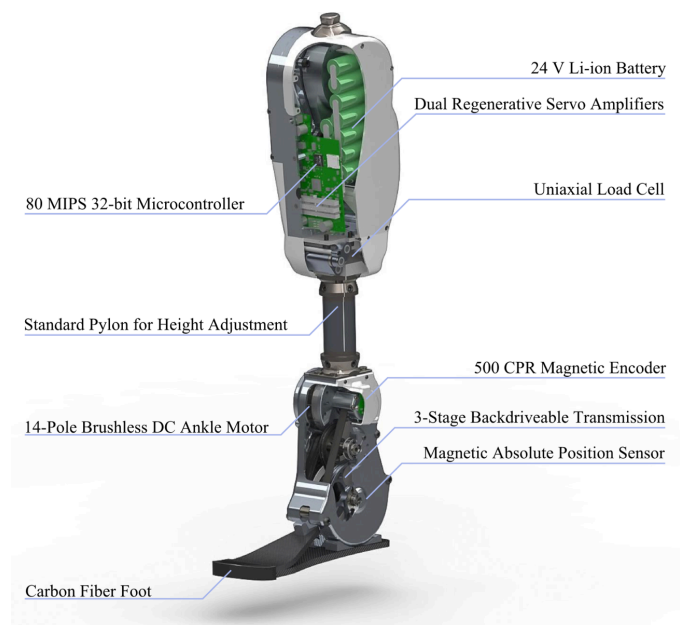


Figure 3: A rendering of the CAD model of the powered prosthesis

linear displacement of a double parallelogram linkage load cell, as seen in Figure 3. The prosthesis was controlled through an embedded electrical system, governed by a main processor (32-bit microcontroller from Microchip Technologies), which executed the joint torque controllers, logged data, and communicated with peripheral controllers, including the servo controller (a 16-bit microcontroller from Microchip Technologies).

Several controllers in the prosthetic leg determined the appropriate joint kinematics. In all controllers, net power generation occurred only during state transitions (e.g. between swing and swing extension). State transitions were dictated by biomechanical cues such as heel strike or push off. However, different gait dynamics occur at different levels of slope. As seen in Figure 1, upslope walking and downslope walking resemble that of level-ground walking; however, joint kinematics were far from consistent over varying slopes. Thus, appropriate alternative controllers must be instigated. The level-ground walking controller is presented in [22]. The upslope-walking controller made several variations to the level-ground walking controller. Several impedance parameters were altered, including increasing knee flexion after heel strike and changing knee stiffness to provide net knee extension, which allowed the user to lift his/her center of mass upwards on a slope. In the upslope controller, ankle flexion trended toward increased dorsiflexion. The downslope-walking controller also followed the general trend of level-ground walking; however, the controller was altered to increase knee flexion and ankle dorsiflexion to allow for increased damping, with the leg functioning as a shock absorber.

Amputee Subject Procedures

The test subject, age 26, was a transfemoral amputee. The subject's metrics are detailed in Table 2, and consent was obtained in the same fashion as with the healthy subjects. The methodology for obtaining data from the amputee subject was similar to healthy subject data acquisition. The subject was instructed to walk up and down a 5-degree incline as well as ambulate on level ground.

The accelerometer and gyroscope signals provided shank data from the prosthesis, while an absolute magnetic encoder and an incremental motor encoder on the knee provided knee angle. The data were then analyzed in MATLAB for the thigh angle push-off condition in the thigh angle by using the inertial measurement data.

EXPERIMENTAL RESULTS

1. Healthy Subject Data Relative to Literature

Healthy subject data were collected relative to existing literature ([5], [17]). After the strides were parsed, the ankle, knee, foot, shank, and thigh angles were calculated for the three healthy subjects. From the averaged data (a mean of 60 total strides for slopes of -10, -5, 0, 5, and 10 degrees), a literature comparison was conducted. The collected data (Figure 1) compared favorably with Winter and McIntosh data. Although there was less stance knee flexion in level ground walking and reduced plantarflexion throughout all of the collected data, the deviations were minimal, as the literature data fell well within one standard deviation of collected data. Given the confirmation of healthy walking biomechanics replication, absolute angles could be extracted from the data.

Characteristic	Value
Maximum Knee Torque	85 Nm
Maximum Ankle Torque	110 Nm
Knee Range of Motion	-5° to 115°
Ankle Range of Motion	-45° to 25°
Battery Capacity	125 W · hr
Maximum Battery Current	30 A
Maximum Motor Current	18 A
Mass	4.75 kg

Table 2: Subject Data and Prosthesis Configuration

Push-Off Condition Analysis

The thigh angle, an absolute angle, was analyzed as potential push-off condition candidate. As seen in Figure 4, the thigh angle condition exhibited a greater average derivative magnitude (1.200°/ % stride) over various slopes than the ankle angle condition (0.4627°/ % stride) over varying slopes. Additionally, the spread of derivative, as indicated by the bow-tie plot on both the thigh and the ankle plots, was markedly less in the thigh angle condition.

Signal Construction and Processing

In order to use the absolute orientation of the thigh as a condition to instigate push off, an estimate of this quantity must be generated by the powered prosthesis. As seen in Figure 5, the thigh angle signal was created through data processing of the IMU signals. The gyroscope data were integrated, yielding a signal representing the shank angle. The x and y components of the accelerometer vector were used as a low frequency estimate of the shank orientation. This estimate was computed using an arctan function. The signals were filtered using complementary filters (time constant of 1.000 s, -3 dB cutoff of 0.1591 Hz), with the high-pass filter being applied to the gyroscope-derived signal and the low-pass filter applied to the accelerometer-derived signal. The fusion of these two processed signals reduced drift (relative to the integrated signal from the gyroscope) and improved the bandwidth (relative to the accelerometer-derived signal) of the shank angle. The shank angle was summed with the knee angle (Figure 2) to create the thigh angle signal. The new controller was tuned for push-off between Phases I (middle stance) and II (late stance and swing) in the prosthesis' control interface.

Amputee Tests and Implementation

The thigh angle push-off condition was implemented in the prosthetic device. Over level ground walking, as indicated in Figure 6, the gait dynamics (represented by the ankle and thigh angles) predominantly fell within one standard deviation of healthy biomechanics. Data for both ankle and thigh flexion in the amputee subject were collected for 65 consecutive strides, with push-off occurring when the thigh angle reached 12°. Among the

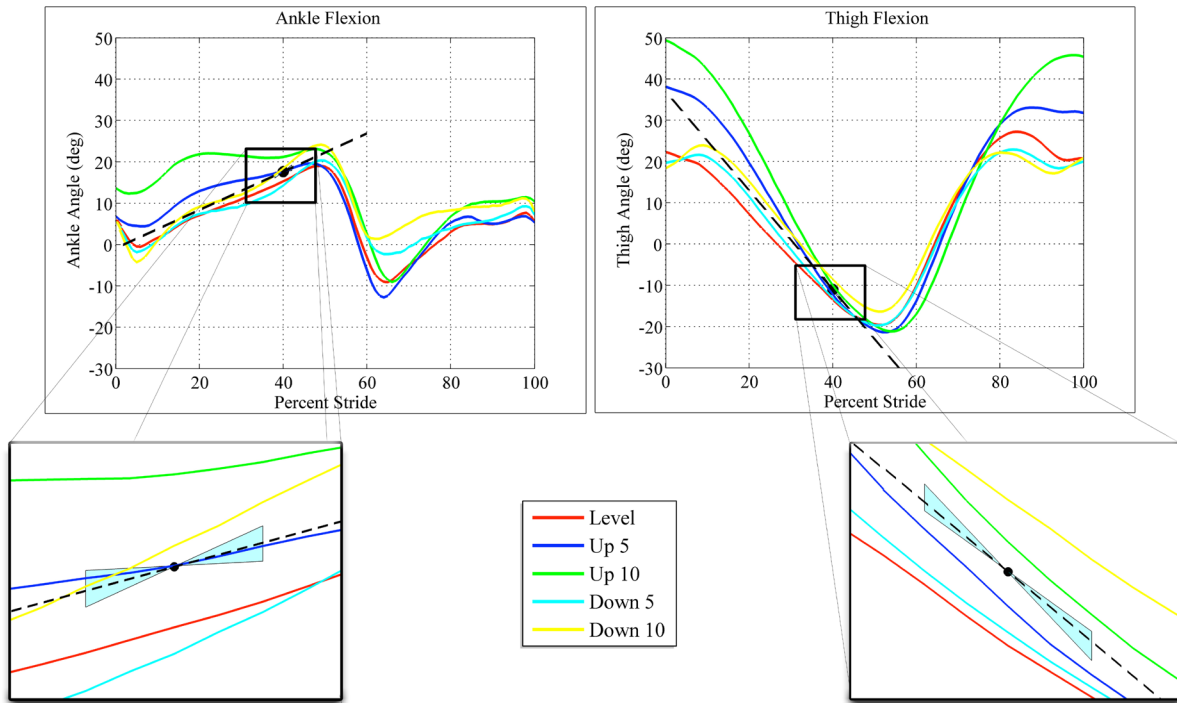


Figure 4: A comparison of the ankle and thigh flexion conditions with a focus on the derivative and derivative spread. The graph on the left (ankle) and the graph of the right (thigh) highlight the derivative and derivative spread at 40°.

The thigh angle exhibited a maximum slope of $-1.038^\circ / \% \text{ stride}$ and a minimum slope of $-1.521^\circ / \% \text{ stride}$. The magnitude of the spread of the derivatives was therefore $0.4830^\circ / \% \text{ stride}$. The percentage variation in the minimum and maximum derivative values was 68.24%. However, the ankle angle exhibited a maximum slope of $0.7826^\circ / \% \text{ stride}$ and a minimum slope of $0.08844^\circ / \% \text{ stride}$. For the ankle condition, the magnitude of the spread of the derivatives was therefore $0.6941^\circ / \% \text{ stride}$. The percentage variation in the ankle angle derivatives over different levels of slope was thus 884.8%. The level of variation in the ankle angle at 40% of stride was 6.910° (21.17° for a 10° incline and 14.26° for a -5° incline). Variation in the thigh angle was 5.018° (-13.55° for level ground and 8.532° for a -10° incline). The results among the three healthy test subjects indicate that the thigh angle condition has a greater derivative magnitude, less spread in the derivative, and less variation in angle over sloped ambulation.

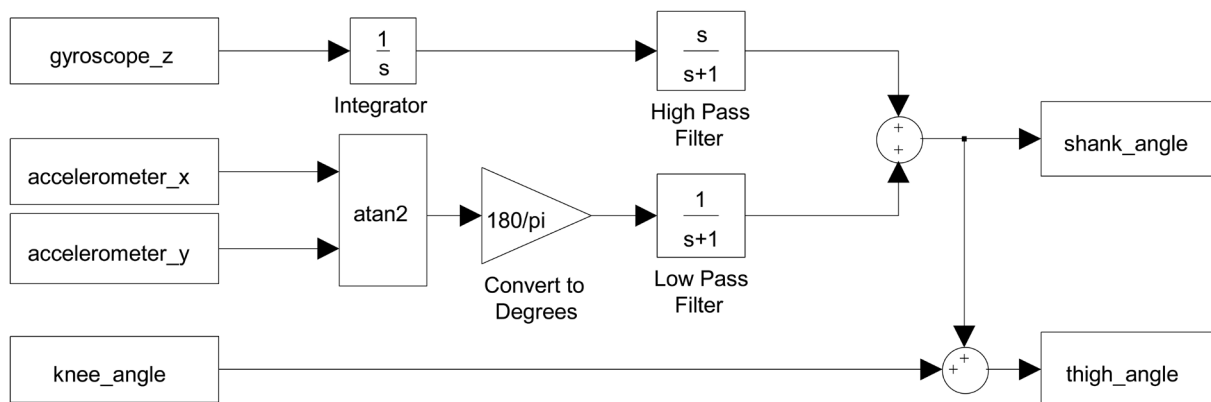


Figure 5: A schematic representing the construction of the thigh angle within the prosthesis

65 strides, there were no pathological outliers, indicating stride consistency. The minimum ankle angle, -26.05° , maximum ankle angle, 13.41° , minimum thigh angle, -24.59° , and maximum thigh angle 34.15° , all compared favorably with healthy subject data. Healthy subject data indicated a minimum ankle angle of -9.072° and a maximum ankle angle of 19.37° . Healthy data indicated a minimum thigh angle of -19.59° and a maximum thigh angle of 27.19° .

Over sloped walking, the amputee subject was instructed to walk at a self-selected gait speed using the thigh angle condition for push-off. Testing was only conducted for 5 and -5 degree inclines, as for safety reasons and practicality concerns (for accurate data to be collected, a ramp of significant length must be used) 10 and -10 degree inclines were not tested. At 5 degrees of incline, the test subject was able to perform walking that reasonably correlated with healthy walking biomechanics. A push-off angle of 10.5° was used. The prosthesis accurately replicated healthy human gait to an accuracy of approximately one standard deviation. Data, collected over 43 strides, were averaged to produce the final figures. Over these 43 strides, two were discarded due to erratic endpoint behavior. This indicates that the thigh angle condition was consistent over the middle 41 strides, reliably timing push-off at the instructed trigger angle. The minimum ankle angle, -19.88° , maximum ankle angle, 18.08° , minimum thigh angle, -29.13° , and maximum thigh angle 34.79° , all fell within the standard deviations of healthy data. Healthy data for 5 degrees of incline indicated a minimum thigh angle of -21.36° and a maximum thigh angle of

38.12° . Healthy ankle angle data at a 5 degree incline indicated a minimum of -12.78° and a maximum of 19.57° , closely paralleling that of the amputee's gait dynamics. At -5 degrees of incline, the amputee subject experienced difficulty replicating healthy walking biomechanics. A thigh angle push-off angle of 11.5° was used. Over 30 strides, 4 were discarded due to erratic behavior. For the ankle angle, the minimum and maximum angles over all strides collected were -28.36° and 17.14° , respectively. Healthy subject ankle data over a -5 degree incline demonstrated a minimum angle of -2.334° and a maximum angle of 20.33° . Healthy subject thigh angle data indicated a minimum angle of -19.71° and a maximum angle of 22.89° . Data for the -5 degree sloped walking suggested that push-off consistently occurred at the appropriate place within the stride; however, other parameters within the leg were not tuned for proper biomechanical function. The data suggest that the user received inadequate power at the ankle and knee for healthy walking, as there is significantly less plantarflexion in early-mid stance within the ankle, and displacement in the angular position of the thigh.

DISCUSSION

The thigh angle push-off condition proves to be effective in the powered transfemoral prosthesis. The thigh angle, exhibiting a greater magnitude of derivative and less spread in derivative than the ankle angle, provides a more accurate and more robust push-off condition. Additionally, at 40% of stride, there is noticeably less spread in the angular position of the thigh than in the angular position of the ankle. These improved metrics allow for increased sensitivity and robustness in the push-off condition, as detailed below.

The thigh angle's greater derivative magnitude, as compared with the ankle angle, suggests that the thigh angle would be a more effective trigger for push-off. As indicated by the results, the thigh angle has 2.590 times the derivative magnitude of the ankle angle. A higher derivative magnitude means an increased change in angle over percent stride. A shallower derivative would entail less change in the angular position for every percent stride, which increases the likelihood of a mistimed push-off. With the mean ankle angle derivative at $0.4627^\circ/\%$ stride, incorrect push-off timing is more probable than with the thigh angle. Given a tolerance of $\pm 1^\circ$, push-off from the ankle angle condition could come at any point within 4.322% of stride. The sensitive thigh angle, with a derivative of $1.200^\circ/\%$ stride, promises accurate and robust push-off timing. Given an identical tolerance, push-off could only come between -0.8333% and 0.8333% of 40% of stride. This reduction in push-off range allows for more accurate and less variant push-off timing. Additionally, less spread in the derivative over slope variations (Figure 4) suggests that the thigh angle would provide a more invariant condition for push-off. With less change in slope (68.24% difference between the minimum and maximum slope) at 40% of stride, the thigh angle would provide a more invariant condition for push-off. The ankle angle condition exhibits substantially larger variations in derivatives, with an 884.8% difference between the minimum and maximum derivative at push-off. This indicates that the thigh angle at different grades of walking shares highly similar behaviors, as compared with the more variant ankle angle push-off condition.

The thigh push-off angles for level-ground walking and sloped walking over 5 degrees and -5 degrees were -12° , -10.5° ,

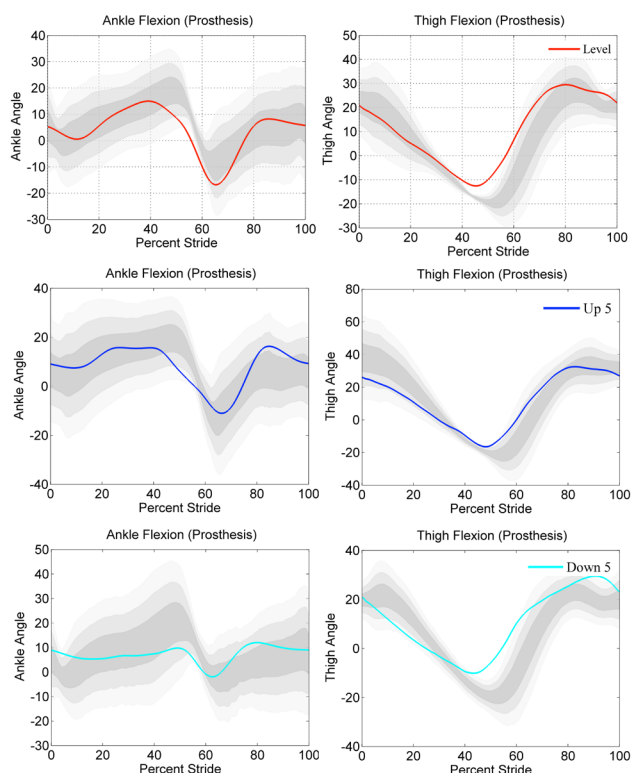


Figure 6: Ankle and thigh angles generated by the powered prosthesis under the revised thigh-angle push-off condition. Shaded regions represent 1, 2, and 3 standard deviations, respectively. Push-off condition presented the ability to push-off at the correct time for all strides within a trial. Additionally, the variation in angle between different gradations of walking was 1° .

and -11.5° . Thus, over three slope grades, there was only 1° of variation over 3 different levels of ambulation suggesting that there is minimized variation between push-off angles over different slope grades. The existing ankle angle condition, which provided a minimum push-off angle of 8° (level-ground) and a maximum push-off angle of 13° (5 degree incline), demonstrates a spread in push-off angle of 5 degrees, far from invariant. The decreased variation in push-off angle allows for a more streamlined controller, with potential for one angle to dictate push-off timing. A single push-off angle increases accuracy and safety of the device, as errors in ground slope estimation are less likely to cause erratic push-off behavior, which has potential to cause stumbling or falling in the amputee [23]. Thus, with less variance in push-off angle conditions, push-off is less likely to occur too early or too late in the gait cycle.

When implemented into the transfemoral prosthetic device, the thigh angle push-off condition falls within the standard deviation of healthy data (Figure 6), indicating the prosthetic device helps the user achieve proper gait. Over level-ground walking, the amputee experiences greater plantarflexion in the ankle at a slightly earlier time in the stride than a healthy subject. Furthermore, thigh data indicate reduced extension in mid-stride. The reference data is the mean of means (an average of a multitude of different subjects' gait), and each individual within that pool exhibits a distinct gait pattern that is still classified as healthy; in this case the unusual height of the subject (1.93m) gives the subject a unique – but not unhealthy – gait pattern.

These behaviors can be attributed to the habitual gait dynamics of amputees. When bending the knee during early stance (using a passive device), amputees tend to experience instability. Thus, as the amputee subject has experienced the adverse effects of flexion during early stance (Table 2) while using his passive prosthetic device, there is a behavioral bias toward less flexion. Over a slope of 5 degrees, both ankle and thigh data fall well within three standard deviations of healthy walking, with one standard deviation being the norm. Again, in the ankle data, plantarflexion of the ankle occurs earlier in the stride (and over a larger percentage of stride). Additionally, the ankle exhibits increased dorsiflexion at the end of stride. In the thigh, extension in mid-stride is reduced. These variations can be accounted for with more precise tuning of the power and torque applied to the joints during push-off. Over a slope of -5 degrees, ankle data generally fall within three standard deviations of healthy walking.

Thigh data follow the pattern of healthy gait, but fall short of healthy kinematics. Joint kinematics at the ankle lack healthy levels of plantarflexion in early stance and dorsiflexion in mid-stride. Dorsiflexion in late stride appears to be more exaggerated than healthy norms. In the thigh, flexion in early stance and extension, as with level-ground and 5 degree walking, is shallow, and comes around 10% too early in stride. There is also slightly increased flexion at the end of stride, as compared to healthy norms. These deviations do not indicate that the thigh angle push-off condition is lacking, as the data indicates that push-off occurs with the correct timing (as indicated by correct timing of ankle plantarflexion). Rather, the uncharacteristic gait is a result of poor tuning of other parameters in the prosthetic device. Time constraints and patient unavailability prevented extensive tuning, resulting in other non-ideal parameters. Notwithstanding, over -5 degrees, the data indicates that the subject did experience healthy gait, as push-off timing was correct, and actuation was provided

at the joints. Thus, data across level-ground, 5 degree, and -5 degree slope walking are indicative of robust push-off timing.

Transfemoral amputee subjects currently have limited control over their prosthetic device. The ankle angle is separated from the user – through the shank and knee - by several degrees, and thus is not within the intuitive control of the user. However, the thigh angle is in direct volitional control of the user, as the user still can actuate thigh muscles. In other words, the residual leg provides the most effective condition for control. By shifting the condition for push-off to a part of the body where the user is in direct control of actuation, it is possible to improve the function and safety of the prosthesis. Given the aforementioned benefits, this trigger not only improves function, sensitivity, and robustness, but also grants the user increased control.

CONCLUSION

This paper presents an alternative condition, the use of absolute limb orientation, to instigate push-off in a powered transfemoral prosthesis. The findings suggest that the thigh angle would provide a better condition for push-off than the ankle angle. The thigh angle not only provides the user with more intuitive control of the powered prosthesis, but also demonstrates increased sensitivity and robustness as a push-off condition. Additionally, the use of an absolute angle to initiate state-transitions was investigated, and the results indicate that absolute angles provide a promising new alternative for instigating state transitions.

Although the results are promising, further testing needs to be done over different slopes to confirm accuracy and robustness. Future study will include additional test subjects, 10 and -10 degree inclined walking, to determine efficacy across a greater spectrum of gait, and refinement of powers and torques applied at the knee and ankle joints. With other potential applications in transhumeral, transradial, and transtibial prostheses and lower limb exoskeletons, we will continue research in the use of absolute limb orientation for other prosthetic devices and systems.

With shark attacks, industrial accidents, medical amputations, and wars all still prevalent in our society today, it is no wonder that the number of amputees continues to rise. Surprisingly, there are over 30,000 above-the-knee amputations done yearly. These amputees currently have only one choice: an unpowered prosthetic leg to aid them walk. However, this unpowered device does not allow the amputee to walk normally – they often have to limp or exert an excessive amount of force on other joints in their body to compensate for their missing limb. In a powered device, this all changes. This research will help amputee subjects by advancing the control interface of a powered prosthesis; this device has potential, through novel actuation and control mechanisms to more closely mimic healthy walking. With this device, amputees are given a second chance.

ACKNOWLEDGMENTS

Special thanks to Brian Lawson, Ph.D and Michael Goldfarb, Ph.D for their dedicated commitment to enriching my experience in and expanding my knowledge of engineering. Funding for this project was provided by the National Institutes of Health.

REFERENCES

1. J. Feinglass, J. L. Brown, A. LoSasso, M. W. Sohn, L. M. Manheim, S. J. Shah, and W. H. Pearce, "Rates of lower-extremity amputation and arterial

- reconstruction in the United States, 1979 to 1996.," Am J Public Health, vol. 89, no. 8, pp. 1222–1227, Aug. 1999.
2. K. Ziegler-Graham, E. J. MacKenzie, and P. L. Ephraim, "Estimating the Prevalence of Limb Loss in the United States: 2005 to 2050," Archives of physical ... , 2008.
 3. M. Goldfarb, B. E. Lawson, and A. H. Shultz, "Realizing the Promise of Robotic Leg Prostheses," Science Translational Medicine, vol. 5, no. 210, pp. 210ps15–210ps15, Nov. 2013.
 4. R. Walters, J. Perry, D. Antonelli, and H. Hislop, "Energy cost of walking of amputees: the influence of level of amputation," The Journal of Bone and Joint Surgery, vol. 58, no. 1, pp. 42–46, Apr. 1976.
 5. D. A. Winter, The biomechanics and motor control of human gait. Waterloo Biomechanics, 1991.
 6. B. E. Lawson, "A Multi-Purpose Finite State-Based Standing Controller for a Powered Transfemoral Prosthesis," Dec. 2011.
 7. S. K. Au, J. Weber, and H. Herr, "Powered Ankle--Foot Prosthesis Improves Walking Metabolic Economy," Robotics, IEEE Transactions on, vol. 25, no. 1, pp. 51–66, 2009.
 8. F. Sup, H. A. Varol, J. Mitchell, T. J. Withrow, and M. Goldfarb, "Preliminary Evaluations of a Self-Contained Anthropomorphic Transfemoral Prosthesis," Mechatronics, IEEE/ASME Transactions on, vol. 14, no. 6, pp. 667–676, 2009.
 9. F. Sup, H. A. Varol, and M. Goldfarb, "Upslope Walking With a Powered Knee and Ankle Prosthesis: Initial Results With an Amputee Subject," Neural Systems and Rehabilitation Engineering, IEEE Transactions on, vol. 19, no. 1, pp. 71–78, 2011.
 10. B. E. Lawson and H. A. Varol, "Standing stability enhancement with an intelligent powered transfemoral prosthesis," ... , 2011.
 11. E. C. Martinez-Villalpando, L. Mooney, G. Elliott, and H. Herr, "Antagonistic active knee prosthesis. A metabolic cost of walking comparison with a variable-damping prosthetic knee," presented at the Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE, 2011, pp. 8519–8522.
 12. W. C. Miller, M. Speechley, and B. Deathe, "The prevalence and risk factors of falling and fear of falling among lower extremity amputees," Archives of Physical Medicine and Rehabilitation, vol. 82, no. 8, pp. 1031–1037, Aug. 2001.
 13. H. A. Varol, F. Sup, and M. Goldfarb, "Multiclass Real-Time Intent Recognition of a Powered Lower Limb Prosthesis," Biomedical Engineering, IEEE Transactions on, vol. 57, no. 3, pp. 542–551, 2010.
 14. Fan Zhang, Zhi Dou, M. Nunnery, and He Huang, "Real-time implementation of an intent recognition system for artificial legs," presented at the 2011 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2011, pp. 2997–3000.
 15. R. D. Bellman, M. A. Holgate, and T. G. Sugar, "SPARKy 3: Design of an active robotic ankle prosthesis with two actuated degrees of freedom using regenerative kinetics," presented at the Biomedical Robotics and Biomechatronics, 2008. BioRob 2008. 2nd IEEE RAS & EMBS International Conference on, 2008, pp. 511–516.
 16. B. Lawson, H. A. Varol, A. Huff, E. Erdemir, and M. Goldfarb, "Control of Stair Ascent and Descent With a Powered Transfemoral Prosthesis," IEEE Trans. Neural Syst. Rehabil. Eng. , vol. 21, no. 3, pp. 466–473, 2013.
 17. A. S. McIntosh, K. T. Beatty, L. N. Dwan, and D. R. Vickers, "Gait dynamics on an inclined walkway," Journal of Biomechanics, vol. 39, no. 13, pp. 2491–2502, Jan. 2006.
 18. A. H. Vrieling, H. G. Van Keeken, T. Schoppen, E. Otten, J. P. K. Halbertsma, A. L. Hof, and K. Postema, "Uphill and downhill walking in unilateral lower limb amputees," Gait & Posture, vol. 28, no. 2, pp. 235–242, Aug. 2008.
 19. A. Leroux, J. Fung, and H. Barbeau, "Postural adaptation to walking on inclined surfaces: I. Normal strategies," Gait & Posture, vol. 15, no. 1, pp. 64–74, Feb. 2002.
 20. A. N. Lay, C. J. Hass, and R. J. Gregor, "The effects of sloped surfaces on locomotion: A kinematic and kinetic analysis," Journal of Biomechanics, vol. 39, no. 9, pp. 1621–1628, Jan. 2006.
 21. M. P. Kadaba, H. K. Ramakrishnan, and M. E. Wootten, "Measurement of lower extremity kinematics during level walking," J. Orthop. Res. , vol. 8, no. 3, pp. 383–392, May 1990.
 22. B. E. Lawson, A. Huff, and M. Goldfarb, "A preliminary investigation of powered prostheses for improved walking biomechanics in bilateral transfemoral amputees," presented at the 2012 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2012, pp. 4164–4167.
 23. B. E. Lawson, H. A. Varol, F. Sup, and M. Goldfarb, "Stumble detection and classification for an intelligent transfemoral prosthesis," presented at the 2010 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2010), 2010, pp. 511–514.



MAURICE CHIANG

Maurice Chiang is a freshman at Stanford University interested in bioengineering. Throughout high school, he was heavily involved in scientific research, investigating the mechanics of human motion, second hand smoking patterns among Chinese populations, and certain aspects of molecular biology. At Stanford, Maurice has been active in Stanford Students in Applied Research, ASES, and other entrepreneurial efforts on campus. Additionally, he remains active in his artistic endeavors, pursuing and creating installations and sculptural works. Looking towards the future, he intends to continue doing research, and hopes to ultimately bring novel discoveries in the lab to market – either in biotechnology or pharmaceuticals – with the goal of improving quality of life.

Interstitial Metal Hydride Alloys: A Systematic Literature Review of Contemporary Hydrogen Storage Applications

Emily E. Petersen

Michigan Technological University

Interstitial metal hydrides have garnered significant attention in recent years due to their promise in facilitating the transition from fossil fuels to cleaner, renewable, and more environmentally responsible energy sources. The high volumetric density of hydrogen in the host lattice, relative safety and cost efficiency compared to alternative sources, reversibility, and tailorability have driven interest in the greater research and industrial communities. Articles were selected for review based on relevance, date of publication (<15 years), and accessibility. Electrochemical applications including nickel-metal hydride batteries and lithium-ion batteries demonstrated the greatest degree of research focus. Other contemporary applications included fuel and neutron shields in compact nuclear power modules, alternative storage materials for thermal energy systems, and hydrogen purification. Independent of application, a more firm understanding of fundamentals is still necessary in order to take full advantage of operation parameters such as volumetric and gravimetric storage density, rate of de/hydrogenation, reversibility, and safety.

ENVIRONMENTAL MOTIVE

In a 2014 study performed by the U.S. Energy Information Administration (EIA), it was estimated that only 11% of world marketed energy consumption was generated from renewable energy sources [1]. These include biofuels, biomass, hydropower, solar, and wind sources. Fossil fuels are currently the world's primary energy source; however they can do irreparable damage to the environment. The burning of these finite resources was found to be responsible for 79% of U.S. greenhouse gas emission in 2010 according to the Environmental Protection Agency [2]. Oil, the world's primary fuel source for transportation, poses significant environmental concerns due to oil spills, extraction, and the release of fine particulates during combustion. Heavier crude oils necessitate the use of even more energy intensive methods. Approximately 39% of the United States power supply is generated by coal, the combustion of which releases harmful air pollutants. Mining further damages the environment by destroying top-soil and vegetation. While natural gas burns cleaner than either coal or oil, it is nonetheless responsible for nearly 30% of greenhouse gas emissions in the United States. A 2012 study revealed that the United States contributes to 18% of the world's total primary energy consumption [3]. Despite the country's dependence on fossil fuels for electricity, heat, and transportation fuel, alternative fuel sources do exist and will prove critical in the transition to cleaner, more environmentally responsible energy use.

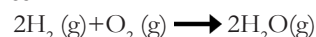
LITERATURE SELECTION

Literature was selected for review based on relevance, date of publication, and availability. Involvement by the greater research

community has remained relatively contained to a few industries therefore the review is organized based on application. Because hydrogen is a relatively new area of interest for energy storage applications, publications were primarily drawn from the last five years, however relevant literature dating back to 2001 was referenced. Finally, the transition to more sustainable energy sources will ultimately rely on collaboration between academia and industry. It is the responsibility of each individual to educate themselves on sustainability and to take steps toward a cleaner future for the environment. In this spirit, the review was conducted entirely with open-source literature. All references are readily available and easily accessed.

HYDROGEN AS AN ENERGY SOURCE

While hydrogen was first identified as an element in 1766 by British scientist Henry Cavendish, it was not until the mid-1800s that it was recognized as a potential fuel source [4]. The most basic of Earth's elements, hydrogen is composed of a single proton and single electron. Energy is required to separate the gas from companion substances, however the result is a powerful, clean energy source.



Not only is hydrogen the most abundant element in the universe, there are no harmful byproducts when burned. It is also one of the most efficient fuel sources, producing more energy per pound of fuel than many traditional fuels and therefore allowing for extended application [5].

Storage has proven to be a challenge in enabling hydrogen as a common alternative energy source; however it will be crucial in enabling the technology to advance. The objective is ultimately to reach the highest possible volumetric density using as little

additional material as possible. One kilogram of hydrogen gas has a volume of 11m³ at ambient temperature and atmospheric pressure [5]. Increasing the gas density requires work to compress it which decreases the temperature below the critical temperature or reduces the repulsion between hydrogen and the secondary material. Another important factor in hydrogen storage is the reversibility of the system.

Hydrogen storage is typically classified as either physical or chemical. Physical storage is characterized by the storage of hydrogen molecules whereas hydrides are a form of chemical storage. Hydrogen stored physically adopts either gas or liquid form. Figure 1 shows a primitive phase diagram for hydrogen,

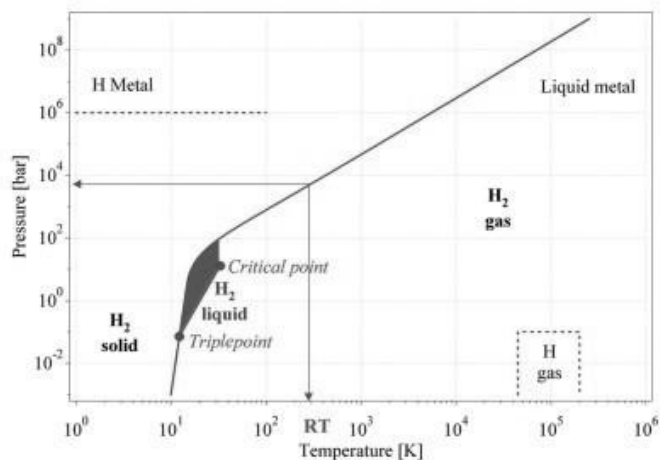


Figure 1: This primitive phase diagram from Materials for Hydrogen Storage in the September 2003 issue of Materials Today shows where liquid hydrogen exists in relation to pressure and temperature conditions. Hydrogen’s unique phase diagram demonstrates the element’s specific pressure and temperature conditions in order to be effective in energy storage applications [5].

clearly indicating the small window where pressure and temperature conditions allow for liquid hydrogen to exist. In gas form, high-pressure tanks between 5,000 and 10,000psi are required [6]. In liquid form, cryogenic temperatures are required due to hydrogen’s low boiling point. Liquefaction sacrifices a large amount of energy, and the insulation needed in the fuel tanks contributes to the overall cost, making it a less preferable option.

In the case of chemical storage, hydrides are hydrogen anions bonded to more electropositive elements or groups. In this manner, hydrogen can also be stored on the surface of or within solid materials. Metal hydrides are one such storage medium. Metal hydrides are preferable to compressed or liquefied hydrogen storage methods due to its relatively safer operation as there is no high pressure containment or boil-off. In this method, hydrogen molecules are stored in materials by physisorption [7]. This phenomena is a property of the primary material, and the hydrogen capacity is proportional to its specific surface area. While most metal hydrides require high temperature to release their hydrogen content due to the strong bonding, pairing weak and strong hydrides reduces the overall energy cost. An alternative method for reducing dissociation temperatures is doping with activators.

Many metal hydride compounds deviate significantly from ideal stoichiometry and therefore can exist as multiphase systems. Typically they exist as ratios of A elements, which are usually rare earth or alkaline earth metals, and B elements, which are often transition metals. Absorption is the process in which hydrogen gas reacts with a metal. Pressure-temperature-composition curves are used to effectively describe the thermodynamics of hydride formation as shown in Figure 2 from the September 2003 article Materials for Hydrogen Storage in Materials Today. The length of the equilibrium pressure plateau determines the amount of hydrogen stored. This pressure is related to changes in entropy and enthalpy as a function of temperature as dictated by the Van’t Hoff equation.

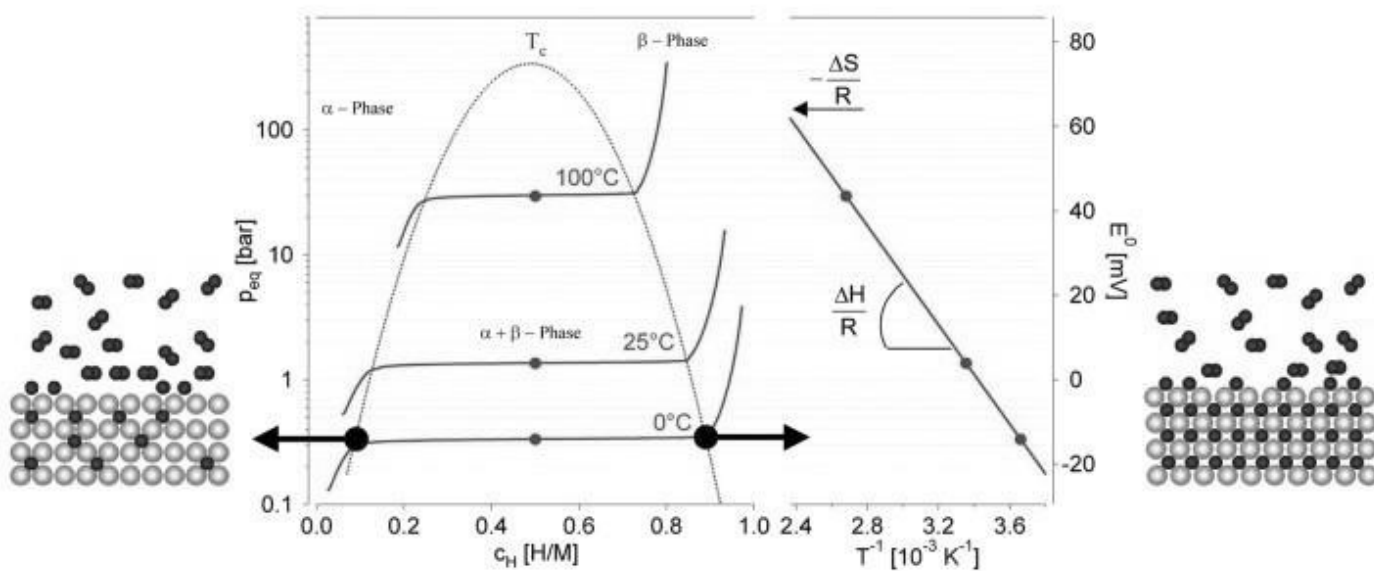


Figure 2: Pressure-temperature-composition curves provide a visual of the thermodynamic aspects of hydride formation. This diagram serves as a representative curve for hydrogen absorption in an intermetallic compound. The Van’t Hoff plot is shown on the right-hand side of the figure, and the slope is the enthalpy of formation divided by the gas constant. The absorption and desorption processes are shown in intermetallic compounds through the sphere diagrams to either side of the graphs [5].

INVOLVEMENT BY THE GREATER RESEARCH COMMUNITY

Metal hydride applications are currently confined to small number of institutions and companies as summarized in Table 1 as found in the December 2007 article Theoretical and Experimental Studies of Metal Hydride Storage Units published by the Norwegian University of Science and Technology. The criticality of exploring hydrogen as an alternative fuel source coupled with the paucity of players in the field highlight the need for greater involvement by the research and industrial community.

NUCLEAR POWER APPLICATIONS

Hydrogen storage spans a wide range of applications. One of the newer applications is the hydrogen-moderated self-regulating nuclear power module. Hydrogen storage in this form exists as a neutron moderator which serves to reduce the speed of neutrons, thereby transforming them into thermal neutrons capable of facilitating a nuclear chain reaction. The fuel and moderator are uranium hydride which is reduced to uranium and hydrogen at elevated temperatures. Hydrogen in gas form exits the reactor core and is absorbed by a secondary hydrogen-absorbing material such as depleted uranium [8]. The Hyperion Power Module, a self-contained, self-regulating reactor model developed at Los Alamos National Laboratory, operates at an optimum temperature of 550°C, allowing for relative ease of transportation without significant high-pressure risk. Figure 3 shows a cut-out view of the generator, clearly labeling the hydride fuel and moderator placement in the system.

Staying in the realm of nuclear engineering, another contemporary application of metal hydrides in energy storage is the neutron shield in generators through the use of compact reactors. Reduction of nuclear heating arising from neutron and gamma interactions is crucial in maintaining the toroidal field coils in a superconductive state. Shortening the distance between the plasma and the inner leg of the coils aids in achieving high power density, necessitating the development of effective and compact neutron shields. Zirconium borohydride and zirconium hydride are candidate materials that both show promising shielding and hydrogen storage capabilities [10]. Studies revealed that high atomic number elements as well as hydrogen atoms are effective in neutron shielding. Additionally, combinations of steel and $Zr(BH_4)_4$ improves shielding capability. Combinations of $(Zr(BH_4)_4 + F_8 2H)$ and $(ZrH_2 + F_8 2H)$ reduce shield thickness by 6.5% and 19% compared to $(water + F_8 2H)$ respectively.

ELECTROCHEMICAL APPLICATIONS

Electrochemical applications are another area of recent advancement for hydrogen storage. Metal-hydrogen bonds form by the electrons accompanying the hydrogen atom entering the lattice. Electrons maintain local charge neutrality by remaining within 3 to 10 angstroms of the protons [11]. During charging, electrons enter the metal through the metal/metal hydride electrode current collector, effectively neutralizing the protons produced at the metal/electrolyte interface. The charging process is determined by the half reaction below:



During discharge, metal hydride protons leave the surface to

Company / institution (country)	Main activities
Ergenics, Inc. (USA)	Hydride Alloys MH-Heat Pumps MH-Heat Storage
Hydrogen Components, Inc. (USA)	MH-H storage units MH-H compressors
ECD Ovonic (USA)	Ni-MH batteries MH alloys and H storage systems
The Japan Steel Works, Ltd. (Japan)	Hydride alloys MH-H storage units Ni-MH batteries Hydrogen fuel cells MH-Heat storage, conversion and transportation MH-Actuators
Japan Metals & Chemicals Co., Ltd. (Japan)	Hydride alloys MH-H storage units MH-heat pumps
HERA Hydrogen Storage Systems, GmbH (Canada - Germany)	MH-H storage units Hydride Alloys
LABTECH Int. Co. Ltd. (Bulgaria)	MH-H storage units Hydride Alloys
Institute of Mechanical Engineering Problems (Ukraine)	MH-H storage units MH-H compressors MH-based evacuation/gas supply systems for vacuum-plasma technologies

Table 1: Hydride alloys, pumps, and storage are prime areas of investigation by a small number of private companies.

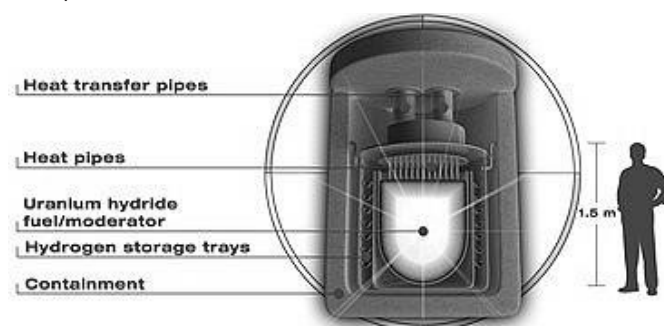
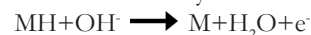


Figure 3: The Hyperion Power Module employs uranium hydride as the fuel and moderator. This unique storage method allows for easy transportation and safe operation as the reactor is self-regulating. As any increase in temperature immediately reduces the criticality parameters of the system, the Hyperion operates safely in a steady state [9].

form water. Charge neutrality pushes electrons out through the current collector, performing work in the attached system. The discharging process is determined by the half reaction below:



The choice of metal hydride determines the standard potential of the half reactions, and with the exception of zirconium batteries, this potential is designed to be as low as possible in order to

maximize the amount of energy stored.

Metal hydrides can also be used in lithium-ion batteries as the anode material. These batteries are preferred over NiMH batteries for their lighter weight and more rapid charge and discharge [13]. Their ability to provide portable energy has made them particularly popular in the automotive and portable electronic industries [14]. Research performed in 2015 demonstrated MgH_2 as one of the most attractive metal hydrides for this application with a reversible capacity of 1480mAhg^{-1} and the lowest electrode polarization of less than 0.2V [15]. Other metal and complex hydrides such as TiH_2 and AlH_3 are in the process of being investigated. Nanocomposite synthesis research is closely tied to this field as the limited electrochemical cycling performance has proven to be the primary drawback of metal hydrides in battery applications. Overall, current research has highlighted the importance of understanding hydride fundamentals in order to optimize the conversion complex process. More effective and involved collaboration between the greater battery and fuel-cell research communities will also be critical in advancing lithium-ion battery technology.

SOLAR POWER APPLICATIONS

Solar energy has taken its place as a major player in the effort to curb fossil fuel use and transition into cleaner energy sources. Currently, high temperature solar plants have the ability to achieve low cost electric power production if equipped with thermal energy storage systems. These systems typically store heat through sensible heat, latent heat, or thermochemical heat [16]. Research has investigated storing sensible heat in the form of molten salts, however sensible heat materials sacrifice energy density and have limited transport distance due to heat loss. Alternatively, latent heat materials, while of higher energy density and smaller temperature range for heat transfer, suffer higher cost and greater corrosive effects. Thermochemical heat storage systems have both very high energy density and low cost, and hydrogenation of metal hydrides provides the additional advantage of being able to reversibly hydrogenate and dehydrogenate by a simple change in pressure at a constant temperature. Metal hydrides offer energy densities 15-20 times higher than molten salt systems, thus lowering the size and cost of many solar power systems [17]. Studies performed in 2015 investigated metal hydride pairs including low-temperature NaAlH_4 and Na_3AlH_6 as well as higher temperature hydrides based on NaMg materials [18]. It was found that the addition of aluminum and expanded natural graphite enhance cycle stability of NaAlH_4 . Furthermore, an economic analysis was performed in the evaluation of two metal hydride pairs. Both pairs (NaMgH_3 ; NaAlH_4 and NaMgH_2F ; Na_3AlH_6) showed promise as low cost, high efficiency target systems.

Fuel cells require purification of hydrogen for operation, offering another exciting application for metal hydrides. In these systems, a metal hydride alloy reacts irreversibly with impurities in the hydrogen gas. These impurities include carbon monoxide, carbon dioxide, methane, water, oxygen and nitrogen, and they affect systems by destroying hydriding ability, decaying the system, and decreasing efficiency, among other damages. Research in this area has sought to investigate the tolerance of metal hydrides for impurities with the objective of minimizing impurity deposition on storage surfaces and improving cycling performance. Due to the absence of funding for this literature review, many current

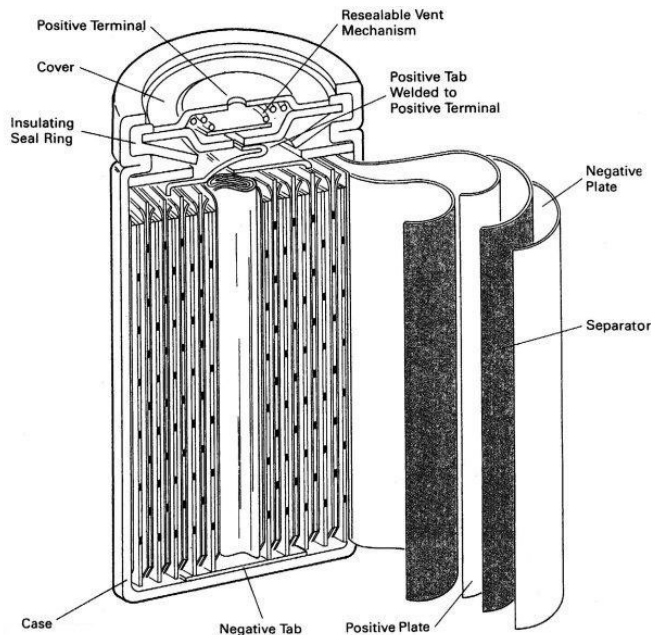


Figure 4: Nickel-metal hydride batteries are available in standard cylindrical sizes as shown in the figure. While similar to nickel-cadmium batteries, NiMH batteries demonstrate higher energy storage capacity as well as minimal memory effect. This effect causes batteries to gradually lose energy capacity through repeated recharge when the battery is not fully discharged. The battery “remembers” the new decreased capacity.

[Image obtained from support.radioshack.com]

publications were not able to be accessed for discussion; therefore contemporary details in this area are not within the scope of this paper.

CARBON ABSORBENTS

Yet another area of recent investigation into hydrogen storage is the use of carbon absorbents. Numerous studies have explored materials such as graphite nanofibers and single and multi-walled carbon nanotubes, however synthetic limitations have spurred the development of materials specifically designed for carbon absorption [19]. Carbon aerogels have been subject to recent research due to their low density and appropriate porosity. A 2012 investigation of metal hydride/carbon aerogel composites revealed that the aerogel was effective in improving hydrogen storage capacity of iron and titanium doped MgH_2 samples. Furthermore, metallic dispersion onto the aerogel showed promise in improving absorption [20]. This area of research interest is yet relatively new, however the currently published data demonstrates great promise for growth.

CONCLUSION

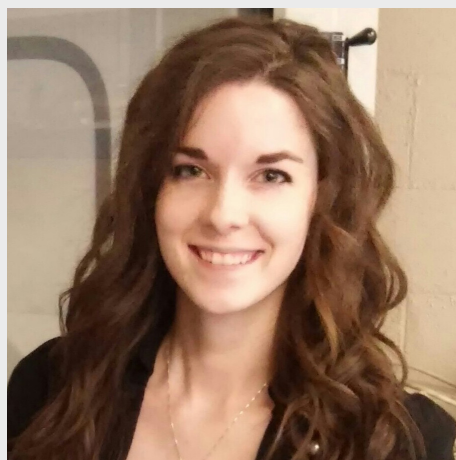
A transition from conventional energy sources such as oil, coal, and natural gas involves a number of uncertainties, however a firm understanding of alternative energy sources is in development and being established through industry and academic alike. A hydrogen economy could provide resource security, economic benefits, and establish precedent for greater environmental responsibility; interstitial metal hydrides have shown great promise as a means

by which to establish hydrogen as a superior alternative energy source. Safe operation, high volumetric and gravimetric hydrogen density, reversibility, ease of transportation, and low raw material cost are recurring themes identified by contemporary research in the development of effective hydrogen storage systems employing metal hydrides.

Significantly more research is still needed, particularly in establishing a fundamental understanding of metal hydrides, their interaction with other compounds and materials matrices, and their properties under varying conditions. The majority of interstitial metal hydride hydrogen storage research has fallen in the realm of electrochemical applications, however optimization of operation parameters is still an area of active focus. Less understood industrial applications including thermal storage systems employed by solar energy plants and neutron shielding systems for nuclear power modules will require even more extensive research and exploration of the full spectrum of potential alloy systems. Nonetheless, the research already performed has shown great promise and is deserving of continued interest and involvement by the greater research and industry communities.

REFERENCES

1. "How much of world energy consumption and electricity generation is from renewable energy?" U.S. Energy Information Administration. Retrieved from <https://www.eia.gov/tools/faqs/faq.cfm?id=527&t=1>
2. "Fossil Fuels." Environmental and Energy Study Institute. Retrieved from <http://www.eesi.org/topics/fossil-fuels/description>
3. "What is the United States' share of world energy consumption?" U.S. Energy Information Administration. Retrieved from <https://www.eia.gov/tools/faqs/faq.cfm?id=87&t=1>
4. "Hydrogen Fact Sheet: History of Hydrogen." New York State Energy Research and Development Authority. Retrieved from <http://www.fuelcelleducation.org/wp-content/themes/sandbox/pdf/History%20of%20Hydrogen.pdf>
5. Zuttel A. Materials for hydrogen storage. *Materials Today* 2003; 1016/ S1369-7021(03)00922-2
6. "Hydrogen Storage." Office of Energy Efficiency and Renewable Energy. Retrieved from <http://energy.gov/eere/fuelcells/hydrogen-storage>
7. Krishna R, Titus E, Salimian M, et al. Hydrogen Storage for Energy Application. *Hydrogen Storage for Energy Application*. 10.5772/51238. 2012.
8. Peterson O. High hopes for hydride. *Nuclear Engineering International* 2009.
9. Hyperion Power Generation. Ganino. Retrieved from http://www.ganino.com/hyperion_power_generation_accumulators
10. Hayashi T, Tobita K, Nakamori Y, et al. Advanced neutron shielding material using zirconium borohydride and zirconium hydride. *Journal of Nuclear Materials* 2009.
11. Young K, Nei J. The current status of hydrogen storage development for electrochemical applications. *Materials* 2013; 10.3390/ma6104574.
12. Hong K. The development of hydrogen storage electrode alloys for nickel hydride batteries. *Journal of Power Sources* 2001.
13. "How can lithium-ion batteries improve hybrids?" How Stuff Works: Auto. Retrieved from <http://auto.howstuffworks.com/lithium-ion-batteries-improve-hybrids1.htm>
14. Wakihara, M. Recent developments in lithium ion batteries. *Materials Science and Engineering* 2001.
15. Aymard L, Oumellal Y, Bonnet J. Metal hydrides: an innovative and challenging conversion reaction anode for lithium-ion batteries. *Beilstein Journal of Nanotechnology* 2015.
16. Reinhardt B. Thermal Energy Storage. Stanford University, Physics 240. Fall 2010. Retrieved from <http://large.stanford.edu/courses/2010/ph240/reinhardt1/>
17. Harries D, et al. Concentrating solar thermal heat storage using metal hydrides. *Proceedings of the IEEE* 2012. 10.1109/JPROC.2011.2158509.
18. Ward P, Corgnale C, Teprovich J, et al. High performance metal hydride based thermal energy storage systems for concentration solar power applications. *Journal of Alloys and Compounds* 2015.
19. Dillon A, Heben M. Hydrogen storage using carbon adsorbents: past, present and future. *Applied Physics A Materials Science and Processing*. 10.1007/s003390100788. 2011
20. Lin K, Mai Y, Chiu S, et al. Synthesis and characterization of metal hydride/carbon aerogel composites for hydrogen storage. *Journal of Nanomaterials* 2012.



EMILY PETERSEN

Emily Petersen will graduate in spring 2017 with a BS in Materials Science & Engineering from Michigan Technological University. Upon graduation she intends to pursue a graduate degree before starting a career in the aerospace industry, a passion inspired by her internships at NASA Langley Research Center and Lockheed Martin. She published with and presented at the 2014 and 2015 Materials Science & Technology Conferences, spoke at the 2015 National Conference on Undergraduate Research, and expects to publish again in the coming months. Outside of academics, she pursues interests in dance, languages, and community service.

CROSSING THE RUBICON INTO CLASSICS RESEARCH

DEAN RICHARD SALLER & ROMAN PATRONAGE AND ECONOMICS

by Kyle Lin

Richard Saller will always remember the reception his early research received at Harvard. Newly arrived to the U.S. from Cambridge, Saller had been invited to speak about Roman history before an audience of Harvard faculty. After his lecture, Saller was standing in the elevator when a chairman next to him spoke up.

“Well, we’ve never heard anything like that before,” the Harvard chair said. He didn’t mean it in a good way.

But it was Saller and his peers who were ahead of the times in ancient history. “I don’t think [the chairman] had any idea what I was talking about,” Saller remarked.

The path that Richard P. Saller, Dean of the Humanities and Sciences, has taken in his pursuit of history has been anything but conventional. Originally an engineering major, he stumbled upon ancient history in an introductory class on Rome. An undergraduate student during the Vietnam War, Saller found himself drawn to the study of Roman imperialism as a lens for understanding America’s own projection of military force worldwide. The die was cast: Saller began studying classics. In retrospect, Saller realized his decision to switch majors was “ill-considered.”

“I don’t think I ever really seriously thought about whether I would have a job doing this,” he said, “but I focused on it and continued.”

A few years later at Cambridge, Saller first met Moses Finley, a renowned historian and well-versed economist. By Saller’s account, “he was the central figure in British classics, who took classics from being as kind of elite-schoolboy exercise...to what I would regard as a more serious historical study of the past.”

To Saller, a Cambridge PhD student, Finley was “intimidating.” Finley was also just the person he wanted for a mentor. Saller vaguely recalled their first interaction at a drop-in meeting. “I remember the office-hour visit being pretty brusque, matter-of-fact, to the point – and I was out of the office,” he said. “He wasn’t going to sit and chat with me.”

Over time, however, the two scholars warmed to each other. Ultimately, Finley’s example – his bold presentation of ideas, and his ability to see implications in the past for the present – would have a great influence on Saller. During his time at Cambridge, Saller became determined to pursue “questions that [were] interesting, that nobody’s bothered to ask before.”

For Saller, these questions would involve the

social sciences: how society and economics functioned for the Romans. Socioscientific issues were abnormal topics to pursue in 1980’s classics academia. But Saller forewent a traditional research focus to pursue topics he found more novel and intriguing. “In reinterpreting Thucydides for the hundredth time,” he concluded, “I was unlikely to have anything original to say.”

Finding information on ancient economics and patronage was difficult, especially because social sciences did not even exist as concepts 2000 years ago. “No Roman wrote a work on social history,” Saller said. “They didn’t even have a concept of economic history as such.”

Saller learned to scour vast amounts of resources, to find each rare piece of relevant material: “often just paragraphs, or maybe a letter of Pliny.” He used computer simulation technology to acquire data, employing digital humanities “before the phrase was ever invented.”

Saller also studied language use. He combed through texts to locate every usage of certain Latin keywords related to the topic at hand. When researching Roman patronage, he focused on words like patronus (patron), beneficium (favor), and gratia (gratitude). Peter Garnsey, an advisor at Cambridge, told him, “You need to start from understanding the vocabulary that the ancient authors use.” Saller did just that.

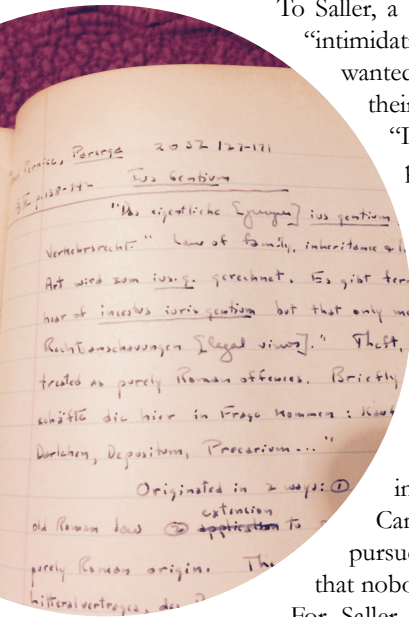
“Because I started Latin and Greek much later in life, I didn’t take it for granted that I knew what these words meant,” he said. “And as a result of this systematic study, I was able to...propel an argument that was different from what people had said before.”

Saller and his peers found themselves “close to the leading edge” of ancient historical research. “And then I think [we] became sort-of standard-bearers” – the aquiliferi – “for opening up the field of family history.”

After leaving Cambridge and forging a unique line of research, Saller has since departed from the aggressive, polemical style and methods of his mentor Finley. However, he does not forget the “huge” impact Finley had on his career. In fact, it was Finley who reassured Saller’s wife about his prospects as a historian: “He apparently told her that she shouldn’t worry about me, because when he was a young scholar my age, he was even brasher...than I was.”

Ultimately, Finley left Saller with something more than his example and advice. Shortly before his death in 1986, Finley gave Saller a keepsake: an old bound notebook, containing Finley’s own Roman law notes from school at Columbia University. The notes have no “biographical or intellectual value,” but Saller still treasures them.

“In the end, I was very touched. I’m not quite sure why he did this,” Saller said. “I still have that [notebook], and I prize that...At some point I need to deposit it in his collection at Cambridge, but I haven’t been willing to give it up.”



humanities

Sociology of Joyce's Women in *A Portrait of the Artist as a Young Man* and the Irish Parallel

Mariam Sulakian

Duke University

In his novel, *A Portrait of the Artist as a Young Man*, James Joyce traces the traditional roles of women as mothers and sexually objectified beings. Particular evaluation of the Christmas dinner party scene and the Ballyhoura Hills woman encounter provides evidence of women continuously fighting their conventional, subordinate positions to men. Yet, there seems to be apparent confusion between the expectations of silence and self-expression, as demonstrated by the dual nature that women take on in the novel. Joyce further exemplifies the estrangement of his female characters from social norms by paralleling their behavior to his own sentiments toward Mother Ireland, which he and his novel's protagonist Stephen both come to resent and leave. With the subtlety of biblical references and the characters' duplicitous behaviors, *Portrait* paints a confounded dynamic between the desire for personal liberation and institutional incarceration within both the novel's women and Joyce himself.

Humanities
Oscar Wilde once wrote, "A woman's life revolves in curves of emotions. It is upon lines of intellect that a man's life progresses" ("An Ideal Husband" 4.246). Similarly, the early 20th century Irish novelist James Joyce often depicts the estrangement of men's and women's spheres. *The Portrait of the Artist as a Young Man* (1916), Joyce's first novel, traces the maturation of a young Stephen Dedalus who gradually rejects social, familial, and religious pressures and instead dedicates himself to a career of writing. Set in the final years of nineteenth century Ireland, the literary work emerged in the midst of the female suffrage movement, during which time infant mortality rates lowered and female citizens increasingly sought paid employment; it should be noted, however, that Catholic women were often more limited in job opportunities than those in the Protestant Ascendancy due to the Catholic church's strict employment restrictions (Wainwright 654). Heated politics, with nationalist political leader Charles Parnell at the forefront, also had radical effects on the women's movement in the late 1800s and early 1900s as women fought to claim more rights through groups such as the Ladies' Land League, headed by Parnell's sister (657). Though *Portrait* begins around the time of Parnell's death, his political reign initiated a shift in women's roles that largely shapes the characterization of females throughout the novel. Joyce portrays the dual roles of women as mothers and temptresses in his illustration of the Ballyhoura Hills woman; likewise, he draws a contrast between the female characters' naïve subordination and their threatening lack of silence towards men in his depiction of Dante in the dinner party scene. Together, the two characterizations parallel Joyce's sociological perspective on the duplicity of the feminized Mother Ireland that nurtures but also confines.

The woman of the Ballyhoura Hills is both a maternal figure and a temptress. Davin, Stephen's friend and fellow peer, describes the duality in her appearance: "She was half undressed as if she

was going to bed...and I thought by her figure and by something in the look of her eyes that she must be carrying a child...I thought it strange because her breast and her shoulders were bare" (Joyce 160). The woman's double appearance reflects her psychosocial position. While her role in society is to serve as a mother, she still holds an innate purpose of seduction that is pointed out by her bare composure. In contrast, Dante's final, and rather destructive, expression in the dinner scene argument contributes to Joyce's unstable, feminized view of women. As Stephen remembers of his governess, "Dante shoved her chair violently aside and left the table...At the the door Dante turned around violently and shouted down the room, her cheeks flushed and quivering with rage...The door slammed behind her" (Joyce 34). Supplemented by her emotions, Dante takes physical action here. If a man had behaved as Dante, his actions would probably have been viewed as an assertion of power. However, as exemplified by Mr. Dedalus' "guffaw of coarse scorn," Dante's angry movements are perceived as signs of her instability, and even rudeness (44). The other woman in this passage, Mrs. Dedalus, is completely without voice and mindlessly follows the typical hostess protocol. Though contrasting characterizations, the three women in these two scenes adhere to the image of the subdominant Joycean female.

In the same manner, the Ballyhoura Hills woman could be seen as an analogy for Ireland, an entity with two sides: the sustenance-providing country and the home of artistic confinement. According to literary scholar Martha Fodaski Blac, Joyce's departure from Ireland—as also represented by Stephen's own escape in *Portrait*—represents his attempts to liberate himself from a sickly enslavement to his birth land and its "patriotic institutions," which include the church and the social restrictions such as gender-based expectations (Blac 86). When Davin presses Stephen to accept the Irish culture, Stephen replies, "When the soul of a man is born in this country there are nets flung at it to

hold it back from flight. You talk to me of nationality, language, religion. I shall try to fly by those nets” (Joyce 179). Stephen condemns the webs of Irish society that handcuff him. These bondages that confine Joyce to the domestic realm also seek to fashion a uniform society by setting a standard within the domains to which each person must adhere in return for acceptance by others in the same penitentiary. It seems that by bidding to “fly by those nets” Joyce yearns to escape conventionality. Eventually, Joyce did leave Ireland, yet he remained enslaved to it in another sense: he continually returned to it in his writing. However, some ambiguity rests in Davin’s analysis of the woman and thus leaves one to question Joyce’s own certainty in the level of Ireland’s social incarceration. As Marian Eide mentions in her critical analysis, Davin addresses the fact that the woman seems to be pregnant, yet he lacks actual proof of his observation. Her pregnancy forms a sort of barrier between Davin and herself—she becomes out of reach. Her actual intentions of partaking in any sexual activity with the young man are also left to speculation, as Davin cannot comment on much except a look in her eyes that can easily, by any human error, be miscalculated. Still, Davin’s perception of the encounter leaves him to see the woman as a provocative figure that is the perfect embodiment of Irish motherhood, and it is with this conviction of the woman’s seemingly suggestive manners and her possible pregnancy that Davin depicts his story (306).

Unlike the seeming eroticism the Ballyhoura Hills woman may suggest, her state of pregnancy and her actions attest to her inherent maternal qualities. Davin recollects, “A voice asked who was there and I answered...that I’d be thankful for a glass of water. After a while a young woman [of the Ballyhoura Hills] opened the door and brought me out a big mug of milk” (Joyce 160). Giving Davin a glass of milk instead of water is emblematic of woman’s motherhood, since milk is often associated with a mother’s nourishment. By forming her own interpretation of the request, the woman demonstrates a maternal-like instinct while confirming her role as a primary provider of life and sustenance. Thus, the strange woman, if duly acting as a temptress, may first use her maternal façade as a welcoming allure for the young man. The milk, too, could be seen as a means of deception. While the milk may signify the beginning (mothers provide milk to their kin at the start of their lives), it may actually be a means to an end, into the temptress’ bed. This analysis is especially relevant in Davin’s situation, who as Nehama Aschkenasy notes, “is [consciously] not afraid for his life. Subconsciously, however, he is reenacting an archetypal male experience, in which the female is conceived of as a threat and a mortal enemy.” Undoubtedly, there lies a clear differentiation between the struggles of Davin’s subconscious and conscious minds; Davin does not knowingly acknowledge his fear of death at the hands of the strange woman, but perhaps looking into the depths of his mind, he undergoes the “archetypal” experience and succumbs to the ominous, antagonistic image of women which arouses his anxiety (Aschkenasy 31). Women are again viewed as conniving, adversary forces. Stephen’s particular, detailed focus on Davin’s experience with the cottage woman reflects his own dismay towards females—though women did not destroy Stephen physically, his sexual encounters with them expedited his spiritual demise. Thus, as natural temptresses, women become the threat and “mortal enemy” referred to by Aschkenasy.

Religion holds as a vital force in the way women are perceived throughout the novel. As Stephen receives the Communion



Figure 1. A 1915 image of James Joyce in Zurich.

wafer, he ruminates, “Another life! A life of grace and virtue and happiness! It was true...The past was past” (Joyce 127). This scene follows Stephen’s sexual undertakings with women. He looks to the wafer as means of forgiveness from God and as a savior from himself. Stephen’s epiphany-like reflection arises from desperation. To him, the “past was past” and he could begin once more with a clean slate, though that slate too is corrupted soon after. The milk versus water passage serves as a reminder to Stephen of his tainted purity and his ephemeral religious epiphanies in an uncanny resemblance to the way mother Ireland inhibits his artistic self.

Joyce parallels the “milk not water” detail in the Ballyhoura Hills woman scene and the biblical books of Proverbs and Judges, both of which portray a threatening image of women that does not stray much from the image presented at the dinner party. Jael of Judges, like the strange cottage woman Davin encounters, draws a man into her home and answers his request for water with milk; Jael, however, murders her guest and is heralded for her actions against the enemy (Judges 4:17-22). While the Ballyhoura Hills woman does coax her guest and attempt to keep him through the night, as Jael had, Davin refuses to comply with her request. A devout Protestant, he would perhaps recall the story of Jael and take warning against the possible dangers imposed by women. While Davin may take caution from the book of Judges, Stephen heeds the Proverbial image of women, by which females are perceived as whores who essentially turn men—like

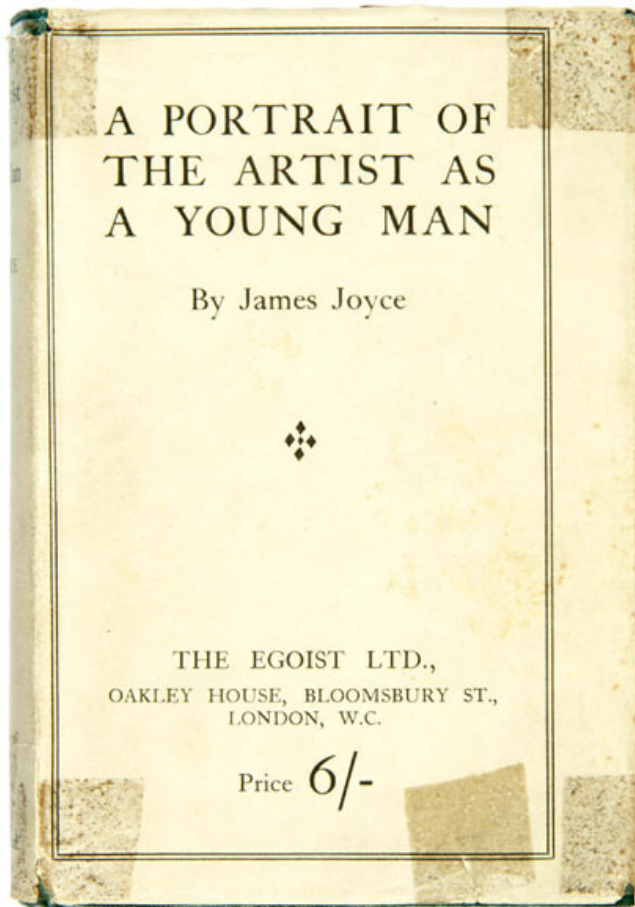


Figure 2. The cover for the first UK book edition of *A Portrait of the Artist as a Young Man*.

King Solomon—away from God. Stated in Proverbs, as was Solomon’s knowledge during his affairs, is a warning against the evasive veneer of women: “For the lips of the adulterous woman drip honey, and her speech is smoother than oil” (Proverbs 5:3). This Proverbial depiction of females mirrors Stephen’s first encounter with a prostitute. Joyce writes of Stephen, “He closed his eyes, surrendering himself to her, body and mind, conscious of nothing in the world but the dark pressure of her softly parting lips” (Joyce 88-89). By resigning to the temptation of the woman, Stephen frees himself from any moral constraints he previously kept, specifically his religious morals of purity. He abandons not only his body, but his mind. His entire being is surrendered in exchange for the sensual pleasure of “softly parting lips.” It is this “dark pressure” that ultimately corrupts his soul, transitioning him from a life of purity, often associated with brightness, to that of sin and darkness. Hence, like the misleadingly sweet words of Solomon’s harlots, the cottage woman’s invitation to her home is but a means of deceit, a point Joyce attempts to subtly hint at through his allusions.

The hesitation Davin experiences, in light of the biblical recollections, mimics Joyce’s precautionary attitude toward both women and Ireland. Joyce decentralizes the focus on the cottage woman as a person and instead focuses on her as a “race.” As Stephen recounts in the novel, he sees the woman “as a type of her race and his own, a batlike soul waking to the consciousness of itself in darkness and secrecy and loneliness” (Joyce 183). The peasant woman, to Stephen, stands for all the constrictions



Figure 3. Statue of James Joyce on North Earl Street in Dublin.

set upon him by his environment. The female figure is not only a fleshly reminder of corruption and decay, but also a symbol of all the forces—home, country, and bodily lust—that have coalesced to paralyze the young artist and imprison his free spirit. He associates femininity with his home and his country, and thus his enemy as he yearns to flee from these mothering aspects of society. As Aschkenasy indicates, Stephen’s disdain for the opposite gender represents his perspective that women are an attacking force that aim to keep their kin nearby before destroying them. In Stephen’s case, feminine figures—home and country—that keep him where he is now gradually cause his demise by not allowing him to liberate the spirit within him by diverging from his repressing environment (Aschkenasy 37). As Davin turned his head away from the dangers of unholy women, Stephen attempts to flee from his own harlot and Jael, Ireland.

Similarly, the dinner party uses the topic of religion as a backdrop to contrasting women who are pressured into subordination and women who threaten men with their lack of silence. Stephen recollects, “Mrs Dedalus spoke to Dante in a low voice but Dante said loudly: —I will not say nothing. I will defend my church and my religion when it is insulted and spit on” (Joyce 29). Particularly apparent in this scene are the characters mirroring Joyce’s governess Charlotte Stoker and his mother Delia Parnell. Dante and Stephen’s mother, the two mirrored characters, offer two stark comparisons between the types of women—Stephen’s mother is quiet and reserves her own potential opinions as a sort of “politeness” toward the men at the dinner table.

The parallelism between women and Ireland is further substantiated through Davin’s final reflection with the imagery of the “batlike soul” which contains a double meaning, one for the woman and one for Ireland. Comparing the woman to a temptress, a “batlike soul” emphasizes a sense of Stephen’s detachment from the woman—by comparing her to a bat consumed in darkness, secrecy, and loneliness, Stephen justifies his isolation from women by placing the blame on their own estrangement. In fact, only the unattainable women, such as Virgin Mary, seem to attract Stephen, who notes that “the glories of Mary [hold] his soul captive” (Joyce 91). It is no surprise, then, that Stephen chooses to compare the simple and very attainable cottage woman to a bat which, as Laurie Teal indicates, was an animal frequently used in the 1800s as reference to a harlot or whore. Thus, Joyce seems to sexually taint the natural essence of motherhood. Speaking of “her race and his own” forms a comparison not only between the woman and Stephen, but also Stephen and Ireland, which, like the woman, Stephen views as an antagonistic entity. To Joyce, the woman voices herself as Ireland and not as the individual woman who she is (Teal 72). The “batlike soul” therefore may not represent the virginal image of the woman but rather that of Ireland; while Ireland through a physical sense appears to form a chaste society when it in fact creates a home for eroticism and prostitution, a home of sin and moral failure. During his own encounter with a prostitute, Stephen first notes that “her room was warm and lightsome” before acknowledging the sensuality of the woman and her appearance as she unravels her clothing (Joyce 88). The sequence of observations attests to Stephen’s naive mentality at that point in the novel as he initially feels welcomed by the hospitable atmosphere. But just as Ireland’s fundamental image transgresses into one of impurity, so does the prostitute’s disposition. In a matter of seconds, she is able to transform herself into a sexualized “doll,” as Stephen calls her, and reveal



Figure 4. A 1918 image of James Joyce.

her true self by removing her clothing, which hid her facade of maidenhood.

On the other hand, Dante has no fear of being regarded as “impolite” and touches upon conversational topics throughout the argument. Michael Wainwright evaluates, “The young Joyce’s formative impression of Irish womanhood must have been particularly divided: the stoical silence of his mother, a woman only rarely driven to assertiveness, in stark contrast to the demonstrative vocality of his governess, a woman often compelled to express her views.” Like Joyce’s own governess, Dante is confident of her views and unafraid to voice them to the male parties. This contrast can be paralleled to Joyce’s split image of Ireland. While he believes that Ireland serves as an artistic confinement, his country continues to shape him and in many ways—through his works, for instance—he keeps returning to Ireland (Wainwright 657). Or possibly the contrast in the women’s behaviors demonstrates something else: while Dante demonstrates her volubility and even courage by defying the common norms of women’s behaviors (as depicted by Mrs. Dedalus and her request for Dante to keep silent), her devotion to the church can be seen as a weakness, a sign of naivety that reflects more on her social status than her religion—she cannot form her own opinions on the politics discussed at the table and in turn devoutly adheres to the viewpoints of the church, stating that “the bishops and priests have spoken...and they must be obeyed” (Joyce 27). While Dante on one hand stands up for herself amongst men, she subjects herself to the men of institutions. These clergymen instill a culture of obedience within the vulnerable, specifically women like Dante who devoutly follow the law of the Church. If equating the female characters to Ireland, one can remark on

the potential single-mindedness of a culturally traditional Ireland.

Dante's counteraction with Mr. Casey in the Christmas dinner party scene demonstrates the lack of males' respect for females and their desire to continuously subordinate women. Joyce writes, "Dante started across the table, her cheeks shaking. Mr. Casey struggled up from his chair and bent across the table towards her, scraping the air before his eyes with one hand as though he were tearing aside a cobweb" (Joyce 34). There appears to be an overwhelming dynamic between the man and woman, Casey and Dante. Whereas Dante reacts emotionally in this line, Mr. Casey's reaction is much more physical—he makes the effort to get up and "[scrape] the air," conveying perhaps the man's necessity to assert his control through behaviors. Comparing Dante to a cobweb demeans her to the status of a bug; as men would normally stomp or grab at pests with the intention to get rid of them, Mr. Casey snatches at Dante. Hence, the emotion-centered women are degraded and subdued by men. Even when prompted to hold silence in respect for the men at the table, Dante transcends her "obligations" as a woman and interjects with her own thoughts on the situation at hand—the church v. Parnell—thereby continuing the "finished" conversation. Joyce represents women in a new light, forming one of either two depictions of the female characters: they stand for their own beliefs with courage or they impolitely overstep their duties (Toolan 404).

By continuously attributing feminine qualities to Ireland throughout *Portrait*, Joyce parallels the threatening portrayal of women in the novel to Ireland as a force restraining an individual's, in this case Stephen's, artistic pursuit. While Davin describes his encounter with the Ballyhoura Hills woman to Stephen, he states, "When I handed her back the mug at last she took my hand to draw me in over the threshold and said: Come in and stay the night here...I thanked her and went on my way again, all in fever. At the first bend of the road I looked back and she was standing in the door" (Joyce 160). Davin's reaction to the woman's invitation formulates a threatening image of her. Her forwardness with Davin potentially takes him by alarm. The woman too continues to stand at the door after Davin rejects her invitation, causing Davin to feel as if she is emotionally stalking him.

Much in the same way, Ireland seems to keep a watch over Stephen as he attempts to escape from its confinements. In the last section of the novel, Stephen writes in a letter, "Mother is putting my new second hand clothes in order. She prays now, she says, that I may learn in my own life and away from home and friends what the heart is and what it feels" (Joyce 224). Stephen's mother, another feminization of Ireland, attempts to place her son's life "in order" and to shape his journey away from her though her prayers. Her words can be seen as an insurance to Stephen of her, and thus Ireland's, spiritual watch over him. As a result, despite Stephen's physical departure, his mind is fixed upon mother Ireland and her last artistic appeal—she offers Joyce a confined liberation that, with the societal pressures placed on Joyce also comes a nurturer of literature. Ireland, the topic of much of Joyce's writing as in *Portrait*, becomes in one sense Joyce's muse in his works. Thus, Ireland, like the women of the text, possesses a double nature, one which socially imprisons its people based on gender-driven standards and the other which nurtures a passion—though this ardor comes much from the social frustration experienced by Joyce.

If Joyce in the dinner scene was illustrating the female characters' tendency of subordination, he paints another image

through Stephen's examination of the inescapable "maze," referring to the various forms of beauty in a woman as perceived by men. In a conversation with his peer, Lynch, Stephen comments on two methods of objectification by which one may lessen the position of women:

One is the hypothesis: that every physical quality admired by men in women is in direct connection with the manifold functions of women for propagation of the species...For my part, I dislike that way out. It leads to eugenics rather than esthetic...There remains another way out...all people who admire a beautiful object find in it certain relations which satisfy and coincide with the stages themselves of all esthetic apprehension (Joyce 208-209).

Stephen first clarifies and dismisses the initial claim that female beauty stems from a woman's ability to reproduce. In essence, a man yearns to breed selectively by finding the mate that will provide him with the most capable offspring. The second proposal, which Stephen more aligns with, objectifies women as "beautiful object[s]" which provide visually pleasing satisfaction for their male counterparts. Stephen turns the sphere of nature, characterized by his fear and anguish toward women as producers, to a realm of art where females are perceived of as less threatening and mere objects. This fear of women, also present in the encounter between Davin and the "batlike" Ballyhoura Hills woman, speaks to the need for men to constantly subordinate women, such as in the Casey and Dante scene. Sexualization, one form of subordination, possibly stems from the notion of productivity; while women are key to human reproduction, a country is vital to the reproduction and sustenance of a civilization. Thus, the stigma of women as forces of procreation eclipse into the ideals of a feminized nation. When one fails to do its bidding, all else seems to collapse. Joyce further portrays this belief onto Ireland by turning the land into a form of art and focusing on its aesthetics. His feminization of Ireland already translates the illustrations he writes of women to those he feels of his homeland; in other words, his beautification of the female actually serves as a metaphor for his objectification of Ireland.

As Ireland itself, a jail of artistic pursuit, fabricates itself a façade of motherhood, so too do her daughters; the woman of the Ballyhoura Hills and Dante of the dinner scene both exemplify and defy the social duplicities of 20th century Irish women: the temptresses towards sin and the subordinates to man. With passing time, societal implications fortify and persist. Women, still, are continuously objectified and lessened in a man's search for freedom. If those who, like Joyce, yearn to liberate themselves from social confinements, why must they parallel the faults of their country to the female kind? Perhaps a woman's curves of emotion yield greater fruit than a man's mere lines of intellect.

REFERENCES

1. Aschkenasy, Nehama. "Biblical Females in a Joycean Episode: The 'Strange Woman' Scene in James Joyce's 'A Portrait of the Artist as a Young Man.'" *Modern Language Studies* 15.4 (1985): 2839. JSTOR. Web. 7 Mar. 2015. <<http://www.jstor.org/stable/3194646>>.
2. Blac, Martha Fodaski. *Shaw and Joyce: The Last Word in Stolen Telling*. Tampa: UP of Florida, 1995. Print.
3. Eide, Marian. "The Woman of the Ballyhoura Hills: James Joyce and the Politics of Creativity." *James Joyce's "A Portrait of the Artist as a Young Man": A Casebook*. Ed. Mark A. Wollaeger. New York: Oxford UP, 2003. 305-36. Print.

4. *English Standard Version*. Bible Gateway. Web. 25 Oct. 2012.
5. Joyce, James. *A Portrait of the Artist as a Young Man*. 1916. Ed. John Paul Riquelme. Norton Critical ed. New York: W. W. Norton, 2007. Print.
6. Teal, Laurie. "Batlike Souls and Penile Temptresses: Gender Inversions in 'A Portrait of the Artist as a Young Man.'" *NOVEL: A Forum on Fiction* 29.1 (1995): 6378. JSTOR. Web. 8 Mar. 2015. <<http://www.jstor.org/stable/1345540>>.
7. Toolan, Michael. "Analysing Conversation in Fiction: The Christmas Dinner Scene in Joyce's 'Portrait of the Artist as a Young Man.'" *Poetics Today* 8.2 (1987): 393416. JSTOR. Web. 8 Mar. 2015. <<http://www.jstor.org/stable/1773043>>.
8. Wainwright, Michael. "Female Suffrage in Ireland: James Joyce's Realization of Unrealized Potential." *Criticism* 51.1 (2009): 65182. JSTOR. Web. 8 Mar. 2015. <<http://www.jstor.org/stable/23131535>>.
9. Wilde, Oscar. "An Ideal Husband." 1895. *The Plays of Oscar Wilde*. New York: Modern Library, 1974. 1-125. Print.
10. "The Woman of the Ballyhoura Hills: James Joyce and the Politics of Creativity." Trans. Marian Eide. *James Joyce's "A Portrait of the Artist as a Young Man": A Casebook*. Ed. Mark A. Wollaefer. New York: Oxford UP, 2003. 305-36.

IMAGE SOURCES

FIGURE 1: Ehrenzweig, Alex. "James Joyce, 1915 restored." 28 April 2010. Online image. Wikimedia Commons.

FIGURE 2: The Egoist Ltd., London. "A Portrait of the Artist as a Young Man." 1917, uploaded 7 April 2007. Online image. Wikimedia Commons.

FIGURE 3: Tonihier. "Joyce oconnel dublin." 8 October 2005. Online image. Wikimedia Commons.

FIGURE 4: Ruf, C. "Revolutionary Joyce Better Contrast." 10 July 2007. Online image. Wikimedia Commons.

ACKNOWLEDGEMENTS

I would like to thank Dr. Anne Douglas at The Harker School, who first set me on the incredible exploration of Joyce's novels; Professor David Font at Duke University, who ignited my passion for literary research; and Sona Sulakian, who has over the years helped shape my writing and strengthen the voice in my work.



MARIAM SULAKIAN

Mariam is a freshman at Duke University majoring in Computer Science with minors in Statistical Science and Creative Writing. A native of the Silicon Valley, Mariam has always been fascinated by the expanding intersections of technology, policy, and literature. When she's not coding or writing, she can usually be found working on her startup or advocating for human rights in her political arena. Above all, Mariam hopes to take her findings from different sectors of research, whether in the humanities or STEM, and apply them through social entrepreneurship to the greater international community, such as in Armenia where she has been an active volunteer the past nine years.

“To Confront a Portrait:” Photography, Parts, and the Whole in the Poetry of Walt Whitman

Alex Torres

Stanford University

Walt Whitman was the most photographed poet of the 19th century. How, then, did photography—if at all—impact his poetry? This essay argues that Whitman used photographs and the figure of photography to explore Emerson’s ‘each and all’ concept, and to illuminate how individuals add up to the ‘all’ of America. In order to do so, I parallel two of Whitman’s greatest projects: *Leaves of Grass* and his photographs of the self. I consider the book history of *Leaves of Grass*, paying particularly close attention to the materiality of Whitman’s revision and his photographs. I argue that *Leaves of Grass* functions as an analogy for a “photo album,” the additions of new poems and changing of photographs emblematic of the addition of photographs. Along the way, I make a related claim about the collection’s emphasis on the body in relation to Whitman’s body, and how the emphasis on the body relates to the advent of photography. As the advent of the photograph also coincided with the Civil War, Whitman’s allure of photography likely stemmed from the potential of photographs to mend and heal the nation and the self through memorialization as much as from the failure to conceal the trauma of battlefield carnage.

During his lifetime, American poet Walt Whitman was immersed in a culture of visual art—especially photographs. In a conversation with his biographer and friend Horace Traubel, Whitman once said: “I think the painter has much to do to go ahead of the best photographs” [1]. Whitman’s preference for photography likely stemmed from his own obsession with photographs of the self. As the poet himself observed in another conversation with Traubel: “I have been photographed, photographed, photographed, until the cameras themselves are tired of me” [2]. But as Ed Folsom proposes in *Walt Whitman’s Native Representations* (1994), Whitman’s obsession with photography also extended beyond the curious desire to see the self. He notes, “Whitman was of the first generation to experience the world in photographic images; his poetry emerged at precisely the time photography was literally taking a hold of the American imagination. Whitman’s poetics, of course, were in large part built to meld the mechanical and the spiritual to discover and sing the deeper meaning of science...” [3]. Folsom primarily focuses on how the advent of photography in America fascinated Whitman, and how the circulation of photographs allowed Whitman and other Americans to see previously unexplored sections of the country. Yet Folsom’s work does not fully address *why* Whitman was particularly drawn to the photograph.

To explore this question, I turn to Whitman’s relationship to his contemporary Ralph Waldo Emerson, a transcendentalist whose poetry and essays influenced Whitman’s famous poetry collection, *Leaves of Grass*. More specifically, this essay argues that Whitman used photographs and the figure of photography to explore Emerson’s ‘each and all’ concept, and to illuminate how individuals add up to the ‘all’ of America. In order to do so, I

parallel two of Whitman’s greatest projects: *Leaves of Grass* and his photographs of the self. I consider the book history of *Leaves of Grass*, paying particularly close attention to the materiality of Whitman’s revision and his photographs. Not only did Whitman publish nine versions of *Leaves of Grass*, but he also included a new photograph of himself in almost every edition. I argue that *Leaves of Grass* functions as an analogy for a “photo album,” the additions of new poems and changing of photographs emblematic of the addition of photographs. Along the way, I make a related claim about the collection’s emphasis on the body in relation to Whitman’s body, and how the emphasis on the body relates to the advent of photography. While my analysis considers *Leaves of Grass* as a whole, I pay particularly close attention to “Out from Behind This Mask [To Confront a Portrait].” I argue that the speaker addresses the photographs to mimic the act of parts of becoming whole, the act of the individual becoming universal.

Whitman’s interest in Emerson’s writing and ideals is apparent in Whitman’s earliest edition of *Leaves of Grass*. In the Preface to that 1855 edition, Whitman writes: “Nothing out of its place is good and nothing in its place is bad” [4]. Whitman’s remark echoes the Emersonian concept of “each and all,” which Emerson uses to explain the relationship between the part and the whole, the individual and the universal. As Norman Miller explains, Emerson’s poem “Each and All” moves through “a series of three ‘cases’ in which particulars—a sparrow, sea shells, a virgin—are removed from their proper setting and ‘brought home.’ Each loses its charm and beauty when isolated from its natural environment” [5]. For Emerson and Whitman, the “each and all” concept hinges primarily on wholeness: “Given this nature, [Emerson’s philosophy] resists penetration and probing.

Tear it at one point and the whole construct falls” [5]. For Emerson and Whitman, each part was necessary for an adequate whole.

Whitman’s fascination with the “each and all” concept emerges not only in his poetry but also in his relationship to photography. In his biography of the poet, Horace Traubel details an 1889 encounter with Whitman and his photographs: “If I could get a book to suit me, into which I could put the pictures to suit me, I would be happy. I wonder if it could be done?—a book about this size?”—[Whitman said], measuring about a foot square. ‘Not necessarily larger—or larger at all.’ He had a great mess of pictures around and had often thought to collect them” [1]. Whitman’s desire to put his photographs together into one album implies a desire to make the parts of the self become whole, a desire that often overshadowed his own poetry. Housed in the Beinecke Rare Book and Manuscript Library at Yale University, the Walt Whitman papers offer additional evidence that Whitman was obsessed with collecting photographs of the self together into one locality. On April 14, 1887, while reading at a lecture in remembrance of President Lincoln and Madison Square Theater, Whitman handed out a small pamphlet not of his poetry but of his photographs. The photographs present Whitman in an array of poses, ages, and locations: a younger Whitman looking slightly to the right, an old Whitman to the left, Whitman recovering from his stroke in a wheel-chair, Whitman lying down on the ground with a dog [6].

Whitman’s desire to collect his photographs into one album resembles the process he used to create *Leaves of Grass*. Whitman began to write *Leaves of Grass* in 1855, when the collection consisted of a mere twelve poems, lacking any organization. Over the course of forty years he continued to revise and add new poems to the collection. Following four decades and eight editions, Whitman produced the final “deathbed” edition of *Leaves of Grass*, which included over 400 poems. In addition to adding poems, Whitman also began to organize the collection into a more cohesive whole. As Betsy Erkkila observes, “[the] most significant change in the 1881 *Leaves* is not the addition of new poems but Whitman’s restricting of the entire volume into a final coherent form” [7]. For Whitman, the arranging of poems and photographs into one collection, it seems, was an essential aspect of the artistic process.

While comparing collecting photographs of the self to collecting poems about America may seem to be unrelated, the presence of the body nonetheless resonates throughout *Leaves of Grass*. Building upon this central premise, Erkkila suggests parallels between Whitman’s body and America: “And yet there had always been a curious correspondence between Whitman’s body and the body politic of America: His body seemed at times a National seismograph, registering disturbances in the political sphere” [7]. Although Erkkila neglects the association of photography to the body, it is important to note that the advent of photography brought an emphasis on the body, portraits, and the self.

Whitman reveals the relationship of poet, body, and nation in the edits to his 1855 poem, “I Sing the Body Electric:” “I sing the body electric, / The armies of those I love engirth me and I engirth them” [4]. Here, the speaker “sings the body,” implying a celebration of the body. However, the relationship of the speaker’s “body” appears to be symbiotic with that of the “arm[y]’s”: the speaker’s body “engirth[s]”—surrounds—the “armies” as the

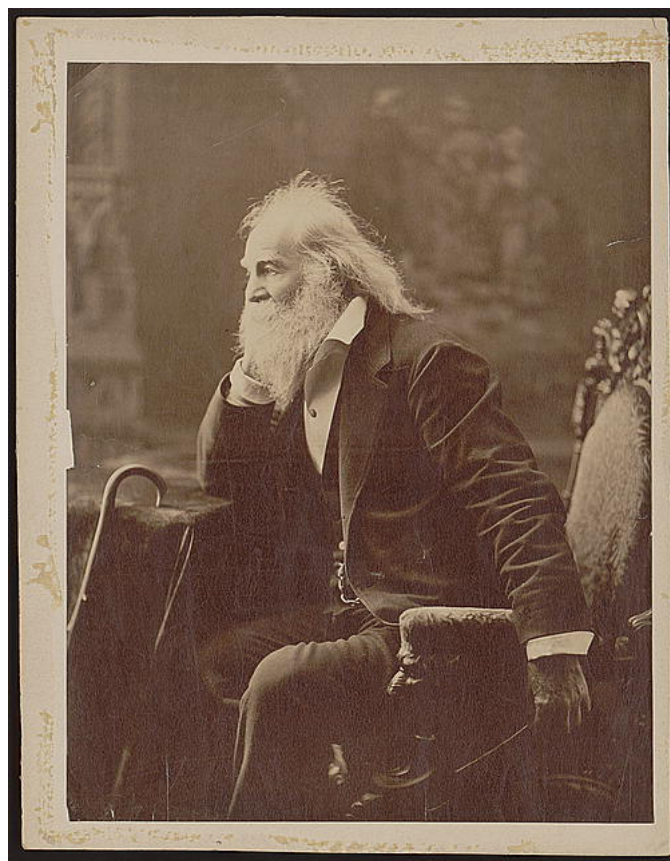


Figure 1. This photograph shows Whitman in 1881, the year he famously reorganized *Leaves of Grass*.

“armies” “engirth” the speaker. In essence, the speaker unifies the part (the individual) and the whole (the armies) through the figure of the body. Whitman’s edits reveal a heightened importance of the body, likely as a result from the carnage of the Civil War. In fact, “electric” might refer to the electricity that was used for the staged photographs of Whitman’s later portraits, and the staged photographs of dead soldiers.

Whitman draws again upon the body to explore the relationship between the self and the nation in “The Sobbing of the Bells,” this time to explore American suffering [4]. The poem begins by the personification of the “sobbing...bells.” Here, the sounds of the city emulate the tears of a person. The speaker continues, describing how the people “(Full well return, respond / within their breasts, their / brains, the sad reverberations).” Structurally, the poem moves from the personified images of the country to the literal descriptions of the citizens, which underscores the subtle equation of nationhood and body that seems to inhere in the initial metaphor. The description returns back to the city scape, as the speaker depicts “The passionate toll and clang, / City to city joining, sounding / passing.” The “joining” of the cities suggests an attempt to mark parts whole, to connect the fragments of the nation. The repetition of noises, then, prime the poem for its final metaphor: “Those heart-beats of a Nation / in the Night.” The subtle equation of body and nation becomes complete, as the “sobbing...bells” have now become “heart-beats of a Nation.” Like the “sobbing...bells,” the speaker personifies the Nation; namely, the “heart-beats” belong not to the individuals but to a vital, unifying force. While the poem eulogizes the assassination President James Garfield, “The Sobbing of the Bells” also seems to emulate Whitman’s famous

1881 photographic portrait, which displays the poet sitting down with his cane, a symbol of the poet's faltering health [4](See Image 1).

Recalling Erkkila's observation of the relationship between Whitman's body and the nation, the parallel between Whitman's 1881 photograph and "The Sobbing of the Bells"—the body and nation—reveals the potential for Whitman's photographs of the self to symbolize the whole of the nation.

In addition to the body's importance to Whitman's poetic aesthetic, the inclusions of Whitman's own body via photograph was also an integral part of each edition of *Leaves of Grass*¹. In an 1881 letter to James Osgood, Whitman reveals his anxiety over selecting the right photograph: "How would the enclosed picture do for a frontispiece? I like it—it is made by Gutekunst" [8]. In his critical work, Sean Meehan reminds us that according to Whitman, "[the frontispiece] represents the book's autobiography soul," and presumably the implicit relationship between a photo album of the self and a book of poetry [9]. And as Ed Folsom proposes, Whitman also included photographs of the self in a more poetic sense: "Whitman's most direct poetic statement about the power of the photographic face is 'Out from behind This Mask [To Confront a Portrait]...Whitman made it clear that the poem referred to W. J. Linton's woodcut of George Potter's 1871 photograph of the poet, which had appeared in the 1876 *Leaves*. But Whitman set the poem floating free of this portrait in *Leaves*, so that the subtitle not only invites the reader to hear the poem as a poetic statement literally confronting the Linton/Potter portrait, but also as a kind of general tutorial on how we should confront any portrait, any photograph of the face" [3]. As Folsom suggests, the first lines of the poem emulate the process of woodcutting, and the "bending rough-cut mask" evokes the process of transferring a photograph into a woodcut, the same process that was used to place Whitman's portrait in *Leaves of Grass*.

While Folsom argues that "Out from behind This Mask [To Confront a Portrait]" depicts the speaker examining a photograph of the self, he does not explore the full implications of the poem's language. Recalling photography's historical moment in context of Whitman's life, Sean Meehan proposes: "The play and provocative uncertainty that Whitman reads in every photo, and makes much of in this instance, derives, most crucially, from the creative potential of photo reproduction, the play and moment of the photographic process that allows the reality of an image to be approximated and continuously revised" [9]. The paradigm of "provocative uncertainty," the "approximated and continuously revised [reality]," emerges here in "Out from behind This Mask [To Confront a Portrait]." To begin, the poem's first four lines are highly evocative of Emerson's "each and all concept:" "this drama of the whole / This common curtain of the face contain'd in me for me, in you / for you, in each for each" [4]. The figure of the "common curtain" alludes to the relationship between individual and universal: the photograph stands in for both the speaker and for the collective American consciousness. The repetition of "me," "you," and "each" emulates the relationship of the speaker and the photograph, implying a sense of reproduction.

The "common curtain" becomes a part of the "me," "you," and "each," all of which come together to form "this drama of the whole"—a "drama" that evokes a paradox of crisis and excitement.

Drawing upon the paradox of the "drama of the whole," the speaker also explores the realms of Heaven and Hell. Emerson's "each and all" concept not only examines the relationship of the individual to the universal but it also challenges the established religious hierarchy, as it argues that the differences among parts are attributed to each part's variety rather than each part's place in an American hierarchy. In the poem, "glaze" and "film" are synonymous, both connoting the cover of the photograph [4]. Further, I would suggest that the repetition mirrors what Meehan describes as the "potential of photo reproduction," a sentiment that resonates throughout the entire poem. More specifically, the speaker mirrors the fungibility through the parallel sentence structure, suggesting that equality exists between God and Satan, Heaven and Hell [4], [9]. The descriptions of Heaven ("serenest purest sky") and Hell ("seething pit") both reveal a sense of extremity, as both descriptions are opposite yet seemingly equal in magnitude [4]. In addition to the equivalence the speaker records

The speaker's imaginative use of the photograph extends beyond seeing aspects of Earth and toward understanding the relationship of part to whole between the self and the infinite.

between Heaven and Hell, the photograph also functions as a window into the afterlife, as it permits the viewer to experience the dead of Heaven and Hell. This ultimately gestures towards photography's ability to eerily mix the past with present, the living with the dead. For Whitman and Emerson, to forget the dead

is to "[t]ear [the whole] at one point" and watch the "construct fall[]" [5]. That is, the dead are as important as the living, as both groups are part of a larger whole.

The photograph allows the reader to examine not only the afterlife but also the natural world, as it becomes "[t]his heart's geography's map, this limitless small continent, this soundless sea" [4]. For the speaker, the photograph becomes a "map," an object that shrinks the whole in order to make it understandable. More specifically, a "geography map" also implies that the photograph functions as a tool to also see the entire natural world. Inadvertently, the speaker becomes a part of nature. The relationship of individual to universal—of self to nature—is evident from Whitman's earliest poetic attempts. In "Song of Myself," the speaker draws upon the figure of the "atom" to demonstrate a sense of equality between the speaker and the nonhuman world: nature [4]. As the speaker proclaims, "My tongue, every atom of my blood, form'd from this soil, this air" [4]. The speaker demonstrates how the composition of his "blood" mimics the composition of the "soil" and the "air," all of which are created by joining atoms together. The idea that the speaker's human body and the facets of nature are formed from equal parts serves as a model for how the relationship of individual to universal functions in "Out from Behind This Mask [To Confront a Portrait]." The part that is the photograph compresses into a smaller, comprehensible whole. By viewing "Out from Behind This Mask" through "Song of Myself," one can conclude that the photograph also symbolizes the entirety of the natural world: that is, the self, in part, functions as a synecdoche for the natural

world.

In “Out from Behind This Mask,” the speaker also describes his own face as a “limitless small continent” and as a “soundless sea,” revealing the allure of nature for Whitman. Returning to the theme of equality, “small continent” and “sea” quite literally describe the actual photograph, the parallel structure highlighting that neither aspect of nature appears superior to the other. Moreover, “limitless” and “soundless” both suggest the photography’s ability to remain unfathomable to the speaker, opening up the speaker’s power of imagination. Together, the “limitless continent” and the “soundless sea” form the “heart’s geography’s map:” the “limitless,” “soundless” whole of the external world’s geography simultaneously becomes a representation for the speaker’s inward experience. Certainly, the speaker’s use of the photograph for the imagination emulates the “provocative uncertainty that Whitman reads in every photo” [9].

The speaker’s imaginative use of the photograph extends beyond seeing aspects of Earth and toward understanding the relationship of part to whole between the self and the infinite. For the speaker, the photograph comes to symbolize “the convolutions of this globe,” which underscores the complexity of the Earth [4]. The speaker, then, uses the photograph to explore how the Earth itself is a part to a greater whole, a “subtler astronomic orb than sun or moon, than Jupiter, Venus, Mars” [4]. In observing the Earth becoming a part to the whole of the universe, the speaker implies too that he is a part of a vast universe. At the end of the stanza, the visions of Heaven and Hell; land and sea; global and universal become “[t]his condensation of the universe”—“a look” [4]. It seems that the “provocative uncertainty” of the photograph ultimately allows the speaker to examine how every part—the afterlife, nature, the universe—come together into one photograph. Or as John Mason proposes, the “reader” becomes involved “in the poet’s movement from “the singular to the cosmic” [10]. Building upon Mason’s premise, it seems that the speaker in “Out from Behind This Mask [To Confront a Portrait]” uses the photograph to become a part of America and a part of the universe—the ultimate form of egalitarianism.

Whitman equally reflects the relationship of part to whole in the poem’s poetic structure. The second stanza reveals a tension between independent present tense verbs (“I greet,” “I... turn”) against noun phrases with past or present participles (“A traveler of thoughts and years,” “Lingering a moment here and now,” “Pausing, inclining, baring my head”) [4]. To borrow Helen Vendler’s term, this demonstrates Whitman’s poetic thinking, his attempt to “compress a multifaceted scene, distributed over sequential time, into a single momentary gestalt” [11]. The model of the “momentary gestalt” emerges through the figure of the photograph: that is, the part (the photograph) takes the role of the whole (a momentary gestalt). The speaker uses the one instance in his life that the photograph depicts to see “A traveler of thoughts and years,” “Of youth long sped and middle age declining” [4]. The photograph allows the speaker to both see a “soul...once inseparable with [his]”—the former self—while also recalling the changes the speaker has undergone since the photograph was taken [4]. Time, then, begins to exist not on a horizontal, temporal plane but rather as a gestalt: the whole that is the speaker is perceived as more than the sum of his time spent living.

The speaker’s effort to interpret the photograph into one “drama[ti]c...whole” parallels Whitman’s own attempts to make meaning out of his collection of photographs. While



Figure 2. This is one of the earliest staged photographs of a dead soldier.

describing his mess of photographs, Whitman explains: “It is hard to extract a man’s real self—any man—from such a chaotic mass—from such historic debris” [1]. Like the speaker who has traveled through “peace and war,” Whitman used his mess of photographs to symbolize the process of understanding the trauma of the Civil War—of the whole splitting apart—upon the self. In this light, Whitman’s poetic speaker might also function as a synecdoche for the nation. As America underwent an identity crisis from antebellum to postbellum, so does the speaker. However, the analysis of the photograph also functions as a tool for the speaker’s introspection—for healing. In his interpretation of the poem, Folsom likens “Out from Behind This Mask [To Confront a Portrait]” to Whitman’s Civil War poems, highlighting the importance of the face: “The faces in Whitman’s poems become more pervasive and haunting in the Civil War poems, where soldiers’ faces again and again look into his own as he peers into and remembers face after face...” [3]. Drawing upon Emerson’s “each and all” concept, the speaker’s face becomes both individual and universal, allowing the portrait to embody the faces of “Tragedies, sorrows, laughter, tears,” “of peace and war”—the soldiers of the Civil War [4]. The relationship of individual to universal, then, might also echo the relationship of photograph to photo album. As Folsom observes, “All of these faces stay in the memory like an album of photographs of the lost and dead, a poetic echo of Brody and Gardner and O’Sullivan photos of the Civil War dead” [3]. (See Image 2) Whitman’s attraction to photography likely stemmed from the potential of photographs to mend and heal the nation and the self through memorialization as much as from the failure to conceal the trauma of battlefield carnage.

While one can only speculate, Whitman’s fascination with photography might have influenced his decision to write in free verse. As Grossman proposes, “The argument that made the meter of Whitman was the unification of the world in the one power of language, the secret authority of the poet...the bestowal of presence across time” [12]. Whitman’s photographs of the self also fulfill a similar function: unifying the public by the shared experience of seeing the poet through his photograph in every edition of *Leaves of Grass*. Certainly, Whitman has succeeded in bestowing his presence among 21st century readers. Whether or not we agree with David S. Reynolds that the photograph is “an

essential metaphor behind [Whitman's] democratic aesthetic," the notion nonetheless opens up new, hermeneutic possibilities for how we might read *Leaves of Grass* [14].

Despite Whitman's unquestionable fascination with photography, Whitmanians have largely ignored this aspect of his work and life, chiefly because of a crucial lack of order with respect to his photographs. Upon Whitman's death, his photographs were divided among three different literary executors, scattering into the abyss of libraries and other collections. As Horace Traubel observes, William Douglas O'Connor sought to make the photographs of Whitman's life an adequate whole: "I shall take care to have it full and complete: it would make a most remarkable presentment: I have always desired to do the thing" [1]. Unfortunately, O'Connor never put the photographs together into an album. Certainly, Whitman may have predicted the fragmentation of his photographs and materials, leaving scholars like Ed Folsom and myself to recollect Whitman's photographs of the self into an adequate whole, and to illuminate how those photographs retell the story of *Leaves of Grass*. Whitman was interested by the power that photographs possessed to trace the evolution of the self, and to explore how the body could function as a metaphor for the United States. These metaphors of the self and the body, alongside photographs, allowed Whitman to make parts whole, to mend and heal the nation from the wounds of the Civil War.

REFERENCES

1. The Walt Whitman Archive. Eds. Folsom E and Price KM. Lincoln: The Center for Digital Research in the Humanities at the University of Nebraska-Lincoln: <http://www.whitmanarchive.org/>
2. Traubel H, and Whitman W. *Intimate with Walt: Selections From Walt Whitman's Conversations with Horace Traubel, 1888-1892*. Ed. Schmidgall G. Iowa City: University of Iowa Press, 2001: 39.
3. E. Folsom *Walt Whitman's Native Representations*. Cambridge: Cambridge University Press, 1994: 102-148.
4. Whitman W. *Leaves of Grass and Other Writings*. 1855. Ed. Michael Moon.

Norton Critical Editions. New York: Norton, 2002: 2-506.

5. Miller N. Emerson's 'Each and All' Concept: A Reexamination. *New England Quarterly* 1968; 41.3: 381-392.
6. The Walt Whitman Collection, Box 1, Folder 22. Printed Ephemera from 1887. Yale Collection of American Literature, Beinecke Rare Book and Manuscript Library.
7. Erikkila B. The Poetics of Reconstruction: Whitman the Political Poet after the Civil War. In: Moon M, ed. *Leaves of Grass and Other Writings*. Norton: New York, 2002: 890-900.
8. The Walt Whitman Collection. Box 1, Folder 4. Letter to James Osgood, 1881. Yale Collection of American Literature, Beinecke Rare Book and Manuscript Library.
9. Meehan SR. *Mediating American Autobiography: Photography In Emerson, Thoreau, Douglass, and Whitman*. Columbia: University of Missouri Press, 2008: 207-210.
10. Mason JB. Walt Whitman's Catalogues: Rhetorical Means for Two Journeys in 'Song of Myself'. *American Literature* 1973; 45.1: 34-49.
11. Vendler H. *Poets Thinking: Pope, Whitman, Dickinson, Yeats*. Cambridge: Harvard University Press, 2004: 47.
12. Grossman A. The Poetics of Union in Whitman and Lincoln: An Inquiry toward the Relationship of Art and Policy. In: Moon M, ed. *Leaves of Grass and Other Writings*. Norton: New York, 2002. 872-889.
13. Vendler H. *Invisible Listeners: Lyric Intimacy In Herbert, Whitman, and Ashbery*. Princeton: Princeton University Press, 2005: 37.
14. Reynolds DS. *Walt Whitman's America: A Cultural Biography*. New York: Knopf, 1995: 285.

IMAGE SOURCES

FIGURE 1: Kenny, Bartlett F. "Walt Whitman." Albumen print, 1881. Online image. The Library of Congress, Prints & Photographs Division.

FIGURE 2: Gardner, Alexander. "Confederate Dead at Devil's Den Gettysburg." July 1863. Online image. Wikimedia Commons.

FOOTNOTES

1. To clarify, in *Leaves of Grass*, the image included was actually an engraving based on a photograph. At the time, it was not possible to effectively transfer a daguerreotype into print.



ALEX TORRES

Born and raised on the border of Illinois and Wisconsin, Alex Torres is a junior at Stanford University pursuing a BA in English Literature and a Minor in Spanish. He studies American poetry, and has conducted archival research on Walt Whitman, Emily Dickinson, Américo Paredes, Jean Toomer, and Allen Ginsberg. A recipient of the Cantor Scholars Award, Alex was the first undergraduate to curate an exhibition at the Cantor Arts Center. Alex is also a Mellon Mays Undergraduate Fellow and a former Chappell Lougee Scholar. He recently presented research on poetry and music at the American Literature Symposium in San Antonio. For the 2015-2016 school year, Alex was selected to receive the Deans' Award for Academic Achievement.

New Atheism: Constructing a New Nonbelieving Identity

Canton Winer

Fordham University

Since the height of the Cold War, atheists have been one of the United States' least-trusted groups. Exclusionary attitudes toward atheists are strong, often outpacing prejudices held against Muslims, possibly America's next least trusted religious identity. Polling shows that Americans were more likely to say they would not vote for an atheist than any other group.

Widespread distrust of and intolerance toward atheists is a commonly-accepted prejudice in much of American public life. Some have blamed this bias on factors including a simple lack of critical mass of nonbelievers, the United States' supposed "Christian" foundations, or even laws explicitly biased against those without religion. Upon examination, a likelier explanation rests on the nonexistence of an identity-based atheist movement in the United States. This absence, however, may soon be filled by "New Atheism," a nascent social movement defined by a more aggressive stance toward defending atheism and secularism, paving the way toward a new, more widely-accepted nonbelieving identity.

While an increasing number of Americans identify as "nones" (that is, as having no religious affiliation), there is a marked absence of atheists in the political sphere. A 2008 survey of 54,461 Americans found that 15.0% had no stated religious preference or were atheist or agnostic — an increase of over 80% from 8.2% in 1990 [1]. Yet Rep. Pete Stark (D-Calif.) left Congress in 2012 as its only self-identifying atheist, leaving publicly declared atheists entirely absent in Congress [2].

Pew Forum and CQ Roll Call released a report in 2010 on "The Religious Composition of the 112th Congress" [3]. The study compares the religious composition of the American public with that of Congress. The greatest disparity between the religious makeup of Congress and the people it represents is in the percentage of the "unaffiliated" (i.e. "those who describe their religion as atheist, agnostic or 'nothing in particular'"). While slightly over 16% of Americans identify as "unaffiliated," not a single member of Congress did.

Atheophobia (fear of or intolerance toward atheists) is a widespread and acceptable prejudice in much of American public life. A slew of factors has contributed to atheists' failure to gain acceptance from the American public. The reasons are not simply lack of numbers, the United States' status as a so-called "Christian Nation," or even a legal code biased against nonbelievers. Instead, these factors — and others — contribute to one larger cause: the absence of an identity-oriented atheist movement in the United States. But this movement may be emerging — and "New Atheism," a fledgling social movement that favors an overtly confrontational brand of atheism and secularism, may be planting the seeds.

It should be noted that I am using a broad definition of "atheism." The issues discussed in this essay relate to a broad nonreligious/nontheistic population, including atheists,

agnostics, humanists, skeptics, and those who simply find religion unimportant or irrelevant. David Niose, former president of the Secular Coalition for America, refers to this collection of populations collectively as "Secular Americans," though perhaps this label is even a bit clumsy [4]. Even within these nonbelieving groups, the exact definitions of these terms are contested.

In order to understand why saplings of an American atheist identity movement only began to grow in the 21st century, we must first examine the roots of the nation's widespread distrust of nonbelievers. It is easy to point to the nation's "Puritan roots," but such an explanation is deceptive. This explanation implies that Americans have widely distrusted and disliked nonbelievers since the earliest days of the Republic, which would seem to suggest that there has been a downward trend in anti-atheist sentiment since then.

As the bombs of WWII faded into a world of two superpowers, the United States was left in a tense relationship with the USSR which curdled into the Cold War. As with many conflicts, each side took to drawing extreme caricatures of the other. In the US, the Soviet Union became defined primarily by three attributes: atheism, totalitarianism, and communism [10]. As noted by Thomas Aiello [10], this rhetoric "fostered a public belief that no nation could positively engage with a counterpart perceived by so many as evil. Popular Christianity became the zenith of popular culture."¹

With the help of alarmist rhetoric from figures like Senator Joseph McCarthy (R-Wisconsin), "godless Communist" became a common epithet. More importantly, however, as Americans increasingly perceived Communism as a threat to the American way of life, the United States' self-image become increasingly tied to Christianity [10]. The choice between Americanism and Communism was also a choice between Christianity and atheism

in popular rhetoric in the mid-20th century. Because Christianity and Communism were seen as irreconcilable, and because the notion of the United States as a fundamentally Christian nation gained traction, it was not much of a stretch to label atheists (as well as Communists) as un-American, placing nonbelievers outside the popular definition of American citizenship.

Collective resistance to Communism was closely intertwined with resistance against atheists, laying the groundwork for an aggressive anti-secular push from the Religious Right beginning in the latter half of the 20th century. Nationalism and evangelical Christianity became inseparable, and powerful public figures like Reverend Billy Graham were eager to capitalize on this blend of patriotism and religiosity. “It is a battle to the death — either Communism must die, or Christianity must die, because it is actually a battle between Christ and anti-Christ,” Graham wrote in 1954 [13]. Not only were Christianity and Communism at battle, but Christianity — according to Graham and other evangelicals — was the key to victory over the Soviet Union. “There is only one antidote for the poisonous venom of Sovietism, and that is the truth of the gospel of Christ,” he wrote [13]. “The greatest and most effective weapon against Communism today is to be born again Christian.” These strong, atheophobic exhortations came from the mainstream as well as the fringes. Frederick Brown Harris, Chaplain of the Senate from 1942-1947 and 1949-1969, wrote in 1954 about “the hideous face of atheistic world Communism... the most monstrous mass of organized evil, that history has known” [14].

It can be argued that the need for a strong, atheist identity movement did not exist until the anti-Communist, anti-atheist hysteria of the 1950s. Yet this hysteria precluded the formation of such a movement. With Christianity a *de facto* prerequisite for patriotic citizenship, raising the very argument of accepting nonbelievers could easily earn advocates of the nonreligious the labels of “godless Communist,” “un-American,” and even “enemy of the State.” Once the Religious Right had the chance to drive the course of public discourse on religion in the United States, it had little incentive to ease the reins.^{II&III}

The animosity toward atheists that was constructed in the 1950s may explain the current dearth of nonbelievers serving openly on Capitol Hill. The lack of openly-atheist lawmakers is likely not due to the fact that nonbelievers simply do not run for office, but instead a result to public distrust of atheists, forcing atheists considering a run for office to either conceal their beliefs or give up running at all. In fact, it is reasonable to assume that Congress contains closeted atheists who remain nonbelievers privately, but demur publicly for the sake of remaining politically viable.

The reason for this hidden disbelief is clear: Atheists are one of the least trusted groups in America, and are less likely to be accepted — both publicly and privately — than most other ethnic, religious, and other minority groups [17]. Exclusionary attitudes toward atheists often outpace prejudices held against Muslims, possibly America’s next least trusted group. It is no surprise, then, that publicly identifying as an atheist is politically toxic; a 2012 Gallup poll revealed that Americans were more likely to say they would not vote for an atheist than any other group [18]. Forty-three percent said they would not vote for an atheist, followed by 40% who said they would not vote for a Muslim and 30% who said they would not vote for a gay or lesbian candidate.^{IV}

Although political office is hardly the only path toward



Figure 1. Religiosity and patriotism were strongly fused together during the Cold War.

national prominence, the general public’s animosity toward atheism has undoubtedly contributed to the lack of a coherent, cohesive, and powerful national voice from within American atheism. As a result, the Religious Right has enjoyed relatively unchallenged dominance in the forging of atheist identity. As with all identities, atheist identity is constructed both by insiders and outsiders. In other words, atheist identity is formed both by believers and nonbelievers. Atheist identity formation, however, has been somewhat stunted due to the lack of significant atheist organizations (in term of membership, funding, and political clout) in the United States. The vast majority of atheists do not belong to any organized atheist group [17]. Most atheists also do not attend meetings with other atheists, though there are “atheist churches” sprouting up around the country [20]. As such, atheist identity formation within the atheist community has largely taken place on an individual basis. With little organized opposition from atheists, the Religious Right has had relatively free reign on the construction of atheist identity in the U.S. The Religious Right — with powerful institutions like Moral Majority, Christian Coalition of America, The Fellowship, and others — was able to dominate conversations on disbelief since no equally powerful atheist groups existed to counter them. Figures of the Religious Right like Pat Roberston, James Dobson, and Jerry Falwell went unmatched from atheists, essentially giving the Religious Right unrivaled freedom to construct atheism without the consent or input of atheists themselves.

This negative construction of atheist identity has compounded the fact that being an atheist marks an individual as an outsider in the United States. Atheists are a fringe group in the country not because of their numbers (or lack thereof), but because of significant discomfort with atheists among the general public [21]. It is worth noting that only 1.7% of Americans identify as Jewish, but 4% identify as atheist or agnostic, and yet Jews are widely (though certainly not universally) accepted around the country and are actually *over*represented in the public sphere compared to the population at large [22]. Atheists are seen as the “Other” in America, and many Americans believe atheists to be “immoral” [17]. They are also often seen as an affront to the “Judeo-Christian tradition” which supposedly defines the United States [17]. Following this line of thinking, many Americans think that atheists are “un-American.” Atheism, in this view, runs

counter to the very fabric of the nation, making atheist identity — as Erving Goffman [23] would call it — a “spoiled identity.” As a result, a great deal of constructing atheist identity in the United States has revolved around combatting and managing a stigmatized identity [24].

Many atheists, however, choose to manage their identity by concealing it. Because atheists are not a visible minority (i.e. nothing “gives away” atheist identity), they are not forced to confront these issues as other groups — including blacks, Italians, and to some extent LGBT people, and others — have historically done. As a result, many nonbelievers may simply find it more convenient to conceal, or be reserved about, their beliefs. Particularly in the case of atheists in politics, these individuals will only face prejudice if they out themselves, so there is actually an incentive to keep their atheism private. The issue is self-reinforcing. Atheists will only be accepted if they out themselves, but if they out themselves, they won’t be accepted.

The challenge of “outing” oneself as atheist is heightened by the reality that, like many advocates for marginalized groups, atheist advocates also face the challenge of being perceived as aggressive and pushy when publicly defending their beliefs. Many religious Americans perceive public atheist sentiment as an assault on their own beliefs — if not an outright attack on the founding principles of the country — and react very negatively to public professions of atheism. As David Niose [4] highlights, atheists who vocally express their views on religion commonly receive the label “militant,” yet Christians who even more aggressively espouse their religious beliefs rarely receive this same distinction, instead earning the label “evangelical” or “devout.” As a result, many atheists simply lie low, “often identifying by default with a religion they don’t believe and don’t practice” [4].

Together, these factors have stifled the creation of a coherent, vocal, and national atheist movement in the United States. “We are where gays were at the time of Stonewall,” Lori Lipman Brown, director of the Secular Coalition of America (a lobbying group based in Washington), said in an interview with *The New York Times* in 2008 [24]. “And the thing we have in common with gays back then is that day to day you’re hidden,” Brown said. “If you make the decision to come out, you’re treated very badly.” Many individuals simply remain “closeted” in their atheist identity because of the social repercussions that accompany publicly declaring oneself an atheist.

The relative scarcity of strong atheist support groups may reveal another reason why electing atheists to public office is so difficult in the United States: atheism remains a dormant political cause. Atheist lobbying groups pale both in number and in clout when compared to religious groups. Additionally, the atheist constituency is difficult to activate simply because atheists are hard to locate. Unlike blacks, Hispanics, gays and lesbians, atheists “do not reside visibly in certain neighborhoods [24]. Atheists also do not come together collectively to worship, adding to the challenge of locating them.

Possibly the most significant obstacle to atheist collective

action, however, is that atheists do not have — or at least have not yet found — anything to coalesce around. The atheist identity is largely composed of things that the group *does not* believe, not what they do believe. There is currently no banner for American atheists to march under. Interestingly, the issue that is probably most pressing for atheists in the United States is probably confronting stigmatization of nonbelievers and challenging political and social discrimination. As I said, however, many atheists choose to deal with this stigmatization by concealing their disbelief instead of by confronting anti-atheist beliefs.

In addition to the problem perpetuated by closeted atheists, many atheists “maintain a diplomatic silence” about their beliefs because they “don’t want to offend our friends and neighbors” [25]. While this might seem like a private issue, Daniel Dennett [25] (a major figure in New Atheism) argues that “the price is political impotence.” Private silence amounts to public silence, and many “[p]oliticians don’t think they even have to pay us lip service, and leaders who wouldn’t be caught dead making religious and ethnic slurs don’t hesitate to disparage the ‘godless’ among us” [25]. In fact, many politicians see atheist-bashing as a “low-risk vote-getter,” largely since most atheists maintain pursed lips when religion appears in conversation. Simply identifying

oneself as atheist and breaking the silence, Dennett argues, will lead to greater political power for the atheist community.

All of this being said, the 21st century holds a fair amount of promise for American atheists. In the past, atheists have largely advanced the cause for equality through the judiciary.^v Social equality, however, cannot be achieved by way of the gavel.

Instead, the nonbelieving community requires a coherent identity-based social movement, and there are several factors which indicate that the nonbelieving community may finally be ready to begin forming this movement.

In many ways, the struggle for atheist rights is comparable to the struggle for lesbian and gay rights. While the courts certainly played a role in the gay and lesbian rights movement’s success, that movement has been primarily waged as an identity movement, not through a series of legal briefs. Because of significant social pressure to keep one’s atheism private, many atheists do not “come out of the closet,” even to their families. Gays and lesbians (though they still have not achieved full social and legal equality in the United States) were largely in a similar position of self-imposed silence mere decades ago. Public opinion shifted, slowly but surely, as more lesbian and gay individuals “outed” themselves. Americans began to realize that they knew people who were gay or lesbian in their workplace, in their circles of friends, and in their homes. Self-enforced silence gradually gave way to gay pride, advancing the cause for equality in a way that court cases simply cannot.

Identity is the source of the successes of the lesbian and gay rights movement. By moving the conversation from the closet to the public square, lesbians and gays reframed what it means to identify as lesbian or gay. Instead of attempting to conform to the norm, gays and lesbians proclaimed their “Gay Pride,” and demanded that American society evolve to incorporate their

“This negative construction of atheist identity has compounded the fact that being an atheist marks an individual as an outsider in the United States.”

identity, not the other way around.

No such “Atheist Pride” movement exists. Similar to gays and lesbians, nonbelieving Americans can hide their identity — there is no physical marker of one’s minority status as an atheist, so it is unlike being black, Hispanic, or female. In fact, atheist identity is even easier to hide than gay and lesbian identity. Closeted lesbian and gay individuals must either conceal their romantic involvements or suppress their sexual feelings; closeted atheists merely keep their mouths shut during certain kinds of conversations, as even failing to regularly attend church does not necessarily mark an individual as an atheist.

Some may note that atheist organizations have existed in the United States since the early 20th century, including the American Humanist Association (AHA); American Atheists; and the Freedom From Religion Foundation, to name a few. This is true, but these organizations have lacked the membership and the power to be a significant player in national politics. Furthermore, these organizations had not, for the most part, been identity-oriented. For example, the AHA — the oldest, and possibly the most prominent organization of nonbelievers in the United States — had focused more on what David Niose calls “general progressive issues, such a reproductive rights, peace, church-state separation, and human rights” [4]. In fact, when Herb Silverman founded the Secular Coalition for American (SCA) in an effort to create a coalition of various groups that could grow into an identity-oriented movement, the AHA initially refused to join, as they found the SCA’s approach potentially too brazen [4]. Even relatively powerful groups which fought against the anti-atheist efforts of the Religious Right, such as the People for the American Way (PFAW) and the American Civil Liberties Union (ACLU), have been quick to emphasize any religious connections, a tacit acknowledgement that secular causes lack legitimacy without religious support, and hardly a resounding cry for atheist pride [4].

Beginning in the 21st century, however, a group of “New Atheists” may be sowing the seeds of a true atheist identity movement. Niose [4] credits the 2000 election of George W. Bush as nonbelievers’ “wake-up call.” Many nonbelievers, Niose claims, were incensed by Bush’s religion-laced public statements, his conservative stances on social issues, and the generous access fundamentalist religious leaders had to the White House under his administration. The attacks on September 11th, 2001 can also be seen as a tipping point for American atheism. After the attacks, the country’s atheists became much more vocal. Many nonbelievers were frustrated by “God is on our side” rhetoric in the face of an attack that they felt was fueled by the same logic — surely the suicide bombers also thought God was on *their* side. David Silverman, president of American Atheists, says that the September 11 attacks inspired many formerly closeted atheists to speak out. “Most people know atheists now,” he says [26]. “They knew them before, but didn’t know they were atheists.”

New Atheism is possibly the most important outgrowth of the Bush era for American atheists. The origins of the term “New Atheism” are unclear; Sam Harris’ 2004 book, *The End of Faith: Religion, Terror, and the Future of Reason*, is often cited as the official “beginning” of New Atheism, though the term appears to have been coined by *Wired* magazine in 2006 (Pharyngula Wiki n.d.). “New Atheism” was at first used mostly as a pejorative used by theists to describe vocal atheists, intending to dismiss these outspoken nonbelievers as an irrational, ill-conceived fad (Pharyngula Wiki n.d.). Those who were described using the term



Figure 2. New Atheists may be laying the groundwork for an identity-oriented movement.



Figure 3. New Atheism has introduced a more aggressive, unapologetic brand of atheism to the American public.

were at first resistant to it, and while some continue resist it, many have since embraced the New Atheist label (Pharyngula Wiki n.d.).

The New Atheists brought a new, much more outspoken approach to public discourse on atheism. Rather than taking the soft, diplomatic approach historically employed by the AHA and other organizations, which were quick to emphasize ties with the religious community and advocated on broader, less controversial issues, the New Atheists forcefully criticized religion and belief in God not only as erroneous, but also as socially harmful. For example, even the title of Christopher Hitchens’ best-selling 2007 book, *God Is Not Great: How Religion Poisons Everything*, makes his distaste for religion quite clear, and Hitchens calls religion “violent, irrational, intolerant, allied to racism, tribalism, and bigotry, invested in ignorance and hostile to free inquiry, contemptuous of women and coercive toward children” [27]. Richard Dawkins wrote in his 2006 bestseller *The God Delusion* that, “The God of the Old Testament is arguably the most unpleasant character in all fiction: jealous and proud of it; a petty, unjust, unforgiving control-freak; a vindictive, bloodthirsty ethnic cleanser; a misogynistic, homophobic, racist, infanticidal, genocidal, filicidal, pestilential, megalomaniacal, sadomasochistic, capriciously malevolent bully” [28].

New Atheism, it should be noted, is somewhat amorphous and hard to define. New Atheists may have brought a sense of pride to some members of the nonbelieving community, but they

have not brought unity. Atheists remain divided on their reception of New Atheism, some finding the approach refreshing and liberating, others put off by its “shrill” or “angry” approach [29].

Certainly, the critics of New Atheism have a point. The gay and lesbian rights movement did not, after all, find success through blunt criticism of the group oppressing them, though prominent adversarial gay rights activists (like Larry Kramer) did exist. Gays and lesbians did not criticize straight communities as harmful to society, but instead simply said “we are here.”^{VI}

Furthermore, New Atheism’s forceful denunciations of religion reinforce the “militant atheist” trope which many American atheists are so careful to avoid. Still, this criticism may be missing the point. While the arguments of the New Atheists are unlikely to convince the devout, the devout are not New Atheists’ target audience. The New Atheists are not trying to convert the religious. They are attempting to activate the religious nones, agnostics, and other “soft” atheists into taking action and actively staking their claim in American identity *and* nonbeliever identity. They are laying the groundwork for an identity-oriented movement.

New Atheism also reorients the relationship between theist and atheist communities. Atheists have historically operated largely from a defensive stance. Instead of responding to attacks, acting essentially in a permanent state of damage control — as American atheism has done in the past — New Atheism attacks religion. New Atheism brings a proactive energy to the community that places religious communities on a newfound defense. No longer merely managing a damaged identity, the New Atheists have gone on the offensive, challenging the very *prima facie* status that the Religious Right worked so hard to construct and uphold. Until the rise of New Atheism, atheist identity in the United States was largely defined by a tacit acceptance of atheism’s inferiority to religion. Instead of rejecting the premise of challenges placed against them, American atheists often responded to them. If someone calls you stupid, and your response is “I’m not stupid,” then we are still talking about stupid. New Atheism does not merely respond to the criticisms of atheism; it reframes the conversation and asks religion to justify itself instead.

In a sense, perhaps New Atheism is nonbelieving America’s Stonewall. While New Atheism is not a singular, defining event as Stonewall was, it may have mobilized the atheist community in a similar (though less potent) way. The violent, spontaneous protests that erupted following a police raid on the Stonewall Inn on June 28, 1969 were the result of New York City’s gay community proclaiming that enough was enough. Rejecting the assimilationist, non-confrontational approach of early homophile groups, the Stonewall demonstrations launched a confrontational, identity-oriented Gay Liberation movement. The antitheistic books and rhetoric of New Atheism may be the momentary “violent” outburst that American atheists need. It seems unlikely that the path to wider public acceptance of America’s nonbelievers will be paved in fiery, antitheistic prose, but New Atheism may be the burst of flame that ignites a steadier, prolonged, and mature atheist identity movement in the United States. Stonewall gave

rise to the Gay Liberation, which in turn gave rise to a gay and lesbian civil rights movement. This movement realized that violent outbursts in the style of Stonewall would not bring sustainable change, but the gay and lesbian rights movement would not have existed without those same outbursts which it sought to avoid. New Atheism may play a similar role in stimulating a movement of its own.

Even under the assumption that New Atheism has laid the groundwork for an identity-oriented atheist movement, there are still several key unresolved issues that may impede the formation of this movement. Perhaps the most important is that atheism remains poorly defined. Even atheists themselves have widely differing views on what it means to identify as an atheist. In the strictest sense, an atheist is someone who believes that God does not exist. Yet, Pew Research found that “14% of those who call themselves atheists also say they believe in God or a universal spirit” [30]. A number of self-identifying atheists also identify as Catholic, Jewish, or some other religious identity (so-called “cultural Catholics” and “Secular Jews,” for example). Inversely, a large number of Americans who say they do not believe in God or a universal spirit do not identify as atheist. Seven percent of respondents said that they do not believe in God or a universal

spirit, yet only 2.4% said they were atheists [30]. Additionally, many atheists are uncomfortable critiquing religion, while others bluntly criticize “superstitious” belief systems. Twelve percent of agnostics and atheists, for example, told Pew Research in a 2014 poll that they felt religious influence was

decreasing, and they thought that was a bad thing [31]. Building a coalition among such a wide-range of identity and objective is daunting, and no clear resolution to these issues is currently apparent.^{VII}

The Internet may be a tool for atheism to overcome several of the obstacles it faces. Most notably, the Internet allows individuals to discuss their beliefs in a space that many find far more welcoming than the physical spaces of the workplace, home, and town hall. As noted by Peter M. Rinaldo, the Internet allows atheists to communicate and express their views without having to identify themselves [36]. Moreover, it also allows atheists to digitally congregate without paying dues or otherwise joining a national organization. This is a double-edged sword for American atheists. On one edge, it allows atheists to form virtual communities without the risk of social retribution, as atheists can operate anonymously online. It also allows communities that feel neglected in “mainstream” atheist discourse, such as women, to congregate in digital communities to make their voices heard. On the other edge, the Internet creates less of a need — or at least an immediately apparent one — for a strong, national atheist organization, thereby prolonging the social conditions which make the anonymity of the Internet appealing.

All of this goes to show that while American atheists have made significant strides in moving toward wider social acceptance, there is still a long way to go. Atheists remain fractured and disorganized, and internal struggles over identity and strategy are far from resolved. New Atheism may have broken the silence

“Beginning in the 21st century, however, a group of “New Atheists” may be sowing the seeds of a true atheist identity movement.”

and given life to a new national discussion on atheism, but it hardly presented a blueprint for the way forward. New Atheism brought candor, but not coherence. Without the formation a strong, identity-oriented social movement, American atheists look destined for continued political impotence and public distrust.

REFERENCES

- Kosmin BA, Keysar A. "American Religious Identification Survey (ARIS 2008)." Mar. 2009. Web.
- Wing N. "Here Are All The Atheists In Congress." *Huffington Post*. N.p., 19 Sept. 2013. Web.
- Miller T. "Faith on the Hill – The Religious Composition of the 112th Congress." *Pew Research Center's Religion & Public Life Project*. N.p., n.d. Web. 11 Dec. 2013.
- Niose D. *Nonbeliever Nation: The Rise of Secular Americans*. New York: Palgrave Macmillan, 2012.
- Arp R. *1001 Ideas That Changed the Way We Think*. New York: Simon and Schuster, 2013.
- Ingersoll RG. "Why I Am an Agnostic." 1896.
- Ingersoll RG, Farrell CP. *The Works of Robert G. Ingersoll*. New York: The Dresden Publishing Co, 1911.
- Leip D. "1912 Presidential General Election Results." *Atlas of US Elections*, 2012. Web.
- Newport F. "In U.S., Four in 10 Report Attending Church in Last Week." Gallup 24 Dec. 2013. Web.
- Aiello T. "Constructing 'Godless Communism': Religion, Politics, and Popular Culture, 1954-1960." *Americana: The Journal of American Popular Culture (1900-present)* 2005; 4.1: n. pag. Web.
- McCarthy, J. "Speech of Joseph McCarthy." Wheeling, West Virginia, February 9, 1950
- Newport F. "Public Opinion on the "4s" Through Recent History." Gallup. 6 June, 2014.
- Graham B. "Satan's Religion." *The American Mercury* 1954; Aug: 41–46.
- Harris FB. "World Communism Is Unveiled." *Sunday Journal and Star* 14 Mar. 1954: 38.
- Sherman R. "Vice President Bush Quote Regarding Atheists." RobSherman.com. N.p., 3 Mar. 2002. Web.
- Bulger M. "Unelectable Atheists: U.S. States That Prohibit Godless Americans From Holding Public Office." American Humanist Association. N.p., n.d. Web. 11 Dec. 2013.
- Edgell P, Gerteis J, Hartmann D. "Atheists as 'Other': Moral Boundaries and Cultural Membership in American Society." *American Sociological Review* 2006; 71.2: 211–234.
- Jones JM. "Atheists, Muslims See Most Bias as Presidential Candidates." Gallup (2012): n. pag. Web.
- Stedman C. "Barney Frank: Why I Didn't Talk about My Atheism While Serving in Congress." Faithist. N.p., n.d. Web. 2 Dec. 2014.
- Sanburn J. "Atheist 'Churches' Gain Popularity—Even in the Bible Belt." *TIME* 2014; July 24: 52–53.
- Guenther KM, Mulligan K, Papp C. "From the Outside In: Crossing Boundaries to Build Collective Identity in the New Atheist Movement." *Social Problems* 2013; 60: 1-19.
- Pew Research Religion & Public Life Project. *Religious Landscape Survey*. N.p., 2008. Web.
- Goffman E. *Stigma: Notes on the Management of Spoiled Identity*. New York: Simon and Schuster 1963.
- Freedman SG. "For Atheists, Politics Proves to Be a Lonely Endeavor: [Metropolitan Desk]." *New York Times*, Late Edition (East Coast) 18 Oct. 2008: A.21. Print.
- Dennett DC. "The Bright Stuff: [Op-Ed]." *New York Times*, Late Edition (East Coast) 12 July 2003: A.11. Print.
- Blake J. "Four Ways 9/11 Changed America's Attitude toward Religion." CNN Belief Blog. N.p., 3 Sept. 2011.
- Hitchens C. *God Is Not Great: How Religion Poisons Everything*. New York: Twelve, 2009.
- Dawkins R. *The God Delusion*. Reprint edition. Boston: Mariner Books, 2008.
- Fincke D. "You Might Be An Atheist Even If You Hate The New Atheists." *Camels With Hammers*. N.p., 13 Feb. 2011. Web. 4 Dec. 2014.
- Lipka M. "5 Facts about Atheists." Pew Research Center. N.p., 23 Oct. 2013.
- Lipka M. "Is Religion's Declining Influence Good or Bad? Those without Religious Affiliation Are Divided." Pew Research Center. N.p., 23 Sept. 2014.
- Marcotte A. "Atheism's Shocking Woman Problem: What's behind the Misogyny of Richard Dawkins and Sam Harris?" *Salon*. N.p., 3 Oct. 2014.
- Shores M. "Will 'New Atheism' Make Room For Women?." *Ms. Magazine* 1 Oct. 2010: Web.
- McKenzie S. "Why Religion and Atheism Need Smart Women." *The Sydney Morning Herald*. Aug 5, 2010.
- Oppenheimer M. "Will Misogyny Bring Down The Atheist Movement?" *BuzzFeed*. N.p., 11 Sept. 2014. Web. 16 Nov. 2014.
- Rinaldo PM. *Atheists, Agnostics, and Deists in America: A Brief History*. Briarcliff Manor, NY: DorPete Press, 2000.

IMAGE SOURCES

FIGURE 1: Riggs, David. "Attention Lunatic Atheists." 12 March 2005. Online image. Flickr.

FIGURE 2: Fibonacci Blue. "Atheists at the Twin Cities Pride Parade 2011." 26 June 2011. Online image. Wikimedia Commons.

FIGURE 3: HeatherKay7. "Minnesota Atheists Billboard." 24 January 2012. Online Image. Wikimedia Commons

ACKNOWLEDGEMENTS

The author thanks Dr. Micki McGee and Dr. Leonard Casutto for their guidance and mentorship. This research was financially supported by a Fordham University Undergraduate Research Grant.

ADDENDUM I

Christianity was framed not only as under attack from "Godless Communism," but also as the strongest weapon in the American arsenal against it. On February 9, 1950, Senator Joseph McCarthy (R-Wisconsin) made this message clear in a Lincoln's birthday address to the Women's Republican Club of Wheeling, West Virginia:

Today we are engaged in a final, all-out battle between communistic atheism and Christianity. The modern champions of communism have selected this as the time. And, ladies and gentlemen, the chips are down — they are truly down [11].

While most political rhetoric was not as fiery as McCarthy's, the public widely supported him. A January 1954 Gallup poll placed McCarthy's favorable rating at 46% and his unfavorable rating at 36%, giving him a net favorable of +10 [12]. Though his favorable ratings would drop irreparably into the negatives by the spring of that year, McCarthy's wide favorability in public polling earlier that year signaled, implicitly if not also explicitly, a widespread acceptance of the Senator's atheist bashing.

ADDENDUM II

While the Communist label is not thrown around as carelessly today as in the 1950s, the Religious Right was keen to capitalize on the distrust of atheism and secularity (i.e. the separation of government and religion), and the anti-atheist sentiment which frothed in the 1950s continued throughout the 20th century and

into the 21st century. In 1980, the so-called “Moral Majority” (a prominent conservative Christian group headed by Jerry Falwell) helped to carry Ronald Reagan to victory over incumbent President Jimmy Carter. In 1987, Vice President George H. W. Bush told reporters, “No, I don’t know that atheists should be regarded as citizens, nor should they be regarded as patriotic. This is one nation under God” [15]. The statement was made one week after he announced his candidacy for the Republican presidential nomination, and though the entire Chicago political press corps was present, only Robert Sherman of American Atheist Press apparently found the comment newsworthy. His son George W. Bush was elevated to the White House 13 years later after running on an explicitly evangelical platform. Even today, controversies over civic displays of nativity scenes, comments made by reality TV stars, and other seemingly insignificant contestations continue to evoke Christian Nation sentiment.

ADDENDUM III

In addition to the development of these attitudinal and rhetorical hostilities, it bears mentioning that there has been some codification of prejudice against atheists in the United States, and there are some laws that directly discriminate against atheists. For example, atheists are explicitly banned from serving in higher office in some states, and seven states even have such bans written in their state constitutions [16]. The North Carolina State Constitution, for example, states that:

The following persons shall be disqualified for office:
First, any person who shall deny the being of Almighty God. (North Carolina State Constitution, Article VI, Sec. 8)

That seven states have language resembling the above in their state constitutions is undoubtedly an indication of prejudice toward and distrust of atheists in the United States. Perhaps even more telling, however, is that these ordinances are often unnoticed, unchallenged, and uncontroversial. It is extremely difficult to find reputable publications that discuss this issue; a simple Google search reveals that no major American news outlets have published articles about these laws.

ADDENDUM IV

Former Rep. Barney Frank (D-Mass.) demonstrates this reality quite clearly. Frank, who served openly as a gay man in Congress, announced his nonbeliever status only after retiring from office. His example is particularly telling, as he came from a fairly liberal district, yet felt more comfortable coming out publicly as an openly-gay man than as a nonbeliever. (It should be noted that Frank, who is of Jewish descent, has said that he was reluctant to “explicitly disavow any religiosity [while in Congress for fear that] it could get distorted into an effort to distance myself from being Jewish — and I thought that was wrong, given that there is anti-Jewish prejudice” [19]. This argument is somewhat suspect, given that Frank came out as atheist shortly after leaving office. Frank remained a person of Jewish heritage; the only thing that changed was his status as an elected official.)

ADDENDUM V

See *McCullum v. Board of Education* Dist. 71, which found religious instruction in public schools unconstitutional; *Torcaso v. Watkins*, which found unconstitutional the requirement that applicants for public office must swear that they believed in the existence of God; *Engel v. Vitale*, which found school-sponsored prayer unconstitutional; *Allegheny County v. ACLU*, which found a nativity scene displayed inside a government building unconstitutional; and many others.

ADDENDUM VI

It should be noted, however, that this is not universally true. Certainly, the life-and-death issues of funding for HIV/AIDS research also changed the stakes and brought many people out of the closet and into the streets. For many in the gay community, the issue was survival, not just pride, and rhetoric sometimes became bitter and urgent in the climate of the AIDS crisis.

ADDENDUM VII

American atheism also faces a gender problem. An overwhelming 67% of self-declared atheists in the United States are male [30]. The prominent figures in New Atheism are all male, and many of them have come under intense fire from feminist circles, who claim that Sam Harris, Richard Dawkins, etc. hold misogynist attitudes [32]. Furthermore, their self-presentation is stereotypically masculine in its aggression, which some critics feel may be driving women away from New Atheism [33]. Journalist Sarah McKenzie [34] suggests that atheism’s gender problem may be more structural, claiming that girls may actually be socialized to keep their distance from the atheist label. “After all, girls are taught to be sensitive and emotional, to not cause trouble or be particularly forthright with their opinions,” McKenzie wrote [34]. “Women who dare to be aggressive or outspoken are often labelled as hysterical harpies, not worthy of being listened to and impossible to take seriously. We should hardly be surprised that some women might be reluctant to come out as atheists.” Some critics allege that when women attempt to discuss their perceptions of sexism within the atheist community, they are met with hostility. “For the past several years, Twitter, Facebook, Reddit, and online forums have become hostile places for women who identify as feminists or express concern about widely circulated tales of sexism in the movement,” Mark Oppenheimer wrote [35]. New Atheist figures like Richard Dawkins, Oppenheimer continued, have “alienated many women — and men — by belittling accusations of sexism in the movement.” Oppenheimer blames this problem, in part, on American atheism’s “roots in very male subcultures.” Women (such as Madalyn Murray O’Hair, who founded the nonprofit group American Atheists in 1963) were certainly involved, but they were exceptions, not the rule. Until the “boys club” in American atheism takes this problem seriously, charges of sexism within the nascent movement will distract from the movement’s overall cause.



CANTON WINER

Canton Winer graduated in 2015 from Fordham University, where he majored in Sociology and American Studies. Originally from Florida, Canton is currently teaching SAT, ACT, and SSAT test prep in Shanghai. Upon leaving China, he intends to pursue a PhD in Sociology.

Dance and the Postmodern Sublime in Jacques Demy's *The Young Girls of Rochefort* (1967)

Carlos Valladares

Stanford University

This paper analyzes the effect of the choreography in Jacques Demy's French New Wave musical-film *The Young Girls of Rochefort* (1967). Critics have dismissed the film because of what they perceive as "sloppy" or inaccurate choreography. This paper finds that not only is the unpolished dancing a conscious reflection on the director Jacques Demy's part, it is a necessary element to the film's internal structure. The paper references modern-day critics such as Jonathan Rosenbaum to clarify the purpose of the film's imperfect choreography, discussing key sequences from the film at length. This paper ultimately finds that Demy's film provokes unusual, indefinable feelings in the viewer that align themselves with postmodernist thinker Jean-François Lyotard's idea of the "postmodern sublime." The film's postmodern bent—of instability, of bizarre juxtapositions, of a lack of a fixed cohesion—unnerves the trained musical viewer who has built certain criteria by which they judge a "successful" musical film. Such a reading of *The Young Girls of Rochefort* is crucial to appreciate the novel combinations of dreams and fantasy in Jacques Demy's larger body of work.

What does it mean to dance in film? When characters stop talking and start moving, what are they communicating to us, their audience? These questions are central to the movie musical, a popular genre of film where logic sits in the passenger's seat and is ceded to illogical bursts of song and dance. Most musical film scholars (including Rick Altman, Jane Feuer, and Ethan Mordden) have discussed the lasting legacy of the musical in terms of the same American movies: *An American in Paris* (1951) and *Singin' in the Rain* (1952), the Radio-Keith-Orpheum (RKO) pictures starring Fred Astaire and Ginger Rogers. However, almost no existing histories bring up the French musicals of Jacques Demy. Those that do, furthermore, are not willing to seriously discuss his exemplary 1967 musical *Les Demoiselles de Rochefort* ("The Young Girls of Rochefort").

A 2011 British Film Institute guide of 100 great movie musicals has a blurb for *Rochefort* which spends more time apologizing for the movie's more obvious flaws than analyzing them:

"*Les Demoiselles* starts promisingly, as the travelers board the Rochefort-Martrou transporter bridge. Some of the other dance sequences, particularly the large group dances, are not so inspired...*Les Demoiselles* is not the masterpiece some (like Jonathan Rosenbaum, 1998) have claimed but it is nevertheless very good, and a fascinating hybrid." [1]

Perhaps no other element in *Rochefort* has attracted more critical dismissal than its dancing. Scores of contemporaneous reviews ridiculed the film's perceived "amateurish" style owing to the lack of finesse in its dance routines. Andrew Sarris of

The Village Voice wrote that the film "too often gets tangled up in its fancy footwork," suggesting that a need for a perfect, choreographic cohesiveness trumps the artist's desire to convey happiness through imperfection. [2] Gary Carey in the late 70s writes that "[*Les Demoiselles*] falls to pieces whenever anyone begins to dance." [3] In addition, influential critic Pauline Kael dismisses the film entirely in her most influential piece "Trash, Art and the Movies," calling it "a movie [that] demonstrates how even a gifted Frenchman who adores American musicals misunderstands their conventions." [4] Kael's suggestion—namely, that films in the American musical style should best be left to the Americans—is not only a reductive way of looking at a foreign film (which is not beholden to "understanding the conventions" of the genre from which they are drawing), it fails to appreciate the singular eccentricity of Demy's film, which goes beyond mere aping of American musicals or lousy footwork. It is self-consciously a disruption of normal movie musicals, and thus cannot be criticized in the traditional fashion (i.e., comparing it to other movie musicals of its kind). It demands of the viewer something more unusual and less intuitive: a rethinking of one's normal definitions of beauty and, above all, sublimity.

What can we learn from *Rochefort's* dance-aesthetics, its eccentric sheen? I hope to use this essay to deeply analyze the dances in *The Young Girls of Rochefort*. To come to understand why these dances serve their purpose in the film is to understand Demy's incredibly generous artistic vision, its fractured complexity, and the deeply unstable emotions it provokes in the viewer that align the film with postmodernist thinker Jean-François Lyotard's idea of a "postmodern sublime."



Figure 1. The opening credits.

THE OPENING SEQUENCE

To understand the film's dance sequences and how they contribute to the film's sublime effect, one need not look further than its opening moments. In the opening sequence, a couple of carnival workers make undramatic plans to cross a transport bridge. No words are exchanged, no smiles are cracked—for all the audience knows, the film is in the realm of documentary realism, since Demy shoots the scene with such plotless heft. Yet, due to Demy's signature style and also since the audience is more than aware what they're about to see is a musical, what grabs the trained musical viewer's attention is the scene's palpable silence, the lack of an immediate or punchy throat-grabber as is expected in the musical's opening. Compared to the one-two sucker punch of the opening to Demy's 1964 pop-opera *The Umbrellas of Cherbourg*—where *Melancholia* (the opening credit sequence) and *Exuberance* (the first scene set in a hopped-up garage) are quickly and economically tied to one another—*Les Demoiselles'* opening seems downright dull. Instead of loud music or singing, only the put-put-puttering of engines and slamming of doors registers on the soundtrack.

Then, the first piano pangs of Michel Legrand's score hit, coinciding with the opening credits. The carnival workers exit their vehicles, stretching their arms and legs in a playful manner. The stretches turn ever so slightly into roughly choreographed actions. The music makes it seem as though they are rehearsing an upcoming number. No sooner does the title credit appear on the screen than the workers begin an exciting but brief dance-step to the tune of the film's theme "La Chanson des Jumelles" ("Song of a Pair of Twins") (Figure 1).

The first dance (aboard the transport bridge) briskly establishes the general mood and tone of the dances in *Les Demoiselles*. It is a simple step, owing to the limited space (i.e., the tiny bridge-square) that the dancers have to move. As a result, the choreography is less bombastic, and therefore easier to master. Furthermore, the dancers are half out-of-step, their high kicks not completely reaching the top of the arc (Figure 1). To the regular eye, such inattention to choreographic pointedness can register as amateurish, potentially an effect the director did not intend.

However, to reduce the dances in *Les Demoiselles* to such a harshly critical degree overlooks their purpose in the larger structure of the film. To be sure, the logic of such a statement—that "lousy" dancing can amount to greatness—seems contradictory. However, I would like to posit that the kicks serve a higher aesthetic purpose than what first meets the eye. These deliberately "sloppy" steps are a realistic look at (and criticism of) traditional movie musicals. The dances in *Les Demoiselles* do not move the viewer through polished, kinetic virtuosity; their effect is of an earthier kind, where we see seemingly non-professional dancers attempting to dance in order to express larger, inexpressible emotions.

As a point of comparison, we can take up the classic Hollywood musicals, which have been canonized and held up as exemplars of the genre. The grueling rehearsals Fred Astaire and Ginger Rogers went through in order to achieve a sleek, unbroken-sweat perfection in their RKO musical numbers [5] are widely noted. Likewise, there are well-known anecdotes concerning choreographer Jerome Robbins's overworked dancer-actor-singers during the making of *West Side Story*. Their bodies and nerves were pushed to critical breaking points, all in the service of a filmed dance that was aesthetically crisp and perfect. Critics attribute the movie-musical's brilliance to the effortless manner in which people in the Hollywood musicals dance, sans flaw or fault. However, what these technically polished dances suppress is the *humanity* of human performance, the nobility of amateurish ballets. In films like *West Side Story* and the Astaire-Rogers films, though technically brilliant and accomplished, there is a divide between the audience and the actor that is widened by the actor's superhuman perfection and poise. By priding aesthetic perfection over an equally legitimate amateurish authenticity, critics like the ones who reject Demy's movie limit the collective understanding of the musical's capabilities. As Rosenbaum so perceptively puts it, "this quantitative aesthetic doesn't allow for the possibility that a musician with limited technique like Thelonious Monk might be a greater pianist than a virtuoso like Oscar Peterson" [6]. If we take his metaphor a step further, one can even favor the more "everyday", ostensibly amateurish feel of Demy's loose footwork over the pristine calculations of a Donen-Kelly or an Astaire-Rogers. To reiterate, this is not to establish preferences;

this is simply to explain that the choreography in *Les Demoiselles* is capable of unexpectedly dazzling the viewers in ways that require a fundamental rethinking of one's conception of a "successful" film musical.

In real life, the average Joe Schmo cannot hope to move with the balletic perfection of Gene Kelly or Leslie Caron. At our best, the regular folks can sashay with the buoyancy and imperfect grace of the characters in *Les Demoiselles*. Demy's world is a democratic one—anyone can dance—and as a result its deeper humanistic implications resonate with us on a more personal level. Furthermore, the imperfect dancing mirrors the film's aesthetic combinations of reality and fantasy. The amateurish quality of the dances prides human error as an aesthetic quality to bear. At the same time, however, there is still a fanciful element of fantasy contained in the musical—or, at the very least, a hyperstylization of normal reality. The blaring colors of the carnival workers' shirts; the Legrand score which penetrates the soundtrack without warning and at any given point in time; the never-ending stream of dancing that spills into the streets of Rochefort: all of these fantastical elements mesh brilliantly with the film's "realistic" vision of dancing.

MIXING IN TWO DANCE SEQUENCES FROM *LES DEMOISELLES*

Mixing (of various kinds) is the film's greatest aspect. Demy constantly juxtaposes disparate modes of thinking, artmaking, and filmmaking in ways that are new and invigorating. Demy's film is a curious oddity of mixed modes: American musical and French art film, studio Hollywood and French New Wave, fantastical fictions and philosophical truths. Where there is realism, there coexists fantasy. Where there is candy-coated happiness (the film's sunny romances), there also lurks horror (the serial-killer subplot). And, as the film's semi-tragic finale indicates, where there is successful romance (Solange finds Andy, Yvonne the café-owner reunites with Monsieur Dame), there are also crushing missed connections (Delphine misses her love, the sailor Maxence, by milliseconds).

It is with the dancing, however, that his "mixing" finds its greatest challenge. In the world of *Rochefort*, the dance never ends: characters and extras alike all move in choreographed routines, even when a scene doesn't specifically call for dance. Consider two sequences from *Les Demoiselles*: the introduction of the American in Paris, played by Gene Kelly, and Catherine Deneuve's walk down the street to meet her fiancée.

In the first, Solange (Dorléac), the brunette composer twin (Fig 2), accidentally trips into Andy Miller (Kelly) (Fig. 3), an American composer on holiday in Rochefort. They experience what seems to be love at first sight, but Solange cuts the meeting short and hurriedly runs in the opposite direction. Andy, who cannot contain his joy from this accidental encounter, runs up to people in the street and tells them of his experience: a young woman, a balloon-stand owner, two sailors, six modish young ladies in miniskirts, and a group of adolescent boys. With each next person, Kelly (the celebrated choreographer-dancer-actor who starred in many MGM musicals of the 1950s, including *Singin' in the Rain* and *An American in Paris*) seemingly improvises new steps on the spot.

What strikes the eye about this sequence is its portrayal of the *metaphysical*, the *atemporal*, and the *emotional release* of dancing. It is metaphysical in the sense that it demonstrates the purpose



Figure 2. Solange (Dorléac), the brunette composer twin.



Figure 3. Andy Miller (Kelly), an American composer on holiday in Rochefort.



Figure 4. Andy attempts to communicate his feelings in Franglish: "I'm in love, I am."

of movie-musical dancing at its most powerful. Words escape the individual when s/he tries to describe the feeling of love, so Andy turns to the realm of dance. This crucial step—away from language, the tangible, and the explanatory—is almost instinctively understood by everyone Andy crosses on the street. Problems do arise when Andy tries to communicate his feelings to the young woman as he sings, "I'm sorry *mademoiselle, mais je suis amoureux*" ("...but I'm in love, I am")—a display of Franglish that mirrors the utter bafflement of tourists in a foreign land. (Fig. 4). However, the language barrier is successfully trumped as meaning-packed words give way to the abstractions of dance, a transition mirrored by the song's ("Andy Amoureux") transition from a traditional musical ditty with lyrics to an instrumental with swelling, romantic strings. It produces a peculiar effect of sublimity that is crucial to understanding this film's peculiar logic.

It is atemporal because Demy collapses time in subtle ways. Not only does the world seem to stop for Solange and Andy, as evidenced by the hazy out-of-focus background that accentuates the two would-be lovers (Figs. 2 and 3), time, too, is frozen, functioning in a more pointedly abstract fashion than has been noticed in the film. One doesn't immediately take notice to a group of dancers already in the throes of their steps (Figure 5) whilst Andy is reading the notes off of the sheet music that Solange hurriedly left behind. It takes a few viewings of the movie for this slight but crucial background detail to be registered. The audience is so enraptured by Gene Kelly's surprise appearance that



Figure 5. Note how the dance has already begun in the background, even while Andy is still reading from Solange's sheet music.

they don't notice the dance has already begun in the background; Demy's CinemaScope widescreen-framing cleverly conceals this otherwise obvious detail. In the world of Rochefort, time has no beginning or end; it merely exists in a realm where dance itself has no temporal basis. In this case, one temporal strand (the dancers' generic joy) collides with the completely different strand of another person (Andy's timeless joy).

The scene is awash with emotion, a distinct concoction that merges Kelly's famous personality with Demy's directorial sensibilities. Though the scene harnesses Kelly's star power (his shiny grin, the perpetual spring in his step, his tap-dancing skills) to give it its primary appeal, it does not work on the same level as a Kelly musical like *An American in Paris* or *On the Town*. In these films, nothing out of the ordinary occurs in the world to his periphery. Kelly himself is the out-of-place, almost mystical source of joy that adds something magical and alive to the relatively banal world around him. In a typical Kelly scene, where he is the focus, the scene's milieu pivots around him. In Demy's world, however, it is Kelly who pivots around the scene's milieu. He is but one element of the scene, which also draws its appeal on its dizzying production design. Extras not only dance in subtle anticipation of Kelly's moves, they dance with the master himself. "Loud" pastel colors scream from the actors' costumes in such overwrought ways that not even Demy's mentors Vincente Minnelli and Stanley Donen would have gone so far. In a daringly pronounced metaphor, the world begins to sing and dance around Gene Kelly.

The subtle emotional power of the sequence derives from its rising action. During Solange and Andy's actual encounter, and just before Andy launches into his song, five couples, with their hands interlocked, pass by (Figs. 6, 7, and 8).

These couples are not dancing, but their intrusive presence in the diegetic space of the Solange-Andy encounter—bolstered by the audience's recognition of them as dancers from what we have seen of them in the film's second opening, the dance in the park-square—*implies* dance-like movements. Their entry signals a psychological extension of the powerful emotions Solange and Andy experience. Their banal chat works on one level of comprehension, but the Legrand score and the non-dancing dancers work on a deeper and unstable level: a level where words fail and movement is the only expression available for lovers. Thus, the audience comes to understand the less intuitive purpose of Demy's dancing: the gamete of emotions (love, sorrow, melancholy, happiness, exuberance) that his musical-films mystically pursue.

Demy's refusal to clarify his characters' emotions through conventional lyrical means signifies his self-conscious interest in a



Figures 6, 7, and 8. Couples passing by.

reaction rooted in feelings of the *sublime*. One's complex reaction to the Solange-Andy encounter (i.e., the romance collides in such a contrived and on-the-nose manner, and yet harmonizes with the peculiar world in which this romance is set) reflects Demy's philosophical sensibility that nothing in one's world is ever of one holistic whole. That is to say, every single scene is contaminated with differing traditions reflective of Demy's ardent cinephilia. *Les Demoiselles* is the astounding result of Demy's daring experimentation: he combines different sensibilities to forge new meanings, new ways to re-think not only how audiences perceive a musical, but how audiences perceive *any* genre film. He both conforms to and rebels against the traditions of a genre he very much admires: the movie-musical.

In the second scene, Delphine (Deneuve), the blonde ballerina twin, walks down the street for a rendezvous with her fiancée Guillaume (Jacques Riberolles). As she walks, scores of dancing sailors and similarly young girls pass Delphine, cavorting and playfully singing around her (Fig. 9). At one point, a pair of sailors picks her up from the ground for a spell, then placing her gently back to the earth to continue her walk (Fig. 10). All the while, Delphine acts nonplussed at the orgiastic fun happening around her; her face registers quiet bemusement, but nothing indicates that what she sees is in any way extraordinary to her. In other words, her reaction and our (i.e., the viewer's) reactions to the dancing are at exact opposites: we are not used to such lusciously indulgent flights of fantasy and so we are dazzled, but to Delphine this is merely an everyday walk along the streets of "boring, old" Rochefort. (The viewer must remember, of course, that the twin sisters want to leave Rochefort and escape what they view as an unbearably quotidian grind.)



Figure 9. Sailors and young girls pass Delphine, dancing and singing around her.



Figure 10. A pair of sailors picks Delphine up, before placing her back on the ground to resume her walk.

In this case, the lines between normal social behavior and flighty unrealistic fantasy are blurred beyond recognition. Dancers spring forth around Delphine seemingly with no specific source, and their radiance and joy rub off on her in spurts. At points, she skips to the rhythms of Legrand's underscoring before returning to her "normal" (i.e., non-dancing) walk. Jonathan Rosenbaum's excellent pieces on Jacques Demy and *Les Demoiselles* introduces crucial terminology that elucidates the elusive effect of these bizarrely choreographed scenes:

The Young Girls of Rochefort often daringly places story and musical numbers on the screen simultaneously, *mixing them* in various ways and in different proportions. One of the stars may be walking down the street, for example, but the pedestrians around her are suddenly dancing, and she slips momentarily in and out of their choreography. This curious mix produces powerful, deeply felt emotions — an exuberance combined with a *sublime* sense of absurdity. [7] (Emphasis mine.)

"Sublime" is certainly the most appropriate term to describe *Les Demoiselles*. However, it is a different form of sublimity than what is traditionally perceived in the movie-musical. Demy's constant shuffling ensures the audience is never on sturdy, familiar ground. His musical-film-world has the look and feel of an American film musical. However, as Rosenbaum observes, it is with a distinctly French, cinephilic twist that does not seek to create a cohesive whole, but rather seeks to combine and re-mix. To put this in the context of the film's choreography, when one sees Delphine walk down the street in *Les Demoiselles*, one observes an irresistible combination of reality and fantasy kick in. The reality is immediately apparent: these folks are simply not adept or particularly refined dancers. They often stumble about in their steps, and the moves are not as synchronized as we would like. And yet, at the same time, fantasy kicks in to tell the viewer, "Though it makes no sense for us to enjoy these deliberately crude steps, we

are in a different logical register than which we are accustomed, and must therefore march to the step of this director's logic." The fantasy enters when one observes the *movement*, regardless of how precise or mannered it is. Demy's dancers move with an intuitive knowledge: dancing feels right in this world. He does not pride the fantasy of a movie world over the reality of his French lower-middle-class upbringing, or vice versa. To him (and perhaps to us, too, after we have finished watching *Les Demoiselles de Rochefort*) they are one and the same. Dancing in *Les Demoiselles* is a larger metonym for life's complex dichotomies: dream-worlds and harsh realities, virtuosity and mediocrity, love and tragedy.

These disparate elements grate against each other to create the film's fascinating tensions, provoking a response that can only be termed—to use Rosenbaum's and the philosopher Jean-François Lyotard's lexicon—as belonging to a "postmodern sublime."

THE POSTMODERN SUBLIME

What does it mean for *Les Demoiselles de Rochefort* to conform to a distinctly postmodern view of the world? And what are the implications of such a connection?

One must first put postmodernism in a larger context of art. For this, we turn to Gary Aysleworth's fantastic summation of Lyotard's initial 1979 conception of postmodernism and "the postmodern sublime":

Lyotard stresses the mode in which faculties (imagination and reason) are in disharmony, i.e. the feeling of the sublime. For Kant, the sublime occurs when our faculties of sensible presentation are overwhelmed by impressions of absolute power and magnitude, and reason is thrown back upon its own power to conceive Ideas (such as the moral law) which surpass the sensible world. For Lyotard, however, the postmodern sublime occurs when we are affected by a multitude of unrepresentables without reference to reason as their unifying origin...In this respect, [the feeling of sublimity] would be more akin to the production of art than a moral judgment in Kant's sense." [8]

Lyotard makes a crucial distinction between the traditionally Kantian conception of the sublime and the postmodern spin on it. In the world of Kant, the sublime fits into a strictly regimented but also naturally formed mold. Chaos is suppressed and the orderliness of the world itself conjures the subliminal release of emotion in the individual. Logic prevails, for the most part. As a result, the highest form of sublimity is reached through the types of moral judgments and the sense of justice that Kant promoted. However, when addressing a "postmodern" sublime, Lyotard refers to a total instability. The world as imagined by the postmodernist is jagged, fractured, without a coherent aesthetic around which everything revolves. As a result, the finished product is purposely impure and imperfect. However, the postmodern sublime still provokes an analogous reaction to the individual as the Kantian sublime; they simply work with different methodologies, one prized on structure, the other on disruption. In addition, Lyotard notes that the postmodern sublime typically springs from a deep engagement with a work of art, not Kant's nature or justice.

Taking this into consideration, we can include *Les Demoiselles* under the larger umbrella term of the postmodern sublime. Demy's world is a motley combination of genres, tones, styles, and modes. Different events that occur in the film are not neatly confined to specific emotional categories. A serial-killer is brought up in the film's second half, but the characters react to the news with unnerving cheerfulness and humor, cracking insane puns about how "what a dirty man he is" but "how cleanly he cuts his girls up." Likewise, as Rosenbaum has observed, the effect of the film's ostensibly happy ending—everybody finds the man/woman they've been looking for—is muted by the much more palpable sense of tragedy revolving around the film's relentless missed connections (for example, potential lovers meant to be together like Delphine and her sailor Maxence, continuously pass by one another, never realizing the other is just within earshot). As a result, such an eclectic mix heightens one's emotional investment in the film. Demy's inability to decide on a set tone excites us. It unnerves the trained musical viewer who has built certain criteria by which they judge a "successful" musical film. Demy's film, however, throws a wrench in these carefully arranged, reasonable plans. It gets by on the pure ephemeral joy its sunny veneer conveys, not necessarily by any virtuoso displays of musical perfection. To be perfect, in the fractured and postmodern world of Rochefort, is an undesirable goal.

The film itself positions itself as a post-modern, post-studio film. After the collapse of the studio system, more and more foreign film imports were being scuttled into the United States. This results in a burst of creative freedom hitherto unseen in America. And Demy's film comes as one of the largest challenges to the perception of the musical as an elitist art that only a few select masters (Minnelli, Donen, Kelly) have mastered. The modernism in Warner Bros. (*Gold Diggers of 1933*), RKO (*Shall We Dance* [1936]), and MGM fare (*An American in Paris*) are forsaken in *Les Demoiselles* for something much more unstable and ruptured: a beast of many colors that ties in all the above studio's aesthetics, combined with a distinctly French sensibility.

We must also consider how dancing enters the equation. Demy presents the viewer an interesting paradox: his choreography looks sloppy, yet we enjoy the film's Frenchified dance-routines. The dancing itself is in accordance with Lyotard's understanding of the sublime. Instead of order, we are presented with disorder. Instead of traditional rehearsal, we get the impression that the dancers, owing to their incessantly smiling faces and cheerful expressions during dancing, are not visibly concerned that their choreography is not refined. Yet one would be hard-pressed to deem these flaws; they are part of the same disrupting structure that distinguishes *Les Demoiselles*.

Taking this all together, we see how the emotional effect of *Les Demoiselles* supplants words. Again, to contextualize it in the realm of the postmodern sublime, the beauty of the experience does not use reason as a guiding force. Reason escapes us. To describe *Les Demoiselles*'s power in words is necessary for a critical reconsideration of this much-misunderstood work. Yet to put that power into words is also antithetical to its true goal: to convey flightiness, joy, tragedy, and sadness through song and dance, but not necessarily words.

CONCLUSION

This analysis of *Les Demoiselles de Rochefort* raises the very important

question of why we should study the film in such a context at all. To dive deeper into *Les Demoiselles* is to seriously engage with the cinema of Jacques Demy to an extent that has not previously been accorded his work. The recent tireless efforts of Demy's widow Agnès Varda to restore her late husband's work have lent his films renewed importance for the next generation of moviegoers and scholars.

Varda's place in film history as one of the French New Wave's finest directors has been solidified, but the same cannot be said of her husband Demy. The ostensibly airy quality of his films—inspired by light entertainment like the American musical, French fairy tales, and pop music—has given the false impression that they do not deserve immediate, serious consideration to unpack their narratives. But as I hope to have illustrated, Demy's cinema goes deeper than its ostensibly sunny surfaces would suggest. Demy's view that dreams and waking life should coexist has immense implications about how the "movie reality" depicted in cinema is not far removed from the reality of our daily lives. His films have much to say about how we live as humans and how we approach an art form as fractured and multifaceted as cinema.

REFERENCES

1. J. Hiller and D. Pye. *100 Film Musicals: BFI Screen Guide*. London: Palgrave Macmillan, 2011.
2. A. Sarris, "The young girls of Rochefort," in *Confessions of a Cultist: On the Cinema: 1955-1969*. New York: Simon and Schuster, 1970, pp. 365-367.
3. J. Rosenbaum. (1998, November 27). *Not the same old song and dance* [Online]. Available: <http://www.jonathanrosenbaum.net/1998/11/not-the-same-old-song-and-dance/>
4. P. Kael, "Trash, art, and the movies," in *The Age of Movies: Selected Writings of Pauline Kael*, S. Schwartz, Ed. New York: Library of America, 2011, pp. 201-234.
5. R. Ebert. (1998, February 15). *Swing time* [Online]. Available: <http://www.rogerebert.com/reviews/great-movie-swing-time-1936>
6. Rosenbaum, "Song and dance."
7. Ibid.
8. G. Aysleworth. (2015). *Postmodernism*, in *The Stanford Encyclopedia of Philosophy* [Online]. Available: <http://plato.stanford.edu/entries/postmodernism/#8>

ACKNOWLEDGEMENTS

I would like to acknowledge Gabrielle Moyer for guiding me to the current vision of my research.



CARLOS VALLADARES

Carlos Valladares is an sophomore double-majoring in Film and American Studies. He has a passion for cinema, music, and filmmaking. He is the Film Desk Editor for the Stanford Daily and soon-to-be Film and TV Editor for the Stanford Arts Review. His current Chappell-Lougee-funded research focuses on a re-appraisal of Richard Lester's films, the director of the Beatles films *A Hard Day's Night* and *Help!* and the Superman sequels. He enjoys heading to the Stanford Theatre every week-end. His favorite filmmakers are Jacques Demy, Jacques Tati, Robert Altman, John Cassavetes, Charlie Chaplin, and Chuck Jones. He hopes to be a professional film critic one day.

Difficult Masses: Femininity and the Female Body in Germaine de Staël's *Corinne, Or Italy (Corinne)* and Mary Wollstonecraft's *The Wrongs of Woman, or Maria (Maria)*

Emma Blauciak

The University of Warwick, England

“- let me but prepare her body and mind to encounter the ills which await her sex.” (*Maria*)

Whilst we typically think of femininity as the defining essence of womanhood, there is no *essential* definition of femininity; only a collection of aesthetic, social, political, sexual and hierarchical notions which concretise into a linguistic descriptive intended to essentialise qualities of ‘woman’. In this vein, femininity can be understood as a social conception that depicts how the female bodily mass should be managed. During the nineteenth century, a period where gendered identities were rigidly defined, the management of the female body was particularly contentious. The evolving notions of femininity and ‘proper’ female behaviour are legible as means of control; there is a pedagogical function of the ‘feminine’ (concept) for the female body (form). Different forms of ‘femininity’ are woven from conceptions of bodily sexual identity and are inextricable from ideas about how the female body should ideally be perceived, and how it should conduct itself, both in the public and private sphere. Feminist cultural theorist, Luce Irigaray, suggests that female social roles are indivisible from male function or need: woman is either ‘mother, virgin [or] prostitute’ [1]. In this way, woman’s body – *feme covert* or *feme sole* - is expropriated by society to be modelled ‘feminine’ in accordance with a specified social order; an order that stipulates a subsidiary role for women. In the social conception of gendered identities, the body and femininity are mutually dependent and reinforcing.

For this order to function, ‘feminine’ must contend with denied and suppressed forms of femininity which are bound to the sexual identity of the body’. Lyn Pykett’s theorisation of ‘proper’ and ‘improper’ feminine is indicative of a social tendency to dichotomise the limits of appropriate femininity and sexual expression. She suggests that the ‘proper’ feminine is a socially desirable model of domesticated femininity characterised by innocence, dependency, passivity and passionlessness’ [2]. The suppressed and subversive shadow of the ‘proper feminine’ is comparatively ‘figured as a demon or wild animal’, ‘a whore’, ‘independent’ and ‘a subversive threat to the family’ [2]. In this way, a woman’s condition on the sexual spectrum is perceived to be indicative of the social function or role she can perform. In Mary Wollstonecraft’s *The Wrongs of Woman, or Maria (Maria)* (1798) and Germaine de Staël’s *Corinne, Or Italy (Corinne)* (1807), the narratives are structured through oppositional representations

of the feminine, expressed through comparisons of the aesthetic, maternal and sexual [3][4]. The reduction of woman to her bodily sexual function across the class scale, through the social classification of ‘proper’ and ‘improper’ feminine is rendered problematic in both texts. Each challenges the constructions of femininity which compound woman’s social condition as politically and emotionally submissive, whilst simultaneously expressing the inescapability of woman’s very *bodily* condition. Written by female authors at the cusp of the nineteenth century, and before the women’s movement mobilized, both fictional works engage with the complex and contradictory ideas that would challenge social thought throughout the century. This essay will examine differential models of femininity and analyse their relationship to social situation in *Corinne* and *Maria*, by exploring how these are manifested and revered or otherwise reviled in the representation of the female body.

In *Corinne*, the inextricability of the female body (as the physical aesthetic form) from social conceptions of femininity is demonstrated through the juxtaposition of the dark haired ‘transgressive’ Corinne, and the appropriately domesticated, pale and innocent Lucile. Pykett identifies detailed ‘writing [of] the [female] body’ [2] as a signal for the deviation from the ‘proper feminine’. The textual germination of the ‘improper’ feminine can thus be explored through opposing degrees of embodiment as the bodily substantiation of Corinne, synonymous with her independence, is juxtaposed against the nebulous metaphorical rendering of Lucile. It is important to note that the contention between oppositional femininities, and how the female body is represented and evaluated in social and ideological terms, is frequently viewed and elucidated from masculine perspectives. Throughout the text, the mental conflict of the ‘distinguished’ Englishman ‘Oswald’, as he deliberates the comparative prospects of a future relationship with the two women, represents a broader social contention about femininity.

Oswald’s first sighting of Corinne coincides with a public acclamation of her talents and literary work; she enters an open public space in the midst of ‘universal enthusiasm’ [3], and it is emphasised that she is of the ‘exceptional’. Oswald’s presence and the narratorial reflections on English custom act as a commentary to the ‘improper’ nature of such a public display of independent

femininity. In this instance, the occasion is intended to draw attention to Corinne's intellectual merits, but textual descriptions repeatedly reduce Corinne to her bodily condition: 'Her arms were dazzlingly beautiful; her tall, slightly plump figure, in the style of a Greek statue, gave a keen impression of youth and happiness [...] all her movements had a charm which aroused interest and curiosity, wonder and affection.' [3]. The exposure of her arms and the fleshy nature of her body increasingly sexualise Corinne and express a level of relief from the strict reserve of English dress and feminine bodily frailty.

Comparatively, when Oswald first perceives Lucile, she is engaged in a private act, on the land inherited by her father, and is partially obscured from view. A bold bodily description is absent and Lucile is instead described through epithets of femininity which express female deficiency, modesty and ethereal beauty:

Lord Nevil was struck by her imposing, modest look and truly angelic face [...] Her features were remarkably delicate, her figure almost too slender, for a little weakness could be seen in her walk [...] paleness gave way to blushes in a moment.' [3]

In this way, Corinne and her half-sister Lucile are aestheticised into female dichotomies. The perception of Lucile's bodily form and the filial and nurturing roles for which this renders her suitable, are manifested in conventional ideas of female weakness as a model of beauty; she escapes complete embodiment by inhabiting hazy social tropes of feminine desirability. Her delicacy and modesty are indicative of the desired sexual chastity of woman and her innocence (Wollstonecraft terms this but a civil word for weakness) insulates her from knowledge and reason, thus perpetually enforcing her need for protection [4]; this highlights the mutual relationship between the women's bodies and notions of their 'proper' relationship to public and private spheres. Lucile, despite her descriptive 'disembodiment' [2], is reduced to bodily reproductive functions in accordance with the domestic role dictated by her femininity. Whilst Corinne, defined by her intellectual productivity, is consecrated in a very public setting, sexualised in accordance with the 'inappropriate' nature of her position, and disregarded in bodily reproductive terms. Both instances illustrate how the female body acts as a signifier for masculine use and social function.

Seen through Oswald's eyes, Corinne's embodiment is a collation of disparate ideals and descriptors are tempered so that she appears to embody oppositional virtues. She possesses both conventionally masculine and conventionally feminine characteristics: 'sound judgement and rapturous emotion, strength and gentleness' [3]. In aesthetic terms, Corinne's attire, the 'Indian turban [...] intertwined with her beautiful black hair' is evocative of the dark and exotic; she is descriptively exposed, sexualised and 'Othered'. However, her dress, 'white with a blue stole fastened beneath her breast', whilst drawing attention to the body beneath the clothing, is also evocative of the conventional aesthetic representation of the Virgin Mary in hieratic art. It is as though, because of her unconventionality, Corinne cannot be satisfactorily placed on the spectrum of female sexual identity; she lives through a 'double existence' [3] which unifies oppositional models of the feminine and masculine but cannot yet be incorporated into social order.

The relationship between femininity and social order is brought into relief through the comparative representations of the female protagonists in *Maria*. Whereas Maria, an upper class



Figure 1. A 19th-century painting of Germaine de Staël.

'lady', is composed of 'supernatural benignity of countenance', 'feminine softness,' and 'infantine ingenueness,' Jemima, a lower class attendant, is physically unfeminine with 'rugged cheeks' and looks 'like a little old woman, or a hag shrivelling into nothing' [4]. In 'Reason and Sensibility in "The Wrongs of Woman,"' Janet Todd argues that Jemima emerged into the world as a 'deracinated adult, born into no family' whereas Maria, through the cultivation of sensibility and romantic sentiments, 'has never truly become an adult.' [7]. The juxtaposition of Maria's hyper-femininity and Jemima's abject bodily condition disputes the idea of woman's 'natural' bodily tendency to be feminine and instead suggests that femininity is dependent upon class and social situation⁸.

The loss (or lack) of bodily purity extinguishes the female potential to be considered 'feminine'. Jemima enters the world without a patronymic as a result of her mother's sexual ruin and father's abandonment - 'born a strumpet; it ran in my blood' - and is thus 'outcast from society' [4]. Her class position enables assumptions about her 'femininity' as modesty and sexual propriety are not inculcated, nor do they have a social function, in the lower classes. As an apprentice, she is beaten and repeatedly forced into sex by her master: '[he] compelled me to submit to his ferocious desire and [...] I was obliged [...] to comply' [4], to exchange her body for bed and board. The contingency of patriarchal protection and female sexual status on all other aspects of social existence limits the female possibility for independent social standing: despite having no institutionalised male signifier for support or definition, Jemima is perpetually dependent on the male. Her femininity is nullified by the absence of a patronymic

and so, with the status of an abject body, she is reduced to sexual function, and forced to donate her body as ‘common property’ [4].

The male’s multiple use of the prostitute’s body negates her ‘feminine’ purity and she is removed from social ‘symbolic order’ except as a ‘diseased’^v or denied body: ‘I was **despised** from my birth and **denied** the chance of obtaining a footing for myself in society’ [4]. Jemima cannot, either, be described as masculine as her ‘sexualisation’ and working status do not bring the implications of power, control and autonomy that they would in male terms. Excluded from the patriarchal protection of the domestic space and shunned to “the street, utterly destitute!”, Jemima is a passive ‘victim of social and economic circumstance,’ rather than an active participant in the public sphere [2]. She is a body despised; valuable, Irigaray asserts, because her body has already been used, her ‘natural assets’ expended, but is excluded from participation within the gendered social order [1]. Metaphorically rendered the ‘filching cat, the ravenous dog, the dumb brute’ [4], Jemima is neither masculine nor feminine, but is condemned to the status of ‘object’ or animal: a body with a social sexual function but without the ability to embody either gendered implication.

The public nature of Corinne’s body and her prohibition (as the ‘improper’ feminine) from refuge in the domestic ‘haven’ parallels Jemima’s social condition. Admiration of Corinne is inseparable from the novelty of her gender and her condition as an object of sexual desire. By aligning Corinne with the artistic, performative and pictorial, her aesthetic objectification gains merit as a public right. In her dance of the *Tarantella*, Oswald becomes scathingly envious of Corinne’s position in front of ‘assembled company’. He feels ambiguously about her superiority; wishing her to be both shy and eloquent, to possess talent but to privatise it for him only. Angered because Corinne seeks self-fulfilment in freedom from him rather than grovelling to please him in melancholy submission: ‘it was almost to escape from his influence that Corinne was making herself so captivating,’ Oswald sighed, ‘as if every one of her successes had separated him from her’ [3]. This sequence provides a clear example of Corinne’s ‘self-construction as a visible public object’ but the description of her body and implications about her femininity are problematic [8]. The internal focalisation of Oswald, and Corinne’s position as the object of public sight, objectifies her from a very masculine perspective. Again, throughout this scene, Corinne’s genius and accomplishments are reduced to bodily referents:

‘she began to dance, and in all her movements there was a graceful litheness, a modesty mingled with sensual delight, giving some idea of the power exercised by the temple dancing girls over the Indian imagination. They are, as it were, poets in their dancing’ [3]

The ‘temple dancing girls’ refers to the Indian Devadasi system whereby a young girl is dedicated or ‘married’ to a deity and conditioned to a life of performativity and enforced prostitution [9], [10]. It is a system which effectively puts virginity on the market [1]^{vi}. Aestheticizing Corinne through mythological, pictorial and historical referents, she is ‘dressed like Domenichino’s Sibly’, has

a body like a ‘Greek Statue’, and charm like ‘Cleopatra’ [3] has a distancing effect which aligns her with anomalies of history and exotic qualities which enforce her condition as a feminine ‘Other’. In this way, along with the construction of her as ‘improper’, the exotic referents embedded in Corinne’s description serve to eroticise, and align her with sexually deviant femininity.

This places Corinne as a sexual object of sight and feeds into Nancy Miller’s argument that Oswald’s ‘inability to see except *as a man*’ is central to Staël’s critique of patriarchy [8]. As a man, he views Corinne as a body inseparable from his needs; his sexual desire; and his compulsion to enforce supremacy through female objectification. Thus, Corinne, aligned with temple dancing girls, is limited to the expression of female creativity in bodily terms as a ‘poet in [her] dancing’ [3]. The struggle between biological, bodily creativity through reproduction and an alternative self-sustaining intellectual form of creativity is a *topos* of feminist literature and here, expressed through Oswald’s objectification, reveals his desire to control, and assert male sovereignty over, the independent intellectual feminine. Oswald refutes Corinne’s ability to stand as a body alone and enforces the feminine need for protection by asking: ‘amongst all these admirers and enthusiasts [...] Is there a lifelong protector?’ [3]. She is positioned as a body

that is public property to be analysed, judged and admired but is perpetually excluded from conventional domesticity. In this way, Corinne and Jemima appear to reside outside of the social ‘symbolic order’ as they cannot be incorporated into the workings of the patriarchal family system^{vii}.

However, somewhat contradictorily, Corinne’s success relies not only on her ability to inhabit conventionally masculine positions, but also to diminish threats to male supremacy by providing very public displays of femininity. Her objectification is seldom separate from a social evaluation of her femininity; her renown is inseparable from feminine embodiment. During the crowning scene, Corinne goes down on one knee according to the custom of the ceremony, and it is as she is in this position of deference that Oswald feels his eyes fill with tears:

‘in the midst of all this splendour and success it seemed to him that Corinne’s eyes had sought the protection of a man friend, a protection no woman, however superior she may be, can ever dispense with.’ [3]

Here Oswald constructs female dependency as innate, rather than a social condition. The perceived vulnerability of woman is regardless of intelligence, independence, and self-sufficiency. The conventional elements of Corinne’s feminisation enable Oswald to admire the aesthetic novelty of her ‘extraordinary situation’ [3] whilst maintaining a perception of her weakness and need for protection. This demonstrates Oswald’s very early desire to capture and restrain Corinne’s open display of unconventional femininity. Miller asserts that Oswald privatizes the public meanings of woman’s gaze ‘in order to consolidate his own identity’ [8]. He limits the possibility of sustained female independence by conditioning Corinne with the need to be rescued by a masculine figure. Corinne’s genius and success are implicated in the conditions of her body; her femininity enables

Corinne’s genius and success are implicated in the conditions of her body; her femininity enables the admiration she receives but also implicitly undermines it.

the admiration she receives but also implicitly undermines it.

In Oswald's overheard and expressed descriptions of Corinne, it is the performative element of her genius that is emphasised ('no one **performed** tragedy like her [...] she **danced** like a nymph' [3]), and this emphasis underlines the perceived sexual expressivity of her body. Comparably, with Lucile, it is the eloquence of modesty and silence that speaks volumes: "[Oswald] wondered how it was possible for the simplest movements and the most ordinary words to reveal a whole soul." [3]. Lucile's 'charming sweetness', domestic activity, and emotional innocence position her as maternal and submissive and prevent her from challenging Oswald's pretensions to supremacy. Oswald's analysis of the effects of 'mysterious veils of silence and modesty' reveals the function of the 'proper' model of femininity; a blank canvas onto which man can project 'whatever virtues and feeling he desires' [3]. Similarly, Lucile's austere life of seclusion invigorates Oswald's attraction to her as it signifies her emotional as well as social and sexual purity; her 'matchless beauty' has been deprived of 'all the homage of society' [3] it deserves. This enables Oswald to envision himself as her sole protector, as only he has possessed her as a sexual object of sight.

Corinne's revelation that Lord Nevil (Oswald's father) had visited her father for a 'viewing' similarly constructs her through the lens of social utility-value as a potential, but ultimately rejected, marital commodity for Oswald. Lord Nevil, as a signifier for the patriarchal institution, determines: 'It would be better to give him Lucile' [3]. Corinne's 'exceptional talents', 'active imagination' and 'foreign ways' diminish her social value as they are considered potential threats to the 'calm hopes of [a] domestic life' which serves to enforce masculine supremacy and maintains woman's position in the 'shade' [3]. There is a suggestion that Corinne's model of femininity - as a balance of oppositional virtues - would provide Oswald with a more fulfilling partnership, but fears for her reputation and the domestication of such femininity, refute this possibility. On the contrary, the youth and 'modesty' perceived in Lucile's features already construct her as a secondary, shaded figure in keeping with the 'English' feminine model of domestic virtue.

The construction of vulnerability and weakness in the feminine social condition enables the effective commodification - in the way of protection, ownership, and possession - of the female body. In relation to the social sexual order, prostitutes enable male sexual satisfaction whilst preserving other chaste innocents for marriage. However, Pykett asserts that '[t]he dominant discourses on marriage and respectable femininity [in the period] were [...] destabilised by a tendency to see contemporary marriage as [...] a form of prostitution' [2]. Maria's profligate husband's perception of the female body is extreme, as he appears not to dichotomise 'proper' and 'improper' femininity in terms of sexual use-function^{viii}, but debases all females equally by considering them solely in bodily terms. His favourites are 'profligate' 'wantons of the lowest class' and he despises 'modest' women possessed of 'female endowments' [4]. Beyond the constructions which aid his abuse of women, feminine 'romantic sentiments' are little but a moralising threat to his position of authority [4]. In his terms, female bodies are disposable objects to seduce, sexually utilise and abandon.

Maria and Jemima's reduction to the sexual function of their bodies through prostitution and rape (legalised or through social necessity) is expressive of the cross-class social function of the



Figure 2. A 1797 painting of Mary Wollstonecraft by John Opie.

female body. In marriage, women maintain their social reputation by espousing the requirements of femininity (compassion, patience, silence) and debasing their bodies, as Maria does in her marriage to George Venables. Bound to an adulterous drunkard who squanders all her economic wealth and social worth, she is repeatedly forced, through feelings of social necessity and emotional weakness, to submit her body against her will: 'compassion, and the fear of insulting his supposed feelings [...] made me dissemble, and do violence to my delicacy. What a task!' [4]. The principles of her femininity are sacrificed to the necessity to submit her body. This system of 'legalised prostitution' is enabled by social constructions of 'proper feminine asexuality' [2] and 'a woman's coldness of constitution, and want of passion,' [4], leaving women with no other recourse but to be manipulated into submitting their bodies as commodities for exchange.

Beyond Maria's marital bodily submission, her husband attempts to reduce her body to 'exchange value' [1] by prostituting her in an economic transaction; he assures her seducer 'that every woman had her price, [...] and concluded with requesting him to lend him five hundred pounds.' [4]. Despite being his wife and 'the mother of [his] child', George reduces Maria to absolute sexual function and illustrates the masculine perception of the female body as a form economic property to be pawned. George similarly reveals the social utility of Maria's femininity as he recommends it as an implement for manipulation, advising 'not to shock [her] romantic notions, but to attack [her] credulous generosity, and weak pity' [4]. He is aware that her 'femininity' inspires compassion and leads Maria to be repeatedly 'duped' by his dissimulation.

However, Maria's rejection of prostitution coincides with a resolution to take control over her body and action: "I will

provide for myself and child”. Whilst her compassion and sense of duty caused her to be misled, her feminine principles (modesty and self-worth) imbue her with a sense conviction to escape the tyranny of her husband. To thus degrade herself and the marriage institution by adopting ‘liberal sentiments’ would not be in keeping with her notions of femininity: ‘How had I panted for liberty – liberty, that I would have purchased at any price, but that of my own esteem!’ [4]. Her class position, and the prospect of economic support from her uncle, similarly provide her with the means to reject absolute bodily subjection. Jemima, of the lower class strata, would not have access to the luxury of such imagined independence. In this way an individual’s social status, class and upbringing enable or otherwise exclude femininity.

Despite this, Maria is aware that ‘[a] woman [of any class status][...] resigning what is termed her natural protector [...] is despised and shunned’ [4]. On her escape from ‘her tyrant’, Maria’s working class landlady declares: ‘Women must be submissive, [...] what could most women do? Who had they to maintain them, but their husbands?’ [4]. Patriarchal structures that prohibit female independence similarly construct feminine submission as a virtue. Adopting this model is the only means to attain stability, maintenance and masculine protection. For this reason, the default feminine qualities of silence and submission are adopted by many women as a social necessity to maintain legitimacy, and to avoid becoming abject bodies.

In Maria’s retrospective account, she begins by performing a socially acceptable model of ‘proper’ femininity but transgresses as she rejects her husband’s tyranny and attempts to gain self-assertion. Through this, she relinquishes her status as a ‘lady’ and becomes a denied body in society, having to skulk in ‘shame’ from different abodes. On attempting to visit ‘ladies with whom she had formerly been intimate’, she is ‘refused admittance’ or otherwise forgotten by their society [4], this highlights the social hypocrisy which values the appearance of feminine propriety, over and above its actuality. The perceived connection between femininity and female corporeality is similarly expressed in *Corinne*. Corinne’s regression from the height of glory to a state of social and bodily impotence – where she craves being Oswald’s ‘slave’ – catalogues her decline into abjection [3]. Previously admired for her vivacity as a very public and feminine object, Corinne, rejected by the British patriarchal institution¹⁸, becomes a distressed, pale and trembling body living in seclusion. Her decline is measured by her aging and debilitated bodily frame compared to Lucile’s ‘delightful youth’ [3] and beauty. Her grief renders her intellectually ‘infertile’ and her previous self-sustenance is diminished by the inextricability of her self-worth from Oswald’s perception of her. Hence, her social performance of femininity is doused alongside the independence and intellectual abilities which caused her to be branded the ‘improper’ feminine.

Relinquishing femininity inescapably coerces the female body into a state of abjection. Regardless of her character or condition on the sexual spectrum, woman is reduced to commodity function and her body is utilised according to the value she presents to masculine need. The eloquence of the female body as a signifier for male social requirements means that, without possessing attributes of ‘proper’ – privatised and submissive – femininity, she is invalidated as a body within the social patriarchal system. Thus excluded from the ‘symbolic order’, woman is reduced to a condition of absolute bodily abjection; sexuality debased, physically debilitated and socially condemned.

REFERENCES

1. Irigaray, Luce. *This Sex Which Is Not One*. Trans. Catherine Porter and Carolyn Burke. Ithaca: New York: Cornell University Press, 1985. Print. pp.186.
2. Pykett, Lyn. *The ‘Improper’ Feminine: The Women’s Sensation Novel and the New Woman Writing*. London: New York: Routledge, 1992. Print. pp. 12.
3. Stael, Madame de. *Corinne, Or Italy*. Trans. Sylvia Raphael. Oxford: Oxford University Press, 1998. Print.
4. Wollstonecraft, Mary. *A Vindication of the Rights of Woman and The Wrongs of Woman, or Maria*. ed. Anne K. Mellor and Noelle Chao. New York: Pearson: Longman, 2007. Print.
5. “abject, adj. and n.” *OED Online*. Oxford University Press, December 2014. Web. 16 December 2015.
6. Gilbert, Pamela K. *Disease, Desire, and the Body in Victorian Women’s Popular Novels*. Cambridge: Cambridge University Press, 1997. Print.
7. Todd, Janet. *Sensibility: An Introduction*. London: New York: Methuen & Co. Ltd., 1986. Print.
8. Miller, Nancy, K. ‘Performances of the Gaze: Stael’s *Corinne, or Italy*’. *The Novel’s Seductions: Stael’s Corinne in Critical Inquiry*. ed. Karyna Szmurlo. London: Associated University Presses, Inc., 1999. Print.
9. Chattoraj, Samantha. “The Devadasi System”. 2002. Accessed 19th December 2014. <http://iml.jou.ufl.edu/projects/Spring02/Chattoraj/genesis.html>
10. Colundalur, Nash. “Devadasis are a cursed community”. *The Guardian*. Friday 21 January 2011. Web. Accessed 19th December 2014.

ENDNOTES

- i. As a result of the cult of sensibility in the period, the discourse surrounding femininity in the late 18th and early 19th century was particularly sensitive to the female body. The very corporeal aspects of sensibility, what Christopher Nagle terms ‘affective excesses’ – fragility or weakness of constitution, sensitivity, delicacy, moral virtue – are largely interchangeable with epithets of femininity.
- ii. The effects of the sexual conception of women’s bodily self is manifested in the historically poor reception of Wollstonecraft’s *Maria* as it was released in the shadow of her ‘racy’ memoir. People feared association with her theories about femininity and female social role because of the perceived tainting of her own bodily sexual condition. As a result of this, she was interpreted as a radical and tainted figure and her work was treated with the same aversion.
- iii. Harriot Taylor’s Westminster Review article of 1851 discusses men as possessing ‘substantive existence’ (57) and women acting as mere ‘humble companions’ (Taylor 1851:300). Pykett cites Taylor’s article which, she suggests, unmasked the tyranny conventionally masquerading as protection by demonstrating whose interests were best served by positioning women as delicate, dependent creatures.’ *The ‘Improper’ Feminine*. pp. 57.
- iv. The dictionary definition of ‘abject’: ‘Cast out, brought low; downtrodden, desperate, inferior, rejected, degraded, base’ is meant for the understanding of ‘abject’ body in this essay. Although not directly used here, Julia Kristeva’s theorisation of the abject as that residing outside of the symbolic social order, and between concepts of the object and subject, is interesting for thinking about Jemima’s condition as a rejected creature-like form; not quite human nor animal. It would be interesting to explore how Kristeva’s ideas engage with Pamela Gilbert’s work on the ‘diseased’ body and Pykett’s ‘Improper’ feminine. [5], [6].
- v. See Gilbert, Pamela. K. *Disease, Desire and the Body in Victorian Women’s Popular Novels*. pp.17.: ‘In the Victorian era two kinds of bodies definable as grotesque were the diseased body and the body of the prostitute. Both were defined chiefly by their permeability, and both became objects of the gaze.’ [6]
- vi. See Irigaray’s ‘Women on the Market’ in *This Sex Which is Not One* for a further analysis of the social exchange value of female virginity.
- vii. ‘Symbolic order’ is a term used by Irigaray in *This Sex Which Is Not One*. She suggests that the ‘trans-formation of women’s bodies into use values and exchange values [is what] inaugurates the symbolic order’ and women thus maintain the social order ‘with their bodies’ (190). Here I suggest that exclusion from the patriarchal family system removes the potential to be considered ‘feminine’ and woman’s body becomes ‘improper’ – denied and suppressed.
- viii. Derived from Irigaray’s theorisation of the ‘use-value’ of woman’s body in

“Women on the Market” in *This Sex Which Is Not One*.

ix. As signified by Oswald and his father’s rejection of her.

IMAGE SOURCES

FIGURE 1: Duyckinck, Evert A. “Portrait of Madame de Staël.” 1873. Originally painted by François Gérard circa 1810. Online image. Wikimedia Commons.

FIGURE 2: Kaldari. “Mary Wollstonecraft by John Opie (c. 1797).” 5 July 2008. Online image. Wikimedia Commons.

ACKNOWLEDGMENTS

I would like to express my gratitude to the fabulous - always happy and down-to-earth - Dr Laura Wood for her guidance and support. I am forever grateful to my Mum for absorbing my self-doubt and continually spurring me on.



EMMA BLAUCIAK

Emma is a recent graduate (2015) of the University of Warwick, England, with a 1.1 BA (Hons) in English Literature. She has a particular interest in gender and social relations in literature and film and has ambitions to work in film and television production. She produced her first short film early this year which won a local Histories Award and is currently working to expand her knowledge and experience of the industry. Emma is a lover of vintage clothing and second hand objects and spends much of her free time rummaging around markets at home and abroad; her next stop is Asia!

SPEAKING THROUGH LIGHT

HANNAH ROSEN & THE HUMBOLDT SQUID

by Hanna Payne

The Humboldt squid communicates using different patterns of flashing light. Its whole body illuminates as it carries on a conversation with its neighbor. Hannah Rosen, a graduate student working in the Gilly Lab at Hopkins Marine Station (HMS), thinks this is pretty darn cool. She spends her days fitting these squids with specialized swimsuits and video cameras to record their every move.

Growing up in the rural suburbs of the northeast, Rosen's childhood was centered on the adventures that waited just outside her front door. She spent years exploring the forests that lined her street, searching under rocks and over stumps for animals to entertain her youthful curiosity. From an early age, Rosen knew she wanted to work with animals, and as she grew, she became more interested in how different organisms think in comparison to each other.

Rosen's passion continued at Bucknell University, where she majored in Animal Behavior. Rosen was initially drawn to the octopus because of its unique cognitive abilities, but she soon came to appreciate the often overlooked squid.

Squid, as opposed to octopus or cuttlefish, lacked research because of the level of difficulty in studying them in the wild. Hannah began reading papers about the creature and wrote to researchers whose work aligned with her own curiosities. She decided she wanted to study Cephalopod behavior further.

Graduate school took Rosen 2,700 miles from the familiar backyard hideaways of Northeast Pennsylvania to sand beaches that now sneak up to her lab door at the Hopkins Marine Station (HMS) in Monterey Bay, California. In Rosen's search for a program that would allow her to work on Cephalopod research, one name featured prominently: William Gilly. She joined Gilly's lab as a PhD candidate at Stanford, and began working in conjunction with Gilly on the connections between neuroscience, behavior, and physiology.

Rosen says that she originally was resistant to the idea of studying the physiology behind squid behavior: "When I first

contacted him, I was just interested in the behavioral aspect of squid and their forms of communication. Gilly mostly does neuroscience, so he pushed me to look at behavior and physiology behind that [communication]".

As she started working at the lab, Rosen began to realize how the physiology of squid and their behavior complemented each other, and how these connections could further her exploration of their communication methods. Rosen and Gilly began to understand the impact of the squid's environment on its behavior, allowing their research to become more well-rounded.

To study squid communication in their natural environment, the Gilly Lab collaborated with National Geographic to capture videos of Humboldt Squid interacting with each other at great depths. A year before Rosen joined the lab, Gilly Lab researchers went to Mexico and fitted squid with a swimsuit-engineered camera to observe their behavior. National Geographic helped develop minimally invasive recording methods that allowed the researchers to reduce bias introduced by human-squid interaction. Now, Rosen and the lab are hoping to recreate the data from the Mexico trip during an upcoming trip to Peru. During this expedition, the team will try new techniques, such as infrared lighting, in an attempt to minimize outside factors.

Rosen is particularly excited for this trip, after having worked with the data collected earlier. "I can't even tell you how many times I've watched this video, over and over again, looking into the world of these animals. It's thrilling," Rosen said, referring to the data video from the Mexico trip. Taken in black and white, the video could easily be confused with an action-packed alien thriller; it's almost hard to believe we're looking into the deep blue below us, and not the never-ending black above us.

Rosen published her study entitled "Chromogenic behaviors of the Humboldt Squid" in the *Journal of Experimental Biology* in January of 2015. "I feel like a legitimate scientist now; I'm making my way, and I'm really doing this," Rosen remarked.

With the Pacific Ocean in her backyard, Rosen's often describes her daily experiences at the Gilly Lab in Monterey as enlightening. Her graduate research has allowed her to travel from the blue sea to the blue screen, and with promising video footage from past trips, she looks forward to more adventures studying squid speech.

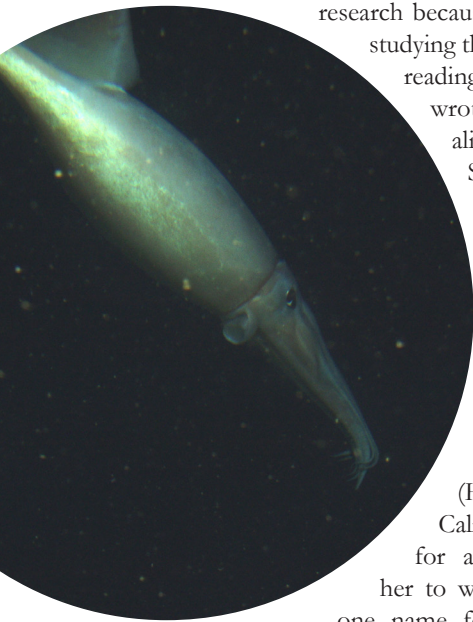


PHOTO: NOAA/CORBIS

natural sciences

Are E-Cigarettes a Safe Alternative to Smoking?

Hayley Tartell

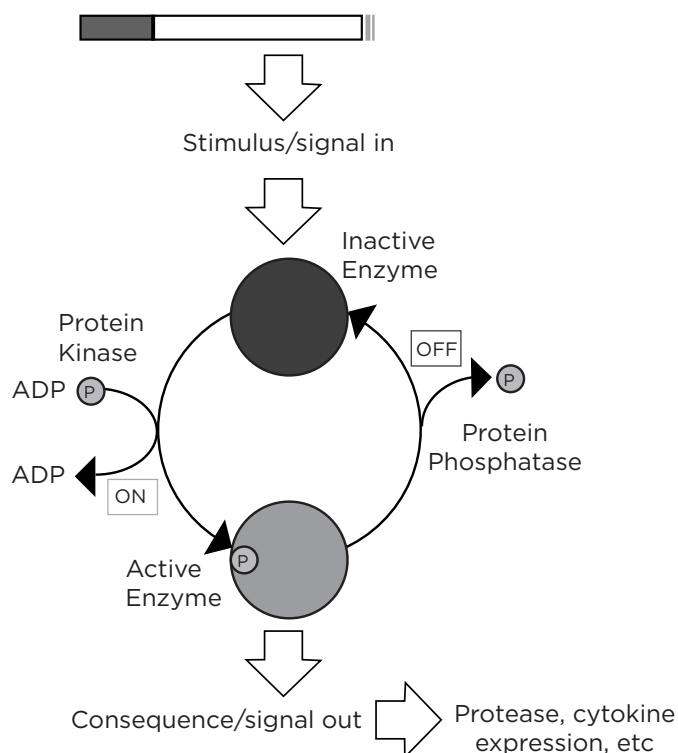
Williams College

The health effects of E-cigarettes, widely used alternatives to tobacco cigarettes, are not well known. The aim of this project was to assess the potential health effects of the products of E-cigarettes in lung cells compared to those of cigarette smoke, and investigate how E-cigarettes affect the progression of chronic obstructive pulmonary disease (COPD). Cigarette smoke is a known activator of proinflammatory MAP (Mitogen-Activated Protein) kinases such as ERK, p38 and JNK, leading to inflammation and potential tissue degradation [1]. The effects of E-cigarettes on these kinases are not well documented. We exposed A549 epithelial lung cells to various concentrations of nicotine alone or American eLiquid™ with or without nicotine. Nicotine was used as a variable, as one of the aims was to determine how the composition of the E-liquid affected these lung cells when exposed. Cell toxicity assays were performed to assess the effect of the liquid found in E-cigarettes (E-liquid) on airway cells and examine E-liquid's potential proinflammatory kinase activation capability to assess E-liquid's role in the progression of COPD. E-liquid appeared to be toxic to cells at concentrations as low as 2% of total media levels and induced activation of every MAP kinase examined. Additionally, similar to cigarette smoke, E-liquid appeared to activate the proinflammatory transcription factor nuclear factor kappaB (NF-kappaB), which is involved in regulating a multitude of cellular pathways involved in the onset of disease [2]. E-liquid exposure from E-cigarette usage could potentially be harmful to the lung and may not be a safe alternative to cigarette smoke.

Introduction

In an effort to find a safer replacement for smoking, many people have turned to E-cigarettes, which are cylindrical devices that have a strikingly similar appearance to conventional cigarettes. E-cigarettes are comprised of batteries and replaceable cartridges with a propylene glycol or glycerin liquid solution containing differing quantities of nicotine. The liquid solution vaporizes with the heat of the battery, and enters the lung upon inhalation. In contrast to traditional cigarettes, this apparatus in theory releases far fewer toxicants by eliminating the smoke component of cigarettes [5, 4]. However, despite their widespread use, the health effects of E-cigarettes and their exposure products have yet to be fully investigated, especially with regard to chronic obstructive pulmonary disease (COPD) progression.

COPD is a progressive disease that induces airflow obstruction and makes it increasingly difficult to breathe [5]. Over time, there is a loss of elasticity in the lung due to oxidative stress, apoptosis and protease expression [5, 6, 7, 8]. COPD is currently the third leading cause of death in the United States [5] and is estimated to develop into the third leading cause of death globally within the next twenty years [8]. Over the past thirty years, the age-adjusted mortality rate correlated with COPD has grown at an alarming rate of 71% despite readily available pharmacological treatments [5]. Cigarette smoke is a primary inducer of COPD, as most people afflicted with the disease are either current or former smokers [5]. According to epidemiologist Prabhat Jha, tobacco is



Figures 1. Cellular mechanism of potential nicotine or cigarette smoke-mediated activation and phosphorylation of MAP kinases.

projected to kill around one billion people in the 21st century (Jha, 2011). [9]. Accordingly, more effective treatment strategies are needed to directly address the fundamental disease pathogenesis and the issue of smoking.

To better understand the issue of smoking, this study investigated the health effects of E-cigarettes by utilizing airway cell culture techniques. This study attempted to test the toxicity of E- cigarettes using epithelial A549 lung cells, a residing lung cell type that would be exposed to both cigarette smoke and E-cigarette vapor *in-vitro*. A549 cells were chosen because these cells are good models for cells that would normally be exposed in the human body. Additionally, this project attempted to determine whether or not the nicotine was the primary factor in E-cigarettes that could activate specific targets and transcription factors associated with inflammatory signaling responses within A549 (lung epithelial) cells and COPD. Cigarette smoke induces phosphorylation, and, thus, activation, of key proteins associated with inflammation (Figure 1). Thus, the following hypothesis was formulated for this experiment: if cells are treated with E-liquid, then MAP kinase activation will induce NF-kappaB activation and subsequent production of cytokines and protease genes.

METHODS

Cell Culture

Human A549 cells (ATCC® CCL-185™, ATCC, Manassas, VA) were grown to 70% confluency on Dulbecco's modified Eagle Medium (gibco® DMEM(1X) + GlutaMAX™-1) treated with Fetal Bovine Serum and Penicillin/Strep. Media was replaced every 2 days during the growth period. After 70% confluence was achieved, media was drained, cells were rinsed with PBS, and trypsin (enzyme) was added. The resulting solution of cells and trypsin was transferred to a test tube and centrifuged for 5 minutes at 1100 rpm at room temperature. Cells were then re-suspended into fresh media and 2 ml of the resulting solution was pipetted into each well of a 96-well plate. Cells were allowed to adhere to the flask overnight. Once cells had reached 60-80% confluence, they were used for experiments with E-liquid. We exposed A549 epithelial lung cells to concentrations of nicotine alone (0.01 mM) or E-liquid (American eLiquid™; 50:50 v/v propylene glycol to vegetable glycerin) with or without nicotine (100% dose has 36

mg/ml (or 222 mM) of nicotine). E-liquid was diluted in culture media to the concentrations outlined in the figures.

LDH Assay

To determine toxic doses of E-liquid, A549 cells were treated with 2 mL of the following concentrations of E-liquid minus nicotine: 0, 0.22, 2.2, 11, 22 & 44 mM (0, 0.1, 1, 5, 10 & 20%, respectively). PBS was used as a negative control. Cell death was colorimetrically assessed by quantifying Lactic Dehydrogenase (LDH) release via the LDH Assay. Media was collected from cells 24 hours after exposure, and LDH release was determined in this media. Using a 96-well plate, 50 µL of media from each condition was examined in triplicate (Figure 2). A 5 ml mixture containing 1666.7ul of each LDH assay substrate solution, LDH assay cofactor, and LDH assay dye solution was prepared, of which 50 µL (standard kit volume) was pipetted into each well containing samples (Figure 2). The plate was then incubated in the dark for 30 minutes, and substrate absorbance was measured at 410nm using a Tecan Genios spectrophotometer. The rate of substrate turnover is measured by change in absorbance, is proportional to LDH release, and corresponds to cellular stress due to toxicity.

E-Liquid Time Study

After performing the LDH Assay, it was determined that 2.2 mM E-liquid was an appropriate sub-lethal dose for assessment of kinase activation. The effects of 1% E-cigarette liquid on MAP kinase signaling were examined. A549 cells were treated once with 1% E-liquid and samples were collected at the following time points: 0, 15, 30, 60, 90, and 120 minutes (Figure 3). At first, the intervals were to be 15 minutes, and then the interval time was increased to 30 minutes to observe effects of prolonged time. The E-liquids (with or without 2.2 mM nicotine) (50%PG/50%VG) solutions were diluted to a final concentration of 1%, which corresponded to 2.2mM nicotine. Cells were treated as outlined in Figure 3.

After 120 minutes, wells were drained of media and washed with PBS. 150 µL of broad range protease inhibitor solution (Fisher Scientific) (27 µL of protease inhibitors, 27 µL of EDTA, and 2700 µL of lysis buffer RIPA) was pipetted into each well. The RIPA buffer caused the cells to burst so that the protein could be readily extracted. These protein samples were stored on ice throughout the extraction process.

Protein Expression: Western blots

The effects of E-liquid on MAP kinase signaling were examined by performing Western blots for the following targets: p-Erk, p

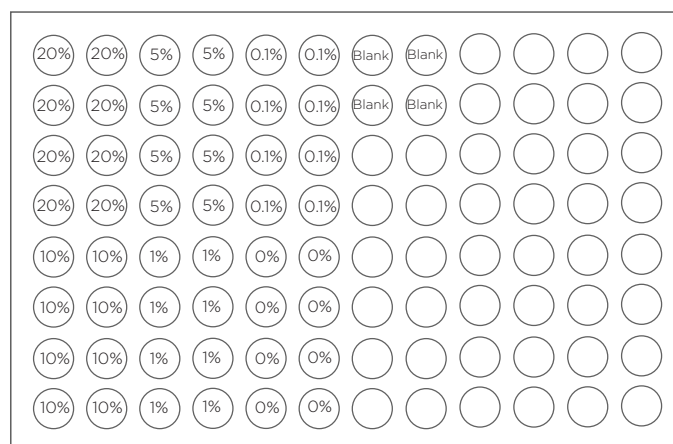


Figure 2. Design for standard 96-well plate used in the Lactic Dehydrogenase Assay.



Figure 3. Example of the experimental plate and concentrations of E-liquid used in the subtoxic assays. A, B, and C represent standard, 6-well plates containing A549 cells. Each plate was treated once with 1% (2.2 mM) E-liquid and samples were collected at the following time points: 0, 15, 30, 60, 90, and 120 minutes.

p-38, p-JNK, p38, JNK, ERK and Actin. 10 μ L of the extracted protein samples were mixed with 3 μ L of a 5X sample buffer in microcentrifugation tubes, heated to 90°C for 5 minutes, centrifuged (2,000 x g for 1 minute) and gel electrophoresis took place at 100 Volts for 60 minutes at room temperature. The proteins were extracted and transferred to a nitrocellulose membrane via three different methods: Semi-dry transfer machine (Trans-Blot® SD Semi-Dry Transfer Cell, Bio-Rad), Wet transfer machine (Wet/Tank Blotting Systems, Bio-Rad), and Turbo-transfer machine (Trans-Blot® Turbo™ Transfer Starter System, Bio-Rad). The nitrocellulose membrane was rinsed in distilled water, Sigma® Ponceau S distilled water again, and then placed in a 5% milk solution containing 2.5 μ L of lactose powder (Fisher Scientific) and 47.5 μ L of PBS solution for one hour. The primary antibody was prepared using 4 μ L of the given antibody solution (p-Erk, p p-38, p-JNK, p38, JNK, ERK and Actin) and 4 ml of a 2% BSA (Bovine Serum Albumin) Solution. The nitrocellulose membrane was placed in the primary antibody solution and left to incubate overnight with mild agitation. The next morning, the nitrocellulose membrane was removed from the primary antibody solution and rinsed in a wash buffer solution three times for 10 minutes each. The nitrocellulose membrane was placed in the prepared second antibody solution containing 4 μ L of Anti-Rabbit HRP solution and 4 ml of a 5% milk solution for two hours. For the phosphorylated proteins, 1,000 μ L each of SuperSignal® West Femto Luminol/Enhancer Solution (Thermo Scientific) and SuperSignal® West Femto Stable Peroxide Buffer (Thermo Scientific) were mixed and pipetted onto the nitrocellulose membrane. For the total proteins, 500 μ L each of SuperSignal® West Pico Luminol Enhancer Solution (Thermo

Scientific), SuperSignal® West Pico Stable Peroxide Solution (Thermo Scientific), SuperSignal® West Femto Luminol/Enhancer Solution (Thermo Scientific) and SuperSignal® West Femto Stable Peroxide Buffer (Thermo Scientific) were mixed and pipetted onto the nitrocellulose membrane. After five minutes, the nitrocellulose membrane was placed in the Bio Rad Gel Dock Station (ChemiDoc™ XRS+ System, Bio-Rad). A chemiluminescence camera was used to detect the concentration of Anti-Rabbit Horseradish Peroxidase HRP (secondary antibody) proportional to the amount of protein present.

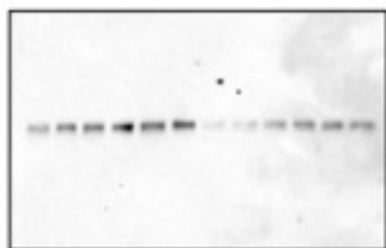
Densitometry Analysis of Proteins

The proteins on the nitrocellulose membrane were analyzed through densitometry analysis (Figure 4). Densitometry was performed on each target and represented as a ratio of pixel intensity compared to actin, using Bio-Rad Laboratories Image Lab software (version 4.0, build 16). This technique was used to quantify the level of activation of proteins by calculating a ratio of phosphorylated protein to total protein. The ratio was proportional to the level of activation of the protein. Graph Pad Prism was used to create a densitometry graph. Results were analyzed using Graph Pad Prism and a T-test, and the statistically significant difference was reported to be any value less than 0.05.

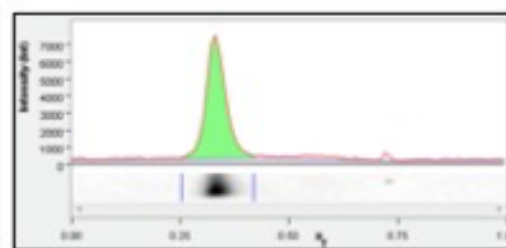
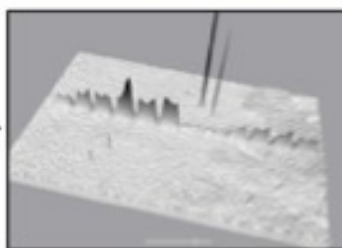
Transcription Factor Assay to Detect Activation of NF-kappaB and AP-1

TransAM™ NF-kappaB p65 (Active Motif) Assay Kit was used to detect and quantify transcription factor (NF-kappaB

1. Record Blot Exposure by chemiluminescence



2. Record band intensity by pixel intensity



3. Present result as intensity of phospho (target) protein over total protein

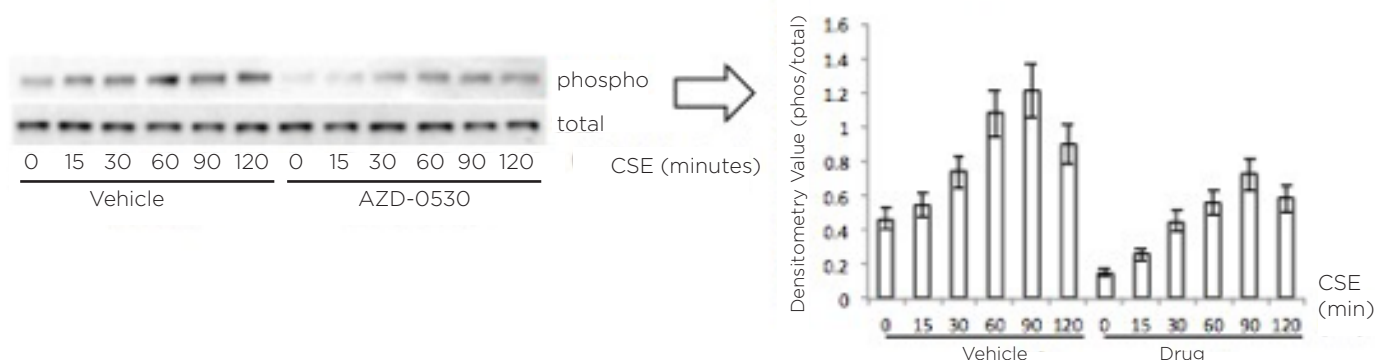


Figure 4. An example schematic diagram of the densitometry analysis. 1) Blot exposure was recorded by chemiluminescence. 2) Band intensity was recorded by pixel intensity. 3) Resulting band represents the intensity of the phospho (target) protein over the total protein.

and AP-1) activation. Simply, cell extract containing activated transcription factor was added to the oligonucleotide-coated plate. Using antibodies directed against the NF-kappaB/AP-1 p65 or c-jun AP-1 subunits and horseradish peroxidase (HRP), a secondary antibody, activated NF-kappa B or AP-1 was detected colorimetrically using a spectrophotometer. The amount of DNA binding was considered proportional to the amount of NF-kappaB activity. We did this by first adding 30µl Complete Binding Buffer and 20uL nuclear extract sample diluted in 1 mL of lysis buffer to each well. 1 µL Jurkat Nuclear extract (positive control of nuclear extract with active NF-kappaB supplied with this kit) in 19 µL of Complete Lysis Buffer was added as the positive control. 15 µL Lysis buffer was added as the negative control. The plate was then incubated for one hour at room temperature with mild agitation (100 rpm on a rocking platform). Each well was washed three times with 200 µL 1X Wash Buffer. 100 µL of diluted NF-kappaB antibody (1:1,000 dilution in 1X Antibody Binding Buffer) was added to each of the 32 wells. The plate was covered and incubated for one hour at room temperature without agitation. The wells were washed three times with 200 µL 1X Wash Buffer, as described previously. 100 µL of diluted HRP-conjugated antibody (1:1,000 dilution in 1X Antibody Binding Buffer) was added to all 32 wells. The plate was covered and incubated for one hour at room temperature without agitation. The wells were washed four times with 200 µL 1X Wash Buffer. 100 µL Developing Solution was added to all 32 wells. The plate was incubated for five minutes at room temperature protected from direct light until the sample wells turned a medium to dark blue. 100 µL Stop Solution was added. In the presence of the acid, the blue color turned yellow. Absorbance was read on a Tecan Genios spectrophotometer within five minutes at 450 nm.

RESULTS

An LDH assay was performed to examine whether E-liquid induced airway epithelial cell death. Significantly elevated levels of LDH were observed in the media from cells exposed to greater than 5% E-liquid, implying that this concentration was toxic to cells. The increase in LDH released was proportional to the amount of cell death.

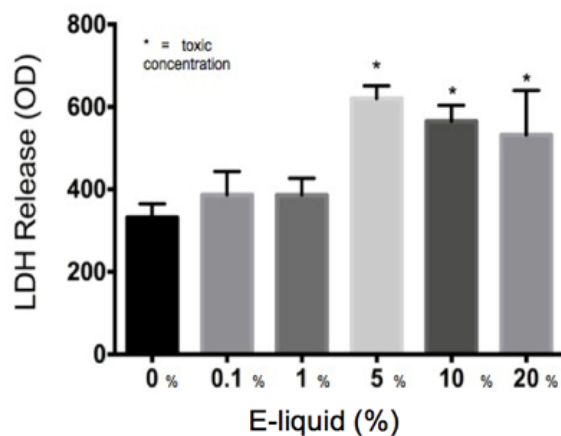


Figure 5. E-liquid (minus nicotine) is significantly toxic to cells at a concentration of 5% or higher.

Subsequently, to assess the effects of subtoxic doses of E-liquid on protein function, cells were treated with 1% E-liquid and MAP kinase activation was examined. 1% E-liquid was the highest dose shown to be non-toxic in the LDH assay. Western blots were performed for the following targets: p-ERK, p p-38, p-JNK, p38, JNK, ERK and Actin in cells exposed to nicotine alone (Figure 6) and E-liquid with nicotine (Figure 8). Actin was used as an internal control because Actin is consistently expressed, so the measure of its protein density is a useful loading control for protein. Nicotine alone and E-liquid minus nicotine were used as controls for the study.

Nicotine appeared to activate MAP Kinases. MAP Kinases were activated both by the E- liquid without nicotine, and nicotine alone. To semi-quantify the level of activation of the MAP Kinases, a ratio of phosphorylated to total proteins was calculated

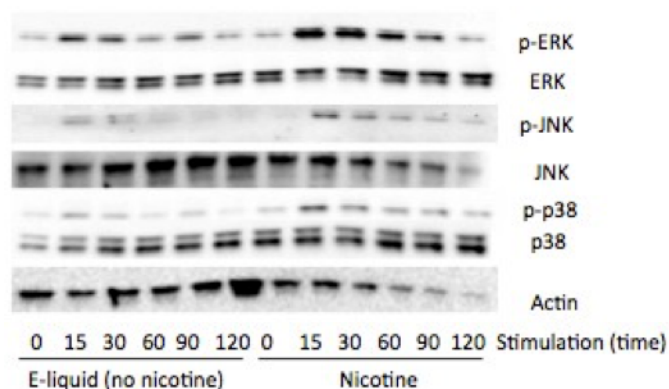


Figure 6. Western blots of MAP kinases. Thickness of the protein bands indicates activation of MAP Kinases using both liquids.

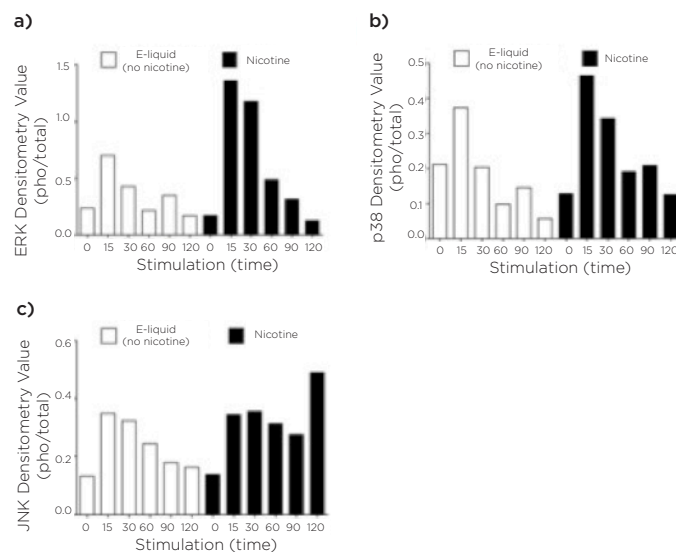


Figure 7. Semi-quantitative analysis of western blots. Time 0 is a negative control. E-liquid (no nicotine) is a control. a) ERK, b) p38 and c) JNK. ERK activation peaked at both the 15-minute mark and the 120-minute mark. p38 activation increased steadily from the 0 minute to 30 minute mark, but remained constant thereafter. JNK activation steadily increased over time but peaked at the 120-minute mark. ERK, JNK and p38 were significantly activated by nicotine alone.

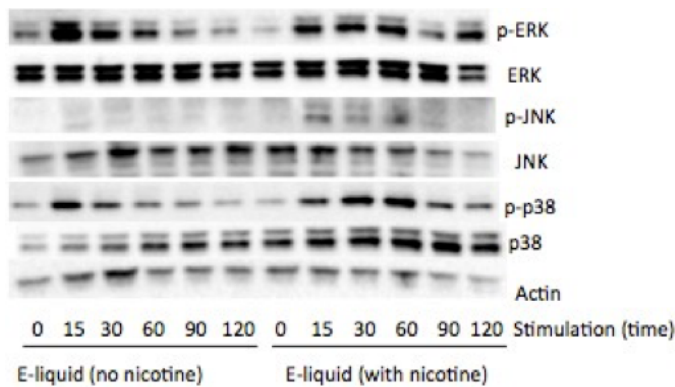


Figure 8. E-liquid with nicotine activates MAP (Mitogen-Activated Protein) Kinases. Western blots of MAP kinases, a highly conserved family of serine/threonine protein kinases involved in fundamental cellular processes such as proliferation, differentiation, motility, stress response, apoptosis, and survival. MAPKs shown below are the extracellular signal-regulated kinase (Erk), the c-Jun N-terminal kinase (JNK), and p38, as well as these phosphorylated kinases. Extracellular stimuli stimulate the activation of one or more MAP Kinase kinases (MAPKKs). MAPKKs then phosphorylate and activate a downstream MAPK kinase (MAPKK), which subsequently phosphorylates and activates MAPKs. Activation of MAPKs leads to the phosphorylation and activation of specific MAPK-activated protein kinases [10]. Activation of MAP Kinases by E-liquid, an external stimulus, is shown by the thickness of the bands of protein.

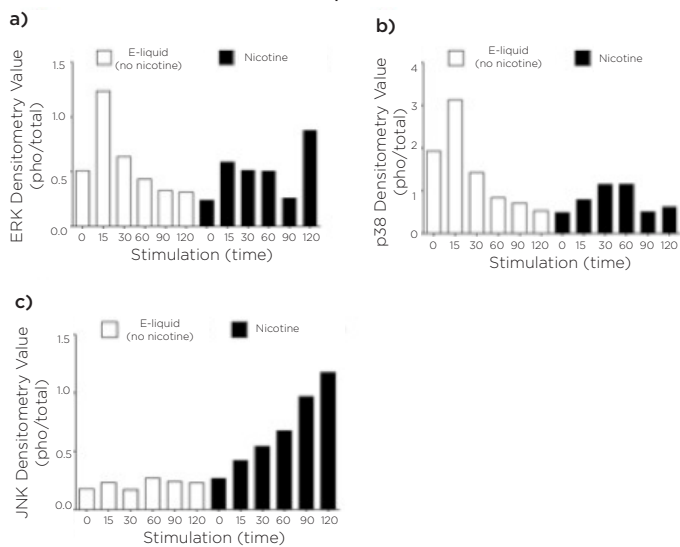


Figure 9. Semi-quantitative analysis of western blots. Time 0 is a negative control. E-liquid (no nicotine) is a control. a) ERK. b) p38. c) JNK. a) ERK activation peaked at the 15-minute mark when activated by E-liquid without nicotine. ERK activation by E-liquid with nicotine remained stable over the 15, 30, and 60 minute intervals. b) p38 activation by E-liquid without nicotine was strongest at the 15-minute mark. p38 was more strongly activated by E-liquid without nicotine. c) Activation of JNK by E-liquid with nicotine was stronger than activation by E-liquid without nicotine. E-liquid with nicotine induced activation of JNK steadily over time.

through densitometry analysis (Figure 7). ERK activation peaked at both the 15-minute mark and the 120-minute mark (Figure 7). p38 activation increased steadily from the 0 minute to 30-minute mark, but remained constant thereafter (Figure 7). JNK activation steadily increased over time (Figure 7). ERK, JNK and p38 were significantly activated by nicotine alone. JNK activation remained relatively stable at the 15 and 30 minute marks, and peaked at the 120-minute mark.

To assess the effects of E-liquid containing nicotine on transcription factor activation, NF-kappaB and AP-1 activations were measured using a Transcription Factor Assay (Figure 10). E-liquid containing nicotine significantly increased activation of NF-kappaB, but not AP-1.

DISCUSSION

The potential impact of safer alternatives to cigarettes products has major public health implications. However, many alternatives on the market have yet to be proven safe. This study was an attempt to assess the toxicity of E-cigarette cigarette vapors (E-liquid) on lung cells and the subsequent impact on inflammation mediators, such as kinases and transcription factors. Our data show that E-liquid significantly induces cell death, activates MAP kinases ERK, p38, JNK and the transcription factor NF-kappaB, suggesting that signaling associated with inflammation is occurring when the cells are exposed to E-cigarette vapor. Cigarette smoke is a known activator of MAP kinases [1], so this finding is not surprising and suggests that nicotine may be involved in kinase activation as both E-liquid and cigarette smoke contain nicotine. Future studies examining whether the degree of protein activation is greater due to nicotine would be useful. Additionally, *in vivo* studies in animals would enhance the results of this study and suggest potential human implications of using E-cigarettes.

The data also strongly demonstrate that NF-kappaB is activated in cells treated with E-liquid. NF-kappaB can regulate the activation of many detrimental genes, such as cytokines and

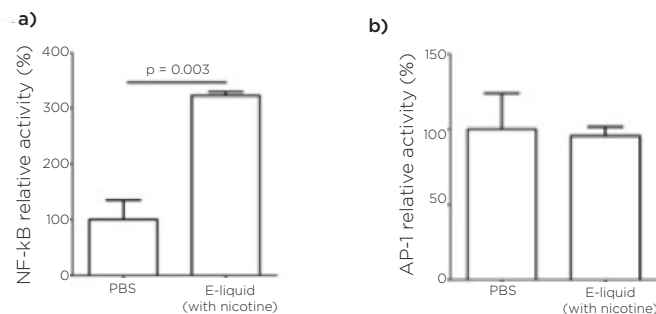


Figure 10. Transcription Factor Assay for the percentage of NF-kB (a) and AP-1 reductase (b) activities following E-liquid exposure. The bar connecting the two columns denotes a statistically significant difference ($p < 0.05$). a) Significant result indicates that the difference observed between these two groups was not random. NF-kB activity was greater when the cells were treated with E-liquid with nicotine as opposed to PBS. b) Statistically insignificant result suggests that the difference across the two groups was due to random chance.

proteases, that are associated with progression of COPD. This study was recently performed in the laboratory of Dr. Foronjy. Dr Foronjy [unpublished data] and demonstrated that when lung epithelial cells are treated with E-liquid, cytokine and protease gene expression (e.g. MMP-7, MMP-9, Cathepsin L1, TLR3, c-Src) associated with disease progression increases. The results supported the hypothesis that E-liquid activates MAP kinases and, subsequently NF-kappaB, inducing cytokine and protease gene production. In future experiments ERK, p38 and JNK could be knocked down using silencing RNA (siRNA), and cytokine and protease levels could be analyzed. These experiments could determine whether the aforementioned MAP kinases are integral to the E-cigarette mediated induction of responses associated with lung damage.

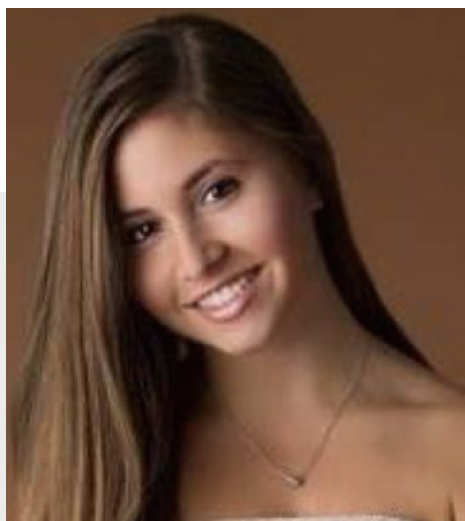
Cigarette smoke is known to be a major risk factor for numerous diseases, including Chronic Obstructive Pulmonary Disease (COPD) [6]. This project suggests that E-liquid used in E-cigarettes could also be potentially harmful to human health, and contribute to COPD progression. This study attempted to assess the lung toxicity of E-liquid *in vitro* and determined that 1% (2.2 mM) E-liquid was a subtoxic concentration of E-liquid. Additionally, this study established that E-cigarette liquid can induce inflammatory signaling by activating MAP Kinase pathways and the transcription factor NF-kappaB, which regulates a number of immune and inflammatory responses associated with COPD [11, 12]. Future assessment of E-cigarette health effects *in vivo* in both animal and human models would be useful in understanding the long-term effects of E-cigarette usage on disease. Furthermore, *in vivo* work may allow scientists to identify the specific proteases and cytokines that could serve as biomarkers for inflammation or disease. Medication to control these protein levels may be developed accordingly, but the best approach would be to minimize the use of tobacco products independent of their delivery method.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Dr. Foronjy for his time and enthusiasm in supervising this project. Moreover, I thank Dr. Geraghty and Dr. Turino for their guidance and encouragement.

REFERENCES

1. Mercer, B. A, D'Armiento. (2006). J. M. Emerging role of MAP kinase pathways as therapeutic targets in COPD. *Int J Chron Obstruct Pulmon Dis.* 1(2): 137–150. Published online 2006 June.
2. Pahl, HL. (1999). Activators and target genes of Rel/NF-kappaB transcription factors. *Oncogene.* 18(49): 6853-66.
3. Pellegrino RM, Tinghino B, Mangiaracina G et al. (2012). Electronic cigarettes: An evaluation of exposure to chemicals and fine particulate matter. *Annali di igiene: medicina preventiva e di comunita*, 24: 279-288.
4. Goniewicz ML, Knysak J, Gawron M et al. (2013). Levels of selected carcinogens and toxicants in vapour from electronic cigarettes. *Tob Control.*
5. "What is COPD?" (2012). *National Heart Lung and Blood Institute.* U.S. Department of Health & Human Services. Web.
6. Andrew J Thorley and Teresa D Tetley. (2007). Pulmonary epithelium, cigarette smoke, and chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis.* 2(4): 409–428. Published online 2007 December. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2699967/>
7. Kim V, Criner GJ. (2013). Chronic bronchitis and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med.* 187: 228-237.
8. "Chronic Obstructive Pulmonary Disease (COPD)." (2012). *World Health Organization.* Web.
9. Jha, P. (2011). Avoidable Deaths from Smoking: A Global Perspective. *Public Health Reviews* (2107-6952), 33(2).
10. "MAP Kinase Signaling Resources." *Cell Signaling Technology.* (2016). Web.
11. Reynolds PR, Cosio MG, Hoidal JR. (2006). Cigarette smoke-induced Egr-1 upregulates proinflammatory cytokines in pulmonary epithelial cells. *Am J Respir Cell Mol Biol.* 2006 Sep;35(3):314-9. Epub.
12. Moodie FM, Marwick JA, Anderson CS, et al. (2004). Oxidative stress and cigarette smoke alter chromatin remodeling but differentially regulate NF-kappaB activation and proinflammatory cytokine release in alveolar epithelial cells;18:1897–9.



HAYLEY TARTELL

Hayley Tartell, a sophomore at Williams College, intends to pursue a double major in Chemistry and English. A self-starter with considerable initiative and focus, she always strives to learn as much about as many things as possible. Intellectually curious and inquisitive, Hayley is passionate about neuroscience and aspires to become an MD/PhD. Her interests include reading, creative writing, tennis, piano, and running.

Carbon Sequestration and Microbial Communities in Mangrove Ecosystems

Daryll Carlson

Stanford University

Mangrove forests play a major role in the global carbon budget by sequestering large amounts of carbon in their sediments. Their ability to sequester carbon is influenced by the impact of varying salinity levels on the microbial communities within the soil, whose health is a determinant of the rate of soil organic carbon loss in these ecosystems. This study examines the mangrove habitats of North Stradbroke Island off the coast of Queensland, Australia to understand how alterations in mangrove biogeochemistry due to global environmental changes may potentially impact their ability to sequester carbon in the future. At two sites, water samples were taken in each of three salinity zones for salinity assessment, and sediment cores were collected for carbon content measurement by loss on ignition analysis as well as for quantification of microbial community species richness, diversity, and catabolic activity utilizing Biolog Ecoplates. Results showed a negative correlation between salinity and both soil microbial community species richness and diversity, as well as a positive correlation between average microbial catabolic activity and both microbial species richness and diversity. These trends suggest that high salinity extremes create soil conditions that are not as suitable for microbial functions in comparison to conditions in zones of low salinity, but also that high levels of microbial diversity and species richness can be at the detriment of carbon sequestration due to higher rates of catabolic activity. Finally, it was shown that there is a significant difference in microbial catabolic activity between different carbon sources, exhibiting the role that carbon source can play in the scope of microbial species richness and diversity that can interact with soil organic carbon in a specific area. This data demonstrates the necessity of maintaining the balances of salinity and microbial community composition in mangrove ecosystems in protecting their capacity for storing carbon and reducing future soil organic carbon loss.

Introduction

Mangrove forests that occur along tropical and subtropical coastlines across the globe contribute a vital set of ecosystem services, including providing support for biodiversity and fisheries, assisting in wave attenuation, supporting nutrient cycling, and trapping large amounts of sediments. Furthermore, mangroves are able to sequester large quantities of carbon in their sediments [1]. Mangroves are capable of accumulating approximately 30-110 kg C m⁻², while adjacent land only accumulates 5-10 kg C m⁻² in the same period of time [2]. As these ecosystems cover large areas of coastlines around the world, their prevalence and exceptional ability to sequester carbon make mangrove habitats a major component of the global carbon budget.

Due to their ability to sequester carbon, mangrove forests are an important resource to monitor in the face of current and future rises in global temperature and CO₂ emissions. With global climatic change, mangrove habitats could experience many environmental disruptions that alter their biogeochemistry and affect their ability to sequester carbon. Extreme tidal events, storm surges, and freshwater surges resulting from large precipitation events are predicted to increase in both frequency and magnitude [3]. Anthropogenic impacts such as pollutant input, nutrient

runoff, and hydrologic modification can have substantial impacts on mangrove habitats. Global warming and sea level rise are also expected to promote coastal wetlands submergence and salinity intrusion [7]. In light of the potential for mangrove habitat disruption, it is important to understand how these events can affect mangrove carbon sequestration.

Organic carbon sequestration is a complex process influenced by many factors. Sediment deposition and biomass production act as carbon inputs, providing organic carbon to be sequestered in the soil, while mineralization and microbial catabolic activity cause sequestered carbon to be lost in atmospheric or other forms. These processes are all highly influenced by salinity conditions [3]. In particular, levels of catabolic activity are affected by the composition and health of microbial communities in the soil, which are highly sensitive to changes in salinity levels due to different biological constraints [3]. The level of microbial catabolic activity in turn influences the rate of soil organic carbon loss, which counteracts carbon sequestration. Increasing salinity results in an increase in ionic strength in soil porewater and a sulfate effect, as sulfate⁻ is abundant in seawater. High ionic strength can disrupt the cellular functions of microbes or may result in lysis by inducing osmotic stress in microorganisms, resulting in reduced catabolic activity [3]. On the other hand, sulfate can act as an alternate electron acceptor during anaerobic respiration, causing

a short-term increase in the rate of soil CO₂ flux and resulting in increases in catabolic activity that are directly proportional to amounts of added sulfate [3]. Freshwater counteracts this increase in sulfate, but also has been shown to provide better conditions for microbial respiration [4]. Modifications to these microbial communities are likely to have a large impact on mangrove carbon exchange [5]. For this reason, studying the influence of salinity on the structure and activity of mangrove microbial communities is critical to better understanding how mangroves' ability to sequester carbon is influenced by differing environmental conditions.

This study examines the mangrove forests on North Stradbroke Island, one of several sand barrier islands that enclose the subtropical embayment of Moreton Bay in southeast Queensland, Australia. The inter-tidal zone along the western side of North Stradbroke Island is characterized by large areas of mangrove habitat, which primarily include the species *Avicennia marina* and *Rhizophora stylosa*. Because Moreton Bay has one of the most rapidly growing populations of any region of Australia, the coastal mangrove habitats of North Stradbroke Island are under intense anthropogenic pressure [6]. By studying the mangrove areas of North Stradbroke Island, the purpose of this study was to elucidate the complex relationships within mangrove soils that influence carbon sequestration in order to better understand how anthropogenic pressure and climate change could impact the ability of mangrove habitats to sequester carbon in the future.

METHODS

Two mangrove sites were sampled on North Stradbroke Island (27°32' S, 153°27' E), one with the dominant species *Avicennia marina* (Amity Point, 27°24'24" S, 153°26'25" E) and one with the dominant species *Rhizophora stylosa* (Myora Springs, 27°27'54" S, 153°25'03" E). At each site, three samples were taken in each of three zones: an area of high salinity (29–35ppt) near the shore, an area of low salinity (1–7ppt) in the mangrove near the transition zone with the saltmarsh, and in intermediate salinity areas (8–28ppt). Locations were selected after preliminary assessment of porewater salinity, or the salinity of groundwater within the soil, in the field. Sediment cores were taken for later measurements of carbon content and microbial community species richness, species diversity, and catabolic activity using modified syringes (27cm³–52cm³ of wet sediments). Water samples (5mL) were taken at the same locations for salinity assessment, which was measured using a handheld refractometer.

Sediment samples were measured for initial wet mass, and approximately 2g (wet weight) were separated from each sample for analysis of microbial communities. Sediments were oven dried at 90°C and reweighed for dry mass after 12 h, 24 h, and 30 h of drying until there was no longer a significant change in mass. Organic carbon content in each sample was estimated using loss on ignition (LOI) mass estimation [8]. Once dry, samples were placed in a muffle furnace for 4 hours at 550°C and reweighed for a post-ignition mass. The LOI masses were standardized by calculating the mass per sediment volume for each sample.

The analysis of microbial communities was conducted on 2 gram sediment subsamples from each sediment core using Biolog Ecoplates (BIOLOG Inc., Hayward, CA, USA). Sediments were suspended in 20mL of filtered seawater and grinded with a Tissue Tearor for 30 seconds followed by centrifugation at 1000rpm

and 25°C for 30 seconds to remove sediment. The supernatants were placed in Biolog Ecoplates, each of which contains a set of 31 carbon sources ecologically relevant to soil microbial communities that are repeated three times [5]. Supernatant from each sample was distributed amongst one replicate of the 31 carbon sources in 150µL aliquots. Bacteria in the wells produce NADH during catabolic usage of the carbon, which reduces a dye and produces color in the wells, which can then be measured using a spectrometer. Plates were then run through a plate reader (Bio-Rad Laboratories, Model 680 Microplate Reader, using Microplate Manager 5.2 software) at 595nm at times of 0, 2 h, 12 h, 24 h, and 48 h to examine the catabolic activity of the microbial communities within the samples [9].

Catabolic activity for each carbon source (A_i) was defined as the corrected absorbance value for that well at a time of 48 h. Mean overall catabolic activity per sample was calculated as the average of these A_i values. Negative A_i values were set to zero, as this denotes an absence of catabolic activity in the carbon source. For each sample, A_i values were used to calculate species richness S and the Shannon index of diversity H, which measures microbial species diversity while accounting for both species abundance and evenness, using previously defined formulas in Stephan et al. 2000 that utilize A_i value to quantify these factors [9].

Linear regression was used to examine the relationships between salinity, species richness, soil organic carbon content, Shannon index of diversity, and average catabolic activity. Two-way ANOVA tests were performed in order to identify significant differences among differing conditions, using as fixed factors dominant tree species (*Avicennia marina* or *Rhizophora stylosa*) and salinity zone (low, intermediate, or high). The response variables tested were the Shannon index of diversity, species richness, and average catabolic activity. To test for differences within sites (Myora Springs and Amity Point) in the response variables due to salinity (low, intermediate, and high areas), a one-way ANOVA test was performed. Assumptions were tested using Bartlett's test and the Shapiro-Wilk test. If assumptions were not met, data was log transformed [10].

Following ANOVA analysis, mean comparisons between catabolic activity levels (A_i) of differing carbon sources were performed using Tukey's test in order to determine which carbon sources may be more biologically relevant, and potentially be more frequently lost to microbial activity as atmospheric CO₂. This information was then applied to separate carbon sources with significantly different mean catabolic activity into groups (A-F). Carbon sources were also organized into five categories: carbohydrates, polymers, carboxylic and acetic acids, amino acids, and amines/amides [11].

RESULTS

The two-way ANOVA tests performed for the fixed factors of dominant tree species and salinity zone with the response variables of Shannon index of diversity, species richness, and average catabolic activity all yielded p-values greater than .05, so dominant tree species was considered insignificant, and the data was pooled together between the two sites for the remaining statistical analysis.

Soil samples had an average initial dry weight of 31.12 ± 1.54 grams, and an average post-ignition mass of 28.05 ± 1.60 grams. However, the carbon content estimated by this loss of mass on

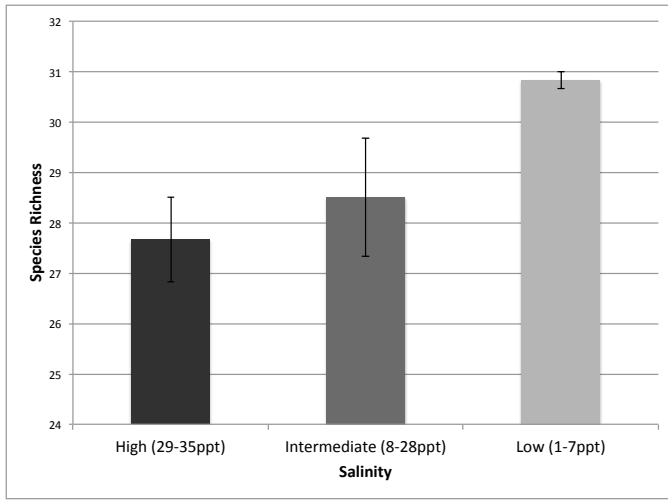


Figure 1. Figure 1. Average richness of microbial species across different salinity zones. Species richness data is shown as means across six samples for each zone, with standard error values given.

ignition did not show any statistically significant relationship with salinity conditions or microbial community demographics.

Microbial species richness tended to decline from low, intermediate, to high salinity zones, and was shown to be statistically significant (p -value = .0459). The resulting average species richness values in each zone from the data collected were 30.83 ± 0.16 , 28.50 ± 1.17 , and 27.67 ± 0.84 respectively (Figure 1). Linear regression analysis shows similar trends in the relationship between salinity and Shannon index of diversity, with levels of diversity decreasing as salinity increased across zones ($R^2 = 0.279$, p -value = 0.02423). Higher levels of diversity were found in low salinity zones, with an average Shannon index of diversity of $3.28 \pm .029$ as compared to $3.12 \pm .039$ in salinity areas between 29 and 35ppt.

Linear regression analyses showed positive correlations between average catabolic activity and both the species richness ($R^2 = 0.578$, p -value = 0.00025) and diversity ($R^2 = 0.591$, p -value = .00019) of microbial communities in the soil samples tested (Figure 2).

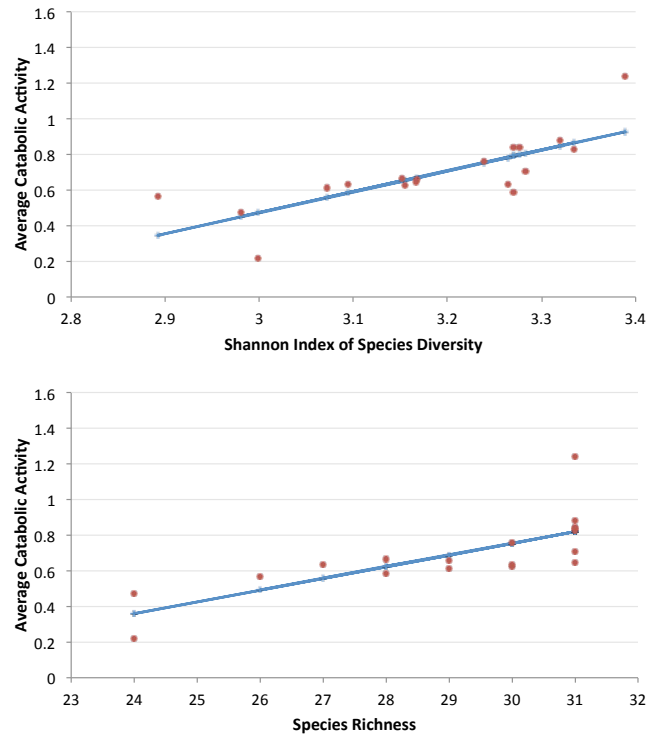


Figure 2. Linear correlations between average catabolic activity and both microbial species richness and diversity across soil samples.

Data also showed statistically significant differences in microbial catabolic activity between various carbon sources, including carbohydrates, polymers, carboxylic and acetic acids, amino acids, and amines/amides (p -value < 0.0001). Group A, the Tukey's test grouping with the highest mean catabolic activity, was dominated by carbohydrates, with a significant proportion of polymers as well. Carboxylic and acetic acids dominated groups E and F on the other end of the spectrum. 5 of 6 amino acid sources were placed in group C, in the upper half of the average catabolic activity values (Figure 3).

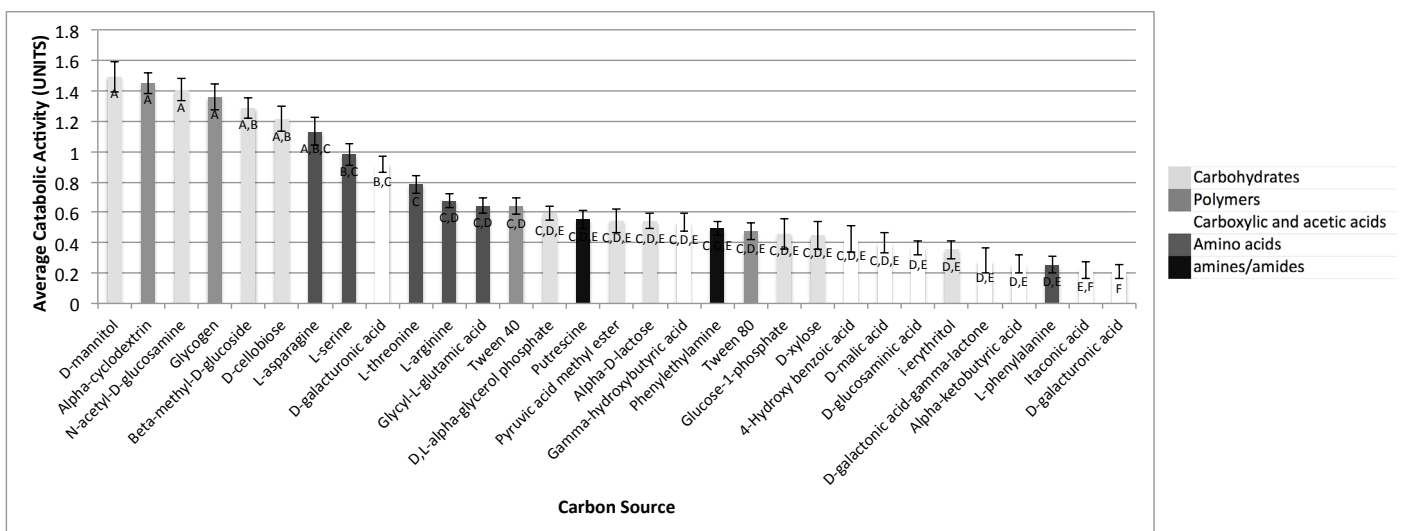


Figure 3. Average catabolic activity of different carbon sources, grouped based on mean comparisons through Tukey's test. Sources are categorized into carbohydrates, polymers, carboxylic & acetic acids, amino acids, and amines/amides (Weber and Legge 2009).

DISCUSSION

The relationship between salinity and microbial community composition, species richness, diversity, and catabolic activity has a major role in determining mangrove capacity for carbon sequestration. With global anthropogenic factors rapidly changing the natural environment, it is important to understand how these alterations will impact the world's habitats and their ability to provide vital ecosystem services. As mangroves may potentially face drastic impacts with the continuation of global climate change, comprehending these complex relationships and how they may be affected could help to reveal potential changes in mangrove carbon sequestration, an important factor of global carbon budget that could occur with future environmental changes.

The results of this study showed a negative correlation between salinity and both soil microbial community species richness and diversity. Low salinity zones had significantly higher values of species richness than high salinity zones. Higher levels of diversity in microbial communities were also found in these low salinity zones. This suggests that high salinity extremes produce soil conditions that are not as suitable for microbial functions, resulting in fewer soil microbes and a less diverse community dominated by species capable of surviving in hypersaline environments. Other studies have supported this finding, with evidence that ions present in seawater disrupt specific microbial cellular processes in organisms not specially adapted to these conditions [3]. Chloride, the most abundant anion in seawater, is a potential inhibitor of both nitrification and denitrification in soil microbial communities, which are important in enabling bacterial respiration and other biological processes [12]. It is predicted that with global warming and sea level rise, mangroves will experience increased submergence and salinity intrusion [7]. Such impacts have the potential to detrimentally influence levels of microbial species richness and diversity, causing areas of the intertidal zone that were previously suitable for certain microbial species to survive in become uninhabitable. As microbial communities are vital to nutrient cycling and making these sources biologically available, this could be devastating to mangrove communities.

Data from this study also suggests that there is a positive correlation between average catabolic activity and both microbial species richness (p -value = .00025) and diversity (p -value = 0.00019). Microbial communities with greater richness and diversity are able to catabolize carbon at a more rapid rate because there is much higher demand for carbon. A more diverse array of microbial species is likely to be able to target a wider range of carbon sources as they fulfill differing niches, and a greater number of microbial individuals will catabolize higher amounts of soil organic carbon. High levels of catabolic activity can be problematic in areas that are important for carbon sequestration. Commonly, 40-60% of carbon taken up by microorganisms is immediately released as atmospheric CO_2 , meaning that high levels of catabolism can increase soil organic carbon loss [13]. Mangrove forests experiencing either increased freshwater pulses from large precipitation events and storm surges or a lower salinity due to physical hydrologic modification with future environmental change are susceptible to increased soil organic carbon loss. The lower salinity levels in the mangrove could increase microbial richness and diversity [4], which would in turn increase catabolic activity, ultimately causing increased organic

carbon loss in the soil. Such changes may disrupt the role of these mangrove habitats in carbon sequestration. This provides an interesting conundrum, as low salinity levels are associated with high levels of catabolic activity that may disrupt the role of mangrove habitats in providing ecological services such as carbon sequestration, but high salinity levels are not optimal for allowing the growth of diverse and rich microbial communities that are vital to nutrient cycling. This emphasizes the importance of the natural gradients within mangroves that provide a range of salinity conditions that allow for a balance between these two sides, but have the potential to be drastically altered by tidal inundation and other hydrological modification.

Finally, results from this study showed dramatic differences in microbial catabolic activity among various carbon sources, despite all of their ecological relevance. The highest levels of catabolic activity were recorded in carbohydrates and polymers, and very low levels were found in carboxylic and acetic acids. D-mannitol, alpha-cyclodextrin, N-acetyl-D-glucosamine, and glycogen were among group A, the Tukey's test grouping with the highest mean catabolic activity, while group F, the group with the lowest mean, was composed of itaconic acid and D-galacturonic acid. This information may speak to which carbon source groupings may be the least metabolically expensive for uptake, or most readily broken down. Carbon sources with higher levels of catabolic activity may also reveal specific adaptations within these particular microbial communities that allow for more efficient or rapid catabolism of certain sources. These carbon sources with the highest microbial catabolic activity will be mostly rapidly lost as atmospheric carbon through respiration, and so may not be as readily sequestered in areas with higher microbial species richness or abundance. Future study of which carbon sources are lost in soil organic carbon loss most frequently may be useful in understanding which specific branches of the carbon cycle are being disrupting by increased soil organic carbon loss in these ecosystems.

There were some limitations in the study as a result of the sampling methods used. Soil samples taken with modified syringes had variation in depth and did not include deeper sediment, which could potentially contain microbial communities that play a large role in carbon cycling. Utilizing Biolog Ecoplates to study and quantify microbial communities limited analysis to bacteria able to grow on plates. Allowing the plates to incubate for only 48 hours may not have shown the full potential for microbial catabolic activity. In future work studying additional sites with different salinity conditions or microbial community compositions would be useful to garner a more thorough understanding of the impact of salinity and other factors on mangrove carbon sequestration.

This study demonstrates the importance of the balance between salinity and microbial community composition within mangroves in ensuring that these habitats are capable of carbon sequestration in the future. Mitigation of human impacts on mangrove forests is necessary in order to preserve the mangroves' role in global carbon storage. Furthermore, to maintain the biological functionality of these ecosystems, protecting microbial species richness and diversity through some mitigation of the impacts of global warming and sea level rise is an important future action in mangrove conservation. These alterations are difficult to address at a local scale, as they often result from global climatic changes; studies have cited global impacts such as sea level rise and temperature change as some of the most

pertinent threats to these ecosystems currently [3,7]. For this reason, mangrove conservation will likely need to involve large-scale and cooperative initiatives for their protection. Much of this work has already been initiated, with mangrove restoration and replantation to maintain their structure and function, the introduction of legislation to prevent further destruction and hydrological modification of these habitats, and the declaration of some mangrove ecosystems as protected zones to reduce other environmental stressors, however there is a great deal of room for more research and cooperation. In addition, insights on which forms of carbon are most biologically useful can help determine which forms are most often sequestered and least impacted by microbial activity. This provides information as to what carbon products mangrove ecosystems may be most able to store, and which portions of carbon cycling are most drastically altered by changes in microbial activity. Overall, understanding how global climate changes and localized impacts may alter conditions in mangroves is vital to protecting their capacity for storing carbon and reducing soil organic carbon loss in the future.

ACKNOWLEDGEMENTS

The author wishes to thank Dr. Catherine Lovelock for her continued assistance throughout this project, as well as the University of Queensland for their added support.

REFERENCES

1. Lovelock CE. Soil Respiration and Belowground Carbon Allocation in Mangrove Forests. *Ecosystems* 2008; 11.2:342-54.
2. Coultas CL. Soils of the intertidal marshes of Florida's Gulf Coast. In: Coultas CL, Hesieh Y-P, eds. *Ecology and management of tidal marshes*. Boca Raton: CRC Press, 1996:61-69.

3. Chambers LG, Osborne TZ, Reddy KR. Effect of Salinity-altering Pulsing Events on Soil Organic Carbon Loss along an Intertidal Wetland Gradient: A Laboratory Experiment. *Biogeochemistry* 2013; 115.1-3:363-83.
4. Jolly ID, Kl. McEwan. A review of groundwater-surface water interactions in arid/semi-arid wetlands and the consequences of salinity for wetland ecology. *Ecology* 2008; 1:43-58.
5. Adame MF, Reef R, Herrera-Silveira JA et al. Sensitivity of Dissolved Organic Carbon Exchange and Sediment Bacteria to Water Quality in Mangrove Forests. *Hydrobiologia* 2012; 69.1:239-53.
6. Manson FJ, Loneragan NR, Phinn SR. Spatial and Temporal Variation in Distribution of Mangroves in Moreton Bay, Subtropical Australia: A Comparison of Pattern Metrics and Change Detection Analyses Based on Aerial Photographs. *Estuarine, Coastal and Shelf Science* 2003; 57.4:653-66.
7. Mendelsohn IA, Sorrell BK, Brix H et al. Controls on Soil Cellulose Decomposition Along a Salinity Gradient in a *Phragmites Australis* Wetland in Denmark. *Aquatic Botany* 1999; 64.3-4:381-98.
8. Heiri O, Lotter AF, Lemcke G. Loss on Ignition as a Method for Estimating Organic and Carbonate Content in Sediments: Reproducibility and Comparability of Results. *J Paleolimnology* 2001; 25.1:101-10.
9. Stephan A, Meyer AH, and Schmid B. Plant Diversity Affects Culturable Soil Bacteria in Experimental Grassland Communities. *J Ecology* 2000; 88.6.
10. Bartlett, M. S., and D. G. Kendall. The Statistical Analysis of Variance-heterogeneity and the Logarithmic Transformation. *J R Stat Soc* 1946; 8.1:128-138.
11. Weber KP, Legge RL. One-dimensional Metric for Tracking Bacterial Community Divergence Using Sole Carbon Source Utilization Patterns. *J of Microbiological Methods* 2009; 79.1:55-61.
12. Seo DC, Yu K, Delaune RD. Influence of salinity level on sediment denitrification in a Louisiana estuary receiving diverted Mississippi River water. *Arch Agron Soil Sci* 2008; 54:249-257.
13. Jackson, LE. Soil Microbial Communities and Carbon Flow in Agroecosystems. In: *Ecology in Agriculture*. San Diego, CA: Academic, 1997:367-413.



DARYLL CARLSON

Daryll is currently a junior at Stanford University majoring in biology, with a concentration in marine biology. Her academic interests include earth systems, oceanography, environmental policy, and marine sciences. She also enjoys photography as both an academic and personal interest in her spare time. After graduating, Daryll plans to eventually pursue a Ph. D. in marine biology, and a career in academic research.

Expanding the Chromatic Range of Galaxies for Weak Gravitational Lensing Simulation

Bradley Emi

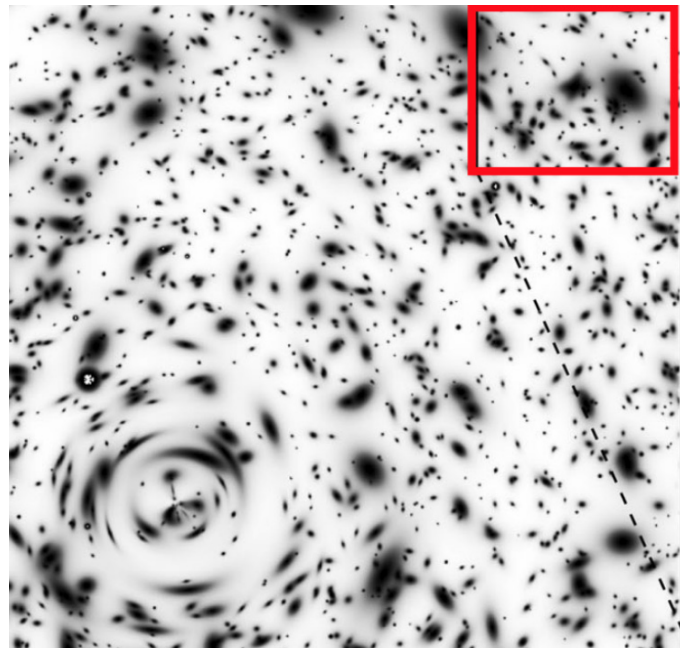
Stanford University

Highly accurate image simulations are needed to study weak gravitational lensing, a tool used in observational cosmology to detect large-scale structure such as dark matter and the causative agent of the observed accelerated expansion of the universe, known as dark energy. Weak gravitational lensing is based on predictions from general relativity that explain how light from distant galaxies bends around large-scale structure. The success of future missions to detect large-scale structure is dependent upon the ability to constrain error in the measurement of galaxy shapes. This project uses a deep sample of Hubble Space Telescope (HST) galaxies in multiple wavelengths to study wavelength-dependent (chromatic) effects of the point-spread-function (PSF) on images of the same galaxies, as they would appear in ground-based telescopes. Because the PSF can differ across images as a function of the spectral energy distribution of the sources, it is important to apply a systematic correction to images based on wavelength. We provide a sample of over 10,000 parametric models based on real galaxies taken in multiple chromatic filters of the HST's Advanced Camera for Surveys Wide Field Channel (ACS-WFC) from the All-Wavelength Extended Groth Strip International Survey (AEGIS). We incorporate these models into GalSim, an open-source image simulation platform for use by the wider astronomical community. The code is also made publicly available for work with future data sets at <https://github.com/bradleyemi/ChromaticGalaxies>.

Introduction

Observations from the last two decades show that the universe is expanding at an accelerating rate. However, according to general relativity, if the universe is truly expanding at an accelerated rate, the curvature of spacetime (and equivalently, the gravitational field) cannot be entirely determined by our existing understanding of mass and energy. The most accepted theoretical hypothesis for resolving this problem, in accordance with the accelerating expansion of the universe and general relativity, is dark energy. If general relativity is correct, it necessitates a new component of energy with negative pressure: this is what is meant by “dark energy.” We are fundamentally interested in knowing whether the accelerating expansion of the universe is due to dark energy, and if so, what the physical properties of dark energy are, or, if not, if general relativity breaks down on cosmological scales.

Wallabies are herbivores and feed mainly on grass. During perProbing dark energy benefits from an understanding of how the shape of the dark matter distribution evolves over time. However, dark matter cannot be observed directly. According to general relativity, large dark matter distributions will bend spacetime between Earth and distant galaxies. The location of dark matter can be determined by observing the distorted shapes of the distant galaxies. Formally, this technique is known as weak gravitational lensing. Weak gravitational lensing takes advantage of the fact that photons travel along the geodesics of spacetime, which are altered by the local gravitational field, determined by the local mass-energy distribution. Figure 1 shows an example



Figures 1. An example of strong gravitational lensing. Gravitational lensing is caused by the bending of spacetime in between an observer and an astronomical object, ordinarily caused by large matter distributions. Weak gravitational lensing occurs when the shape distortion is much less than the intrinsic shape, even less distorted than the galaxies in the red box. Observations of the curvatures of these galaxies allow us to infer the presence of dark matter.

of strong gravitational lensing to demonstrate how large mass distributions can distort background objects. Weak lensing by contrast is much more subtle than the effect on the galaxies in the red box.

By observing the correlated curvature of background galaxies, we can indirectly observe the distribution of the large-scale structure of the universe. With photometric redshifts, we can also determine the growth of the large-scale structure of the universe as a function of time. This information allows us to constrain cosmological parameters and make inferences about how both gravity and dark energy have affected the evolution of structure in the universe.

Given a galaxy shape, orientation, and a gravitational field, it is not difficult to predict the corresponding image. However, the reverse task, predicting the gravitational field from a series of images of galaxies, is a highly nontrivial problem. This is because there are a number of factors contributing to the shape of a galaxy in an image: intrinsic galaxy shape, physical orientation of the galaxy relative to the observer, and a series of systematic atmospheric, thermal, and optical effects contributing to what is defined as a point-spread-function (PSF). Furthermore, the signal from weak gravitational lensing is typically barely detectable over intrinsic noise thresholds.

It is true that we cannot know the intrinsic galaxy shapes and orientation, since we only see the image of the galaxy after it is gravitationally distorted. From one galaxy alone, it is not possible to conclusively determine a gravitational lensing signal. [1] Fortunately, we can exploit the fact that the universe is isotropic on large enough scales, which means that the universe looks the same in every direction. However, assuming an isotropic universe, which follows from the cosmological principle and is consistent with observations to date, we expect the distribution of orientations and shapes to be random. If we notice a correlated lensing signal between a large number of galaxies, however, we can infer that it is due to gravitational lensing. As a result, a large number of observations are needed to beat the “shape noise” of the galaxies’ intrinsic shapes. The Hubble Space Telescope provides a significant number of images for our purposes, and future missions, including the space-based Wide Field Infrared Survey Telescope (WFIRST) and Euclid, and several ground-based telescopes, including the Large Synoptic Survey Telescope (LSST) will allow us to expand the dataset of galaxies with good shape measurements from millions to billions.

The goal of this study is to minimize systematic error associated with the PSF by examining the PSF effects on both ground-based and space-based telescope images. Since we need a large number of galaxies for weak lensing analysis, galaxies from both ground- and space-based telescope surveys are used, and the ability to directly compare galaxies from both types of observations is essential.

One way to test the data analysis pipelines for future weak lensing studies is by creating highly accurate image simulations, with and without weak lensing effects. The current platform for such simulations is known as GalSim, which is available open-source [2]. GalSim currently has the capability to generate galaxy models and simulate weak lensing effects, but it also contains a basis set of noise-free galaxy and PSF models based on Hubble Space Telescope images from the Cosmological Evolution Survey (COSMOS) [3]. This basis set of galaxies ensures that data processing software is adaptable to a realistic distribution

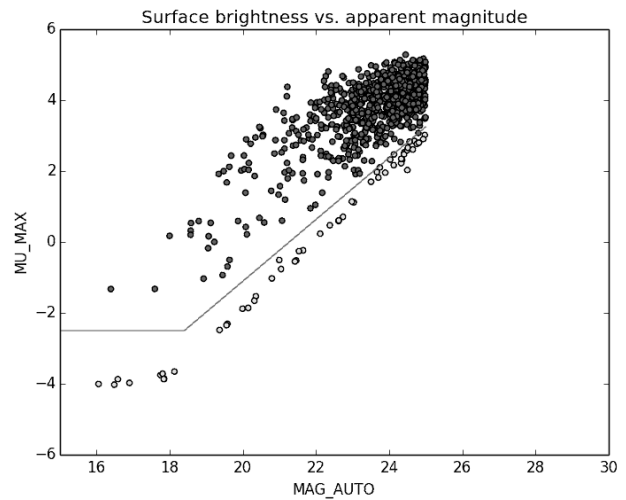


Figure 2. The chosen star-galaxy separation line. The stars are in light gray, the galaxies in dark gray. MU_MAX refers to the surface brightness, roughly how spread out the light from an object is, and MAG_AUTO refers to the total light coming from an object.

of shapes. These galaxies can also be convolved with lower-resolution monochromatic PSFs to simulate how these galaxies would appear in a ground-based telescope. This functionality is part of the RealGalaxy module.

However, the COSMOS basis set is biased due to the fact that the images of the COSMOS galaxies were only taken in the f814w (I-band) filter. Since the PSF is chromatic, the measured shape of the galaxy differs in different wavelengths. Since a galaxy can also have a spectral energy distribution that changes across its surface, the galaxy may even be convoluted differently in different locations across the galaxy.

It is an area of ongoing research how non-negligible this particular effect is on our ability to measure gravitational weak lensing. However, it is important to be able to test our analyses via simulations of multichromatic galaxies.

Our project focuses on providing such a set of multichromatic galaxies from the All-Wavelength Extended Groth Strip International Survey (AEGIS). [4] This will allow us to estimate a realistic galaxy shape distribution for wavelengths from 606 nm to 814 nm. We identify an initial sample of approximately 106 objects from the I- and V-bands of the AEGIS survey images. We then distinguish stars and galaxies based on their surface brightness characteristics. We develop a robust image-processing pipeline to mask the images for defects, ensuring a high-quality sample. We estimate the PSF for each image from the focus position as measured from a least-squares fit of about ten high signal-to-noise stars to simulated stars generated by the TinyTim ray-tracing program. We cut a postage stamp around each galaxy to minimize error from close contaminating objects and cutting off the edges of the galaxy flux distribution. We match galaxies in the two filters by position to get a multichromatic galaxy postage stamp. Finally, we use this information to provide input to a bulge-disk fitting program to generate parametric models of these galaxies. Our final sample contains approximately 30,000 multichromatic galaxies. We provide some initial sample characterization statistics in the results section.

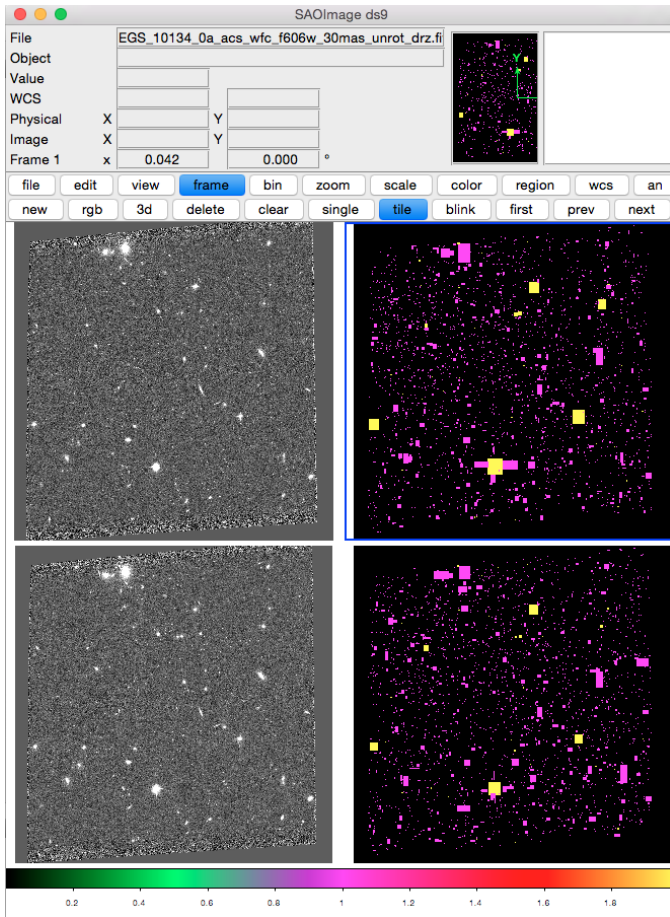


Figure 3. A visual map testing the classifier for both filters. The left hand side contains raw images, and the right hand side is a map showing areas of the raw images containing a star detection in yellow, and an area containing a galaxy detection in pink.

METHODS

Data

The initial data consisted of 63 Hubble Space Telescope Advanced Camera for Surveys Wide Field Channel (HST ACS-WFC) images each for two filters, f606w and f814w. These fields were chosen for their depth and similarity to the COSMOS data, for which galaxy models have already been extracted. The images were arranged in a slightly overlapping 20 by 3 grid entirely covering the Extended Groth Strip, which has an area of 67 arcminutes by 10 arcminutes with a pixel scale of 30 milliarcseconds per pixel width. The images were previously corrected for charge-transfer inefficiency, analyzed to find the magnitude zero-point, and provided with appropriate inverse variance files. [5] More information on AEGIS can be found on the AEGIS website.

Object Selection

SExtractor [6] is software used to automatically detect stars and galaxies in CCD images. However, the automatic detection algorithm used by SExtractor suffers from two problems. Sometimes, it detects fragments of galaxies, such as spiral arms or disk remnants, as whole galaxies. Other times, two overlapping galaxies are mistakenly classified as only one galaxy. Specifically, we

use the hot-cold detection strategy as presented in Leauthaud et. al. (2007) [7] to minimize error from both of these kinds of errors. Roughly, we perform two passes of object detection; with a first pass that detects big bright galaxies, and a second pass that detects smaller galaxies that are not in the vicinity of the large galaxies. The main parameters that we adjust to affect how SExtractor are those that require an object to cover a minimum area on the image in order to be detected, and various other settings related to how bright an object has to be before it is detected. We use the parameters described in the paper for SExtractor configuration, first performing a coarse (cold) detection of all the bright objects in the image, creating a segmentation map with a cushion of 20 pixels around each object. We then perform a fine (hot) detection of all the fainter objects in the image, discarding objects within the highlighted areas of the segmentation map. Combining this set of objects with the cold objects, we merge these two catalogs to create an initial catalog.

Star-Galaxy Separation

It is important to distinguish stars and galaxies in our images, as galaxies are used for weak gravitational lensing shape measurements, while stars are used as important data to calibrate corrections for the varying focus of the camera. Instead of using the built-in neural network classifier, the surface brightness/magnitude plane was separated into two regions, which naturally separate the stars and galaxies due to the fact that galaxies have a much more spread out flux distribution; stars appear more like the light was coming from a laser pointer, while galaxies appear more like the light was coming from a flashlight. The surface brightness/magnitude plane allows us to see roughly how “wide” the light from an object is. The separation is performed visually and is presented in Figure 2. A map generated to test the classification of stars and galaxies is presented in Figure 3.

Quality Filtering of Objects and Stellar Diffraction Spikes

Very bright stars tend to leave “trails” on the CCD that bias our measurements of galaxies that get in the way (Figure 4). This effect is due to diffraction, a well-understood optical phenomenon caused by diffraction off the struts that hold the Hubble Space Telescope’s secondary mirror in place. A correction for this effect would be difficult and unreliable, so we just throw objects in the paths of diffraction spikes away. We simply identify these trails, flag affected areas of the image, and delete objects within the flagged areas. Diffraction is a wavelength-dependent effect, so we define separate automatic masking parameters for each filter. A shape for the mask is predefined (Figure 4). We assume that the length of each mask should be proportional to total flux of the bright stars, and the width is held constant at a relatively conservative estimate of 40 pixels based on visual inspection. We also note that the diffraction mask is rotated by some angle with respect to the image axes. We measure the size of the spike and angle of rotation for 7 bright stars in each filter of various magnitudes, with an upper limit of 15.1, which was the apparent saturation limit as determined by the surface-brightness/magnitude plot from the star-galaxy classification step. Based on these measurements, we determine the best linear fit to the data and obtain a spike size as a function of flux. The r-values for the f606w and f814w filters

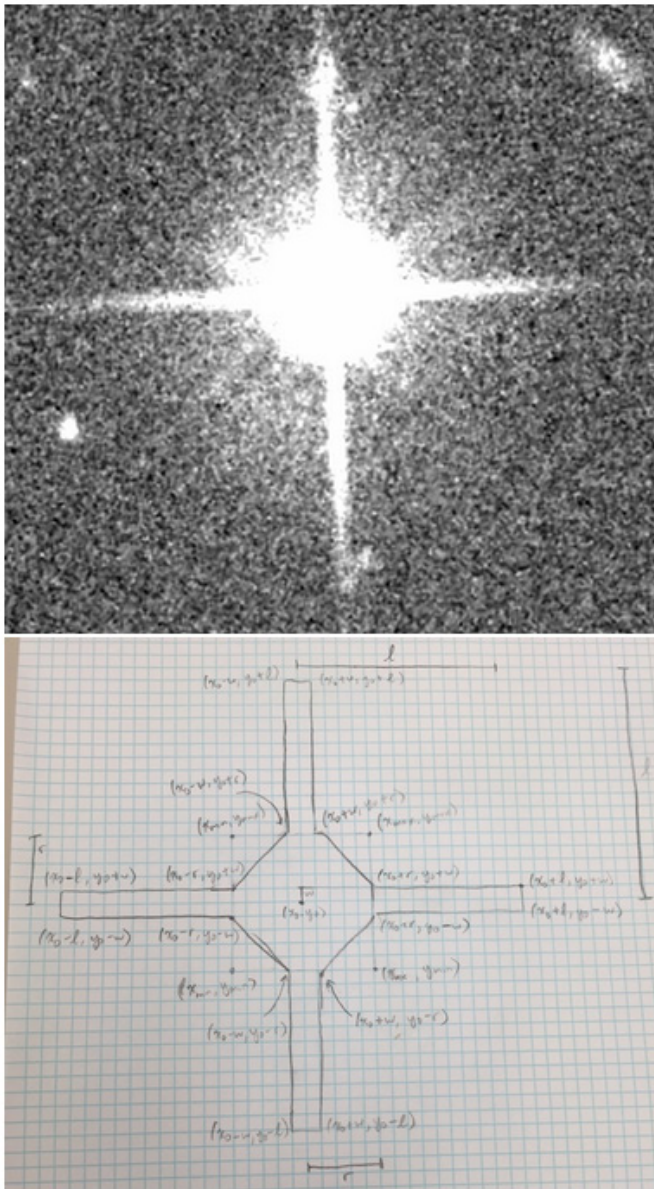


Figure 4. The “spikes” surrounding this star are a result of diffraction, an effect where the struts holding the Hubble Space Telescope in place become visible in the image as spikes. Since these spikes can affect shape measurements of background galaxies, we want to remove stars and galaxies that lie in the path of these spikes from our sample. We place a mask (bottom) over each star and remove objects within the mask from our sample.

respectively are 0.99 and 0.97, indicating a good fit to the data. No relationship was found between the right, left, top, or bottom spike size relative to the other positions. To implement the spike mask, we identify all saturated stars and define a list of vertices of a polygon (mask) for each star. We then iterate over the other objects in each catalog and use the Jordan Curve Theorem (which states a point lies inside a closed curve if and only if, when we draw a line out to infinity from the point, it crosses the curve an odd number of times) to determine whether an object is within a mask by counting the number of “crossings” of the polygon by a line drawn out to infinity from the object. Odd numbers of crossings indicate the object is inside the polygon. We then delete objects within masks. In this step, about 3.2 percent of the total

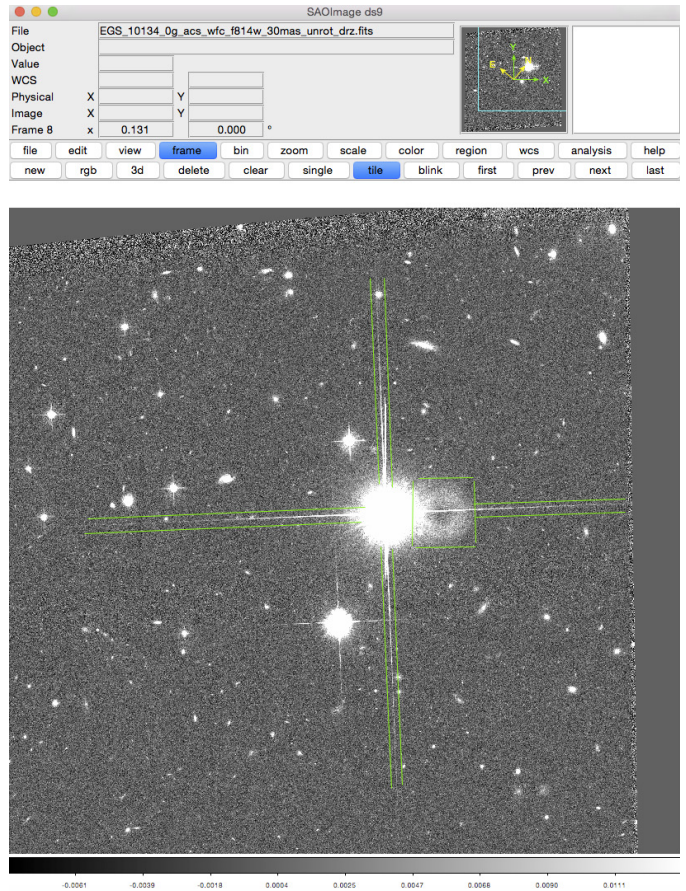


Figure 5. A manual mask of a very large star and the reflection of the Hubble Space Telescope primary mirror. Images were processed by hand and a mask was placed on the large artifacts since the automatic detection possibly detects the wrong centroid. All galaxies within the green lines are removed from the sample.

objects are discarded.

Image Overlap Correction

The individual image tiles in the Extended Groth Strip, the area of the sky imaged in the survey, are slightly overlapping, which means that it is possible to select the same object in two images. The object in one of the two catalogs is deleted if the centroids are less than 0.2 arcseconds apart in the world coordinate system (about 7 pixels).

Manual Masking

The centroid measurement given by SExtractor may be inaccurate for large stars, because the stars saturate and we lose the peak of the flux distribution. In order to flag and mask other areas of the images affected by non-physical defects, a manual mask for large stars and image defects was implemented. Each of our 126 images was visually inspected and boundaries of areas to discard were defined. A sample mask of a large star and mask of the reflection of the Hubble primary mirror are shown (Figure 5). About 0.5 percent of the objects are discarded in this step.

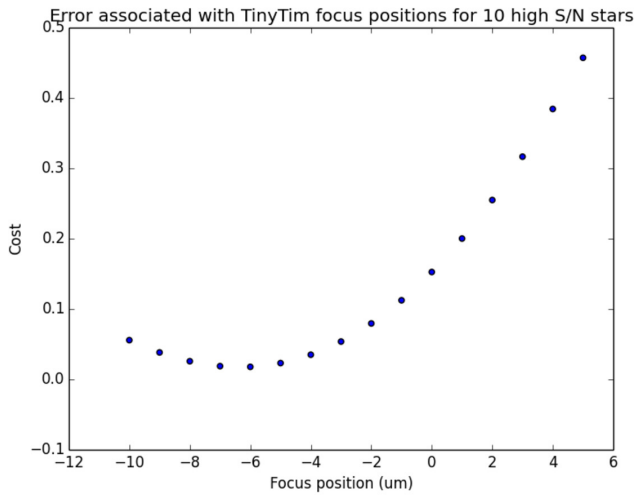


Figure 6. For each focus position, we plotted the cost function, which describes how different the simulated stars at a given focus position are compared to the actual stars in the image. We estimate that the focus position of the image corresponds to the one which minimizes this cost function.

Focus Position and PSF Estimation

In addition to the distortion introduced by gravitational lensing, the light from the galaxies is also modulated by a point-spread-function, as a result of various detector and optical effects. To correct for the effects of the point-spread-function, typically stars (effectively point sources) are used to predict how the light from a point source is affected by the telescope. There are unfortunately not enough stars in our sample to use the stars as effective PSFs; additionally, the stars can be affected by the intrinsic noise of the image. Fortunately the variation of the PSF of the Hubble ACS WFC can be mostly ascribed to a single physical parameter: the focus position, or separation of the primary and secondary mirror, which changes due to the thermal breathing of the telescope. [8] The focus position can be estimated based on a best-fit measurement of about 10 well-measured, high signal to noise stars. Given a focus position, we then use the TinyTim Hubble Scope Telescope simulation software to generate noiseless, simulated PSFs corresponding to the correct focus position, which are ultimately more accurate than using real stars as PSFs.

A number of stars were selected manually out of the 20 highest S/N star detections in each image. If not enough stars were found (i.e., less than 9 stars), which occurred in about 10 percent of the images, an exhaustive search of all the stars in the image was performed. All images contained enough stars to determine the focus position after this second search.

The stars were compared to simulated stars at each focus position. A cost function was defined as the sum of the squared ellipticity component differences between the real and simulated stars, where the ellipticity was defined in terms of the second order moments. The goal is to find the set of simulated stars corresponding to a particular focus position that have the least “error” when compared to the actual stars in the image. One component measured ellipticity along the horizontal and vertical axes, and the other measured ellipticity along the diagonals. The focus was chosen as the one corresponding to the field with the lowest cost function. A sample graph of the cost function is

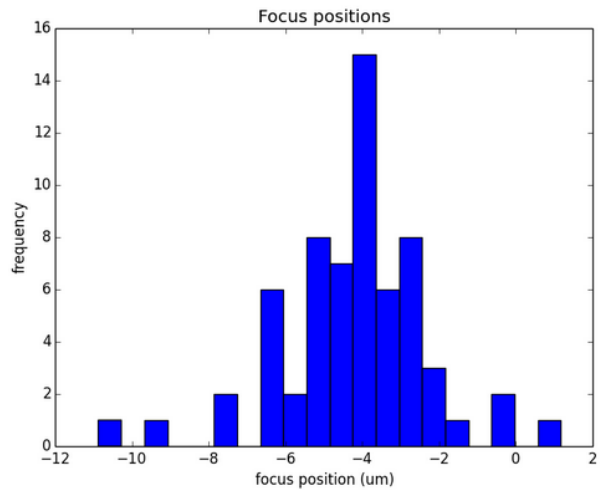
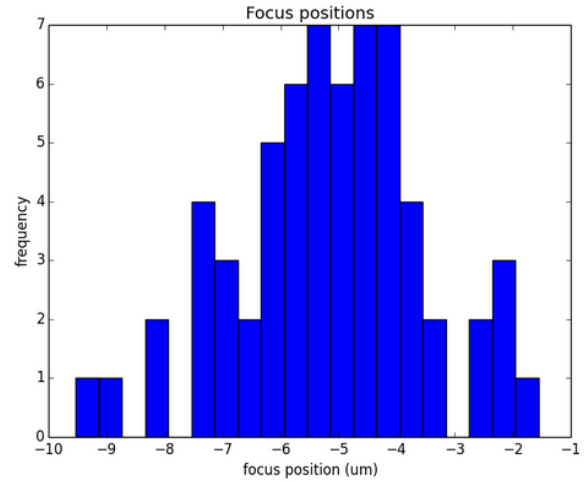


Figure 7. Histograms of focus positions for the f606w filter (left) and the f814w filter (right). This is the number of images we measured at each focus position.

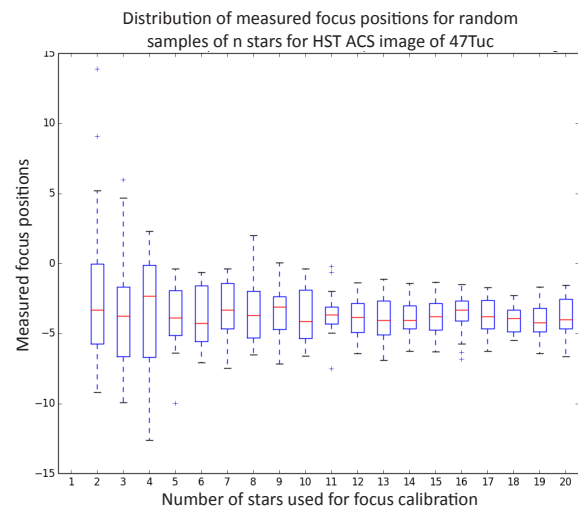


Figure 8. Distributions of focus positions from 20 samples of varying sizes. We notice the additional accuracy of measuring the focus position with more stars is no longer statistically significant after about 10 stars. This result confirms previous work that 10 stars is sufficient to accurately measure the focus position.

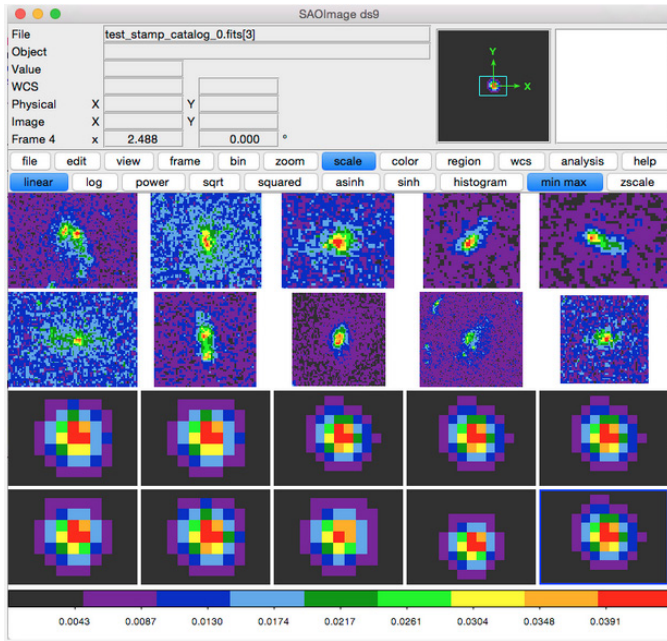


Figure 9. The top 10 images are samples of the postage-stamp cutouts of high signal to noise galaxies in our images. The bottom 10 images are the corresponding PSFs.

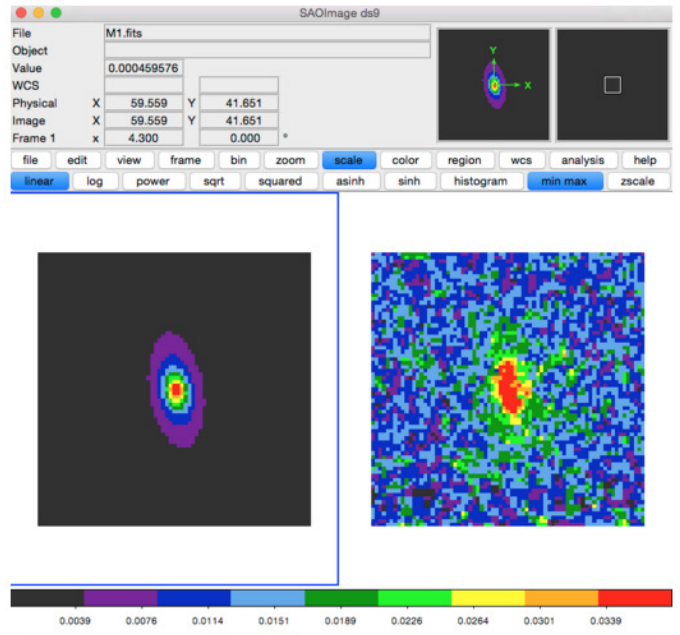


Figure 10. A sample parametric galaxy model. On the right is a noisy raw galaxy image. We used Lackner and Gunn's bulge-disk fitting software to create a model of the galaxy on the left which can be easily parameterized as a simple bulge and disk.

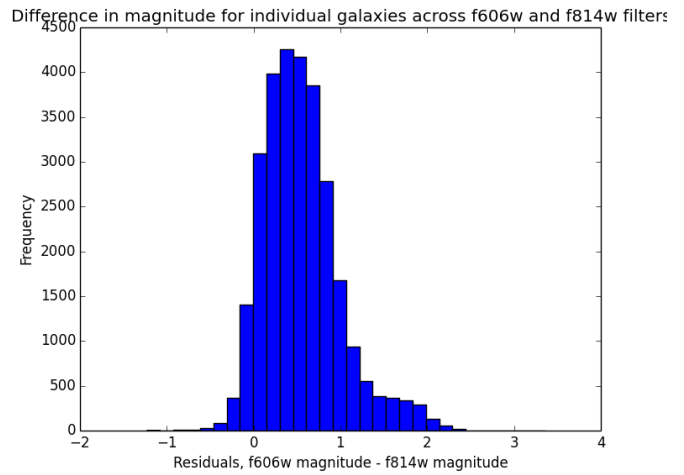
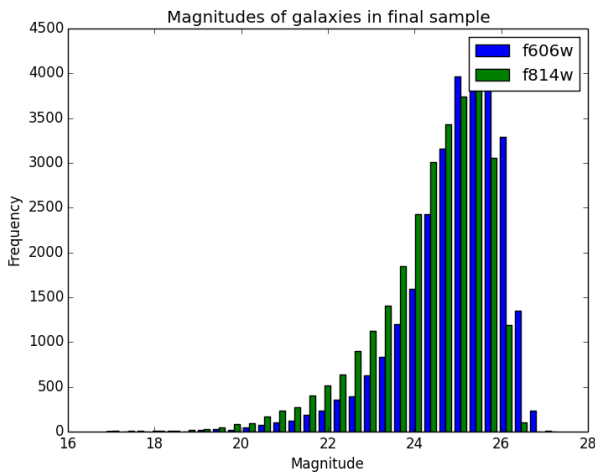


Figure 11. Apparent magnitudes of the galaxies in the sample and the differences of the magnitude across f606w and f814w filters. Figure 11a shows a histogram of the magnitudes, or scaled negative log-brightness, of the galaxies in the sample, by filter. By convention, lower magnitudes represent brighter galaxies. Figure 11b shows a histogram of the differences between the magnitude in one filter versus the magnitude in the other filter.

plotted in Figure 6. Histograms of the measured focus positions are displayed in Figure 7.

Verification of Focus Position Measurement Precision

In order to make sure that the determination of the focus positions is reproducible given different sets of input stars, the focus position pipeline was run on an image of the globular cluster 47Tuc in the f606w filter with the same camera as was used to take our science data. (We did not verify the measurement accuracy for the f814w filter, which was already verified. [8] However, this should be re-verified in future work as our methodology differs

slightly from the RRG method used in the original paper). First, we randomly selected 20 samples of 20 well-measured, high S/N stars in the f606w filter. We then selected a subsample of 2-20 stars from each sample and calculate the focus position of the subsample. We then look at the variation of the focus position as a function of the number of stars in each subsample to determine a minimum number of stars we need to adequately determine the focus position. The results are shown in Figure 8.

We see the variation is no more than ~ 1 micron on average, and ~ 4 microns maximum for different samples of stars when the number of stars is at least 10. This is slightly greater than the uncertainty previously measured [8]. However, we conclude that the uncertainty does not appear to become smaller as the number

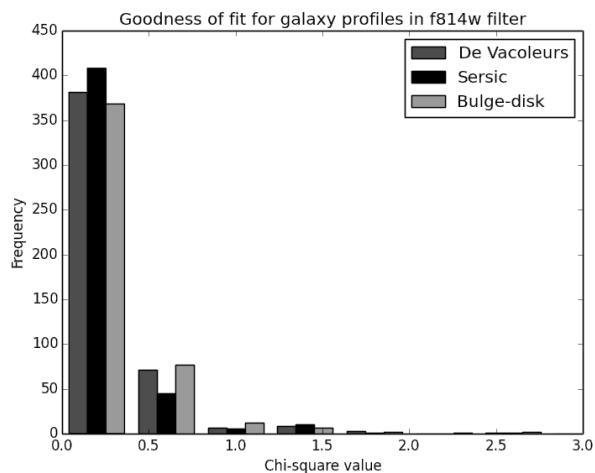
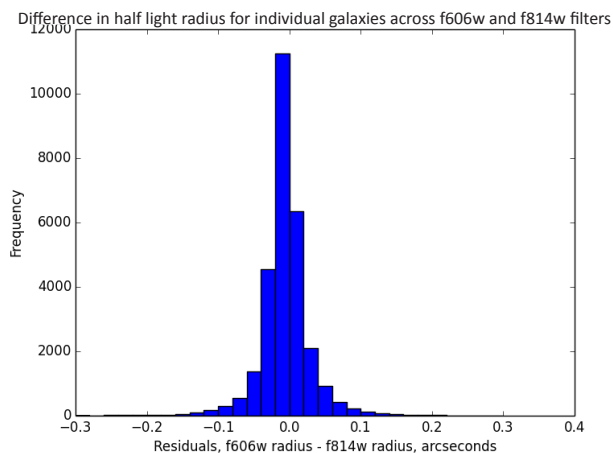
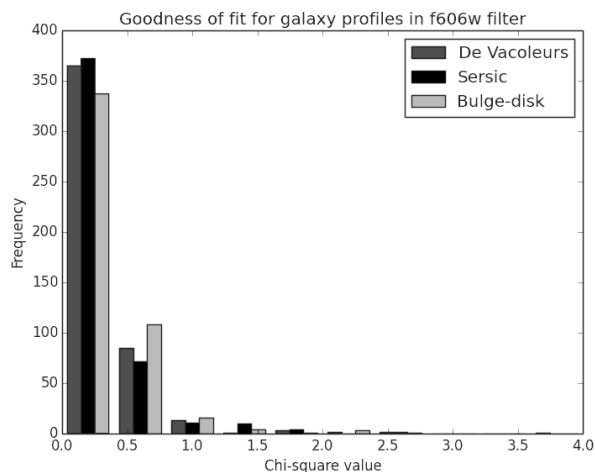
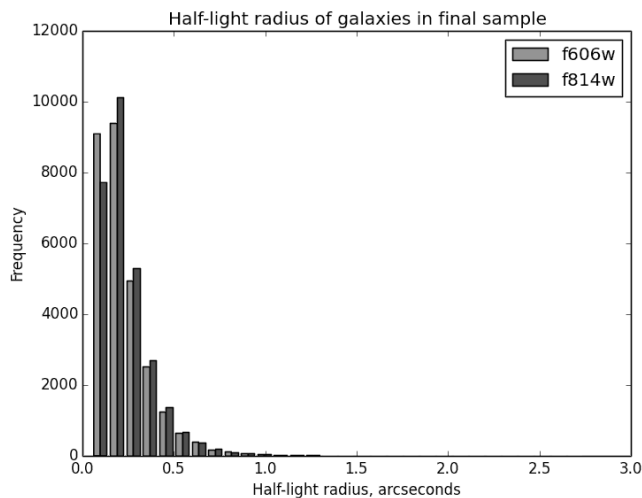


Figure 12. Half-light radius of the galaxies in the sample and the differences of the half-light radius across f606w and f814w filters. The top figure shows a histogram of the widths of the galaxies in arcseconds, measured by the width at the half-maximum of each galaxy flux distribution. The bottom figure shows a histogram of the difference between the width of a galaxy in one filter versus the other.

Figure 13. A histogram of the goodness-of-fit for each of the galaxy models we fit to the data. We fit three different models to each galaxy, and the Chi-square value represents how well the data fits the model. We infer from these plots that no one model outperformed the other two across the entire sample, as expected.

of stars used increases past 10. We therefore conclude that 10 stars is sufficient to reasonably determine the focus position.

Postage Stamp Selection

The remaining galaxies in the catalog were saved as postage stamps, where the postage stamp size was defined as four times the half-light radius, along with their TinyTim PSF. The stamp size was chosen by manually inspecting a test set of galaxies to find the optimal stamp size that does not cut off the galaxy distribution, but also for the most part avoids contaminants. The galaxies were required to have a signal-to-noise ratio of at least 20, consistent with the S/N from the GREAT3 challenge [9], and to be contained within the image. The PSF postage stamp was also obtained from the TinyTim software based on the focus position of the image. SExtractor was run again on each of the postage stamps with nearly default parameters to generate a mask containing the pixels only belonging to the central object for the parametric fitting code. Sample postage stamps and PSFs are

shown in Figure 9.

Obtaining Parametric Models

We began to obtain parametric models as described in the GREAT3 Challenge paper by using the bulge-disk fitting code from Lackner and Gunn (2012) [10]. Models were obtained for the first 472 galaxies and input was prepared for the remainder.

Three models were used to fit the galaxies. The first two were both Sérsic profiles of the form: $\ln(I(R)) = \ln(I_0) - kR^{1/n}$ with I as the intensity as a function of radius and I_0 as the intensity at the center of the profile. For the first profile, n , the Sérsic index, was left as a free parameter, and for the second profile, the de Vaucouleurs profile, n was fixed to 4. Finally, the third profile was a bulge-disk decomposition where n was fixed to 4 for the bulge and fixed to 1 for the disk. An example output of the best-fit model can be seen in Figure 10

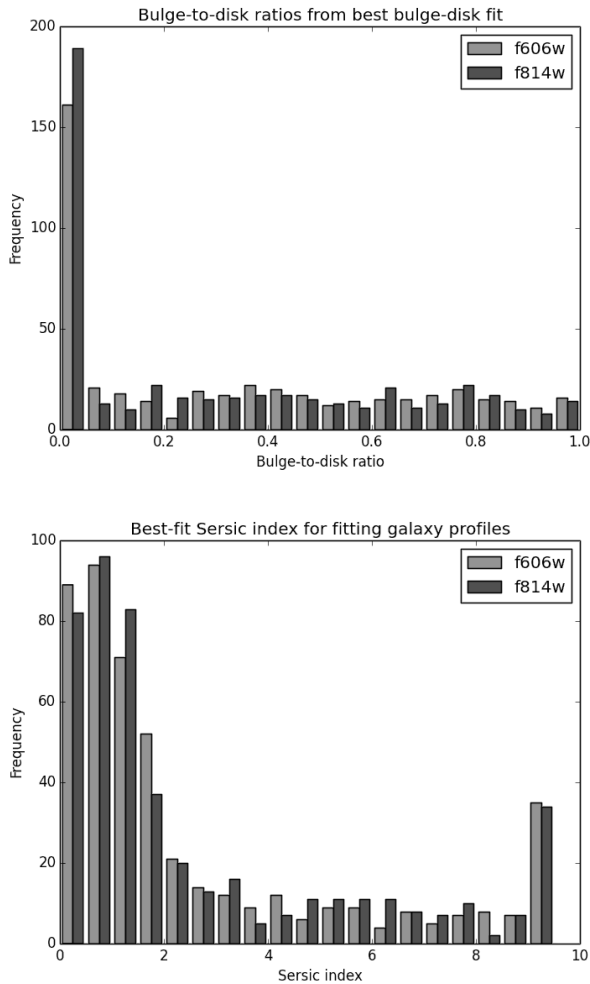


Figure 14. Figure 14a shows a histogram of the ratio of bulge-to-disk for the B-D model. A ratio of 0 indicates that the galaxy was best fit by a simple disk, rather than the two-part bulge-disk. Figure 14b shows a histogram of the best-fit Sérsic index; loosely how quickly the galaxy flux distribution drops off away from the center of the galaxy.

RESULTS

General Sample Statics

We characterized the final sample in terms of magnitude and half-light radius. Plots of the magnitude and difference in magnitudes between filters are given in Figure 11. Plots of the half-light radius and difference in half-light radius between filters are given in Figure 12. The sample in total contains approximately 30,000 multichromatic galaxies.

Statistics on Initial Models

We also characterized the first 472 parametric models. Figure 13 shows the goodness of fit for all three profiles in both filters. We conclude no single fit is best suited to fitting these galaxies and all three models fit the galaxies about equally well. Figure 14 shows the bulge-disk ratio for the bulge-disk fit in both filters and the best fit Sérsic index. In future work, we plan to use this data to

select which model best fits the galaxy for implementation into GalSim.

CONCLUSION

Future work includes generating the full set of parametric models for the dataset, matching our data to catalogs of known photometric redshifts, and using the models as inputs for a ChromaticRealGalaxy module. We were limited by our computing capacity due to time constraints, but the postage stamps are hosted online and are ready to be used as input to the modeling program with no additional modifications needed. Testing and verification of the features of ChromaticRealGalaxy will be needed as well.

Beyond the current implementation of RealGalaxy, we would like to be able to simulate images from WFIRST. This will require expanding the range of wavelengths out to the infrared limit of 2 microns and deepening the sample by about two orders of magnitude. These simulations will lead to a much better understanding of how well we will be able to probe dark matter and dark energy with WFIRST.

CODE REPOSITORY AND DOCUMENTATION

All code is available for public release and documented on GitHub at <https://github.com/bradleyemi/ChromaticGalaxies>. Additionally, the original images, the SExtractor catalogs, the postage stamps, and the input files for the bulge-disk fitting code can be found at that link.

ACKNOWLEDGEMENTS

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, and was sponsored by SURF and the National Aeronautics and Space Administration. Thanks to Jason Rhodes and Andrés Plazas for excellent mentorship through the summer. Additionally, Rachel Mandelbaum was extremely helpful in guiding the process and providing useful advice relating to the process of implementing RealGalaxy with COSMOS data. Finally, thanks to Patricia Burchat and Joshua Meyers from Stanford University for helping with GalSim integration. This research would not be possible without the California Institute of Technology Student-Faculty Programs office for organizing the SURF program and the Jet Propulsion Laboratory for various intern support throughout the summer.

REFERENCES

- Bernstein, G. and Jarvis, M. *Shapes and Shears, Stars and Smears, Optimal Measurements for Weak Lensing*, 2001
- Rowe, B. et. al. *GalSim: The Modular Galaxy Image Simulation Toolkit*, 2014
- The Cosmological Evolution Survey, <http://cosmos.astro.caltech.edu/>
- The All-Wavelength Extended Groth Strip International Survey, <http://aegis.ucolick.org/>
- Hubble Space Telescope ACS Data Analysis, <http://www.stsci.edu/hst/acs/analysis>
- Bertin, E. and Arnouts, S. *SExtractor: Software for Source Extraction*, 1996
- Leauthaud, A. *Weak Gravitational Lensing with COSMOS: Galaxy Selection and Shape Measurements*, 2007
- Rhodes, J. et. al. *The Stability of the Point Spread Function of the Advanced Camera for Surveys on the Hubble Space Telescope and Implications for Weak Gravitational Lensing*, 2007

9. Mandelbaum, R. et. al., *The Third Gravitational Lensing Accuracy Testing (GREAT3) Challenge Handbook*, 2014
10. Lackner, C and Gunn, JE, *Astrophysically Motivated Bulge-Disk Compositions of SDSS Galaxies*, 2012



BRADLEY EMI

Bradley Emi is a second year undergraduate at Stanford University studying physics and computer science. He is interested in techniques from artificial intelligence and their applications in physics and engineering. He is also fascinated by new developments on the theoretical side of physics and their implications for the philosophy of science. This summer, Bradley will be in Pittsburgh, Pennsylvania, working with the Perception team at the Uber Advanced Technologies Center to develop software for self-driving cars. After graduating from Stanford, he plans to attend graduate school and obtain a Ph.D., continuing to be involved with scientific research. Aside from physics, in his free time, Bradley loves cooking, reading philosophy, and playing volleyball.

The combined effects of ocean acidification and sea surface temperature rise on larval sea hare (*Aplysia dactylomela*) development

Anja Malawi Brandon

Stanford University

The increase in partial pressure of CO₂ in the atmosphere, largely from anthropogenic sources, is affecting marine calcifying organisms. This is caused by a reduced availability of carbonate due to ocean acidification. Simultaneously, the sea surface temperature (SST) is rising due to the greenhouse effect associated with increased CO₂ in the atmosphere. This study investigates the effects of decreased pH and increased temperature on the larval development of *Aplysia dactylomela*, an ecologically important marine gastropod that produces a calcium carbonate shell as an embryo. An ocean acidification and simulated sea surface temperature system was used to treat *A. dactylomela* egg masses in 4 treatments based on the Intergovernmental Panel on Climate Change (IPCC) predictions for 2100: control, heated (+3°C), acid (pH 7.6), and acid plus heat (+3°C, pH 7.6). This study found a statistically significant decline in rates of survival and amount of movement for all treatments relative to the control, with the acid plus heat treatment having the lowest rates of survival and movement. These results suggest that over the next century there may be further negative consequences for the larval development and thus species success and survival of *A. dactylomela* as well as other marine calcifying organisms important to coral reef ecosystems.

Introduction

Anthropogenic carbon dioxide (CO₂) emissions have increased to unprecedented levels over the last several decades and are predicted to continue rising [1]. The atmospheric concentration of CO₂ has risen over 100 ppm from pre-industrial levels of 280 ppm to nearly 400 ppm in 2014 [1]. Much of this steady increase in CO₂ in the Earth's atmosphere has been absorbed by the planet's largest CO₂-trap: the ocean. The additional atmospheric CO₂ has led to a greater CO₂ partial pressure (pCO₂) in the oceans, which has reduced oceanic pH by 0.1 units since pre-industrial times [2]. The steadily decreasing oceanic pH, known as ocean acidification, has a detrimental effect throughout the oceans and especially on sensitive coral reef environments [3].

Ocean acidification reduces the concentration of calcium carbonate (CaCO₃) in the ocean and thus limits the ability of coral reefs and other calcifying organisms to grow [2]. The current global oceanic average pH of approximately 8.1 allows for the super saturation of CaCO₃ needed by reef building corals and other calcifying organisms; however, the Intergovernmental Panel on Climate Change (IPCC) predicts that oceanic pH will continue to decrease by as much as 0.3-0.4 units within this century [1]. These changes will affect many calcifying and CaCO₃ skeleton-building organisms, both as mature invertebrates and during embryonic development. The decrease in the concentration of carbonate ion (CO₃²⁻) in the ocean will weaken the ability of marine calcifying

organisms to build shells or exoskeletons out of CaCO₃. This under-saturation can lead to pitting, or the degradation and deformation of the exoskeletons or CaCO₃ shells [2]. Species that are dependent on embryonic shell development are susceptible to population bottlenecks due to decreased mortality as larvae. This is because of their increased vulnerability to acidic conditions [4]. A bottleneck in early life stages may cascade and impact the species overall success and continued persistence [4]. While ocean acidification has been well established as a problem for calcifying corals, it is also known to affect the growth and development of other marine calcifying organisms like mollusks and other benthic invertebrates [3,5].

The long-term increase in atmospheric CO₂ and other greenhouse gases has also been shown to impact global temperatures, which in turn affects the sea-surface temperature of the ocean [6]. The IPCC predicts a sea surface temperature (SST) increase of 2.4°C to 6.4°C by 2099 assuming greenhouse gas emissions continue to increase as they have over the last several decades [1]. An increase in temperature can lead to an increase in rates of metabolic reactions and the acceleration of embryonic development until a species-specific temperature threshold is reached [4]. Increases in temperature above this threshold are detrimental and can result in a dramatic increase in mortality [4]. This increase in temperature is especially concerning to tropical organisms that are adapted to a relatively small temperature range.

One such organism potentially affected by the increase in SST and ocean acidification is the spotted sea hare, *Aplysia*

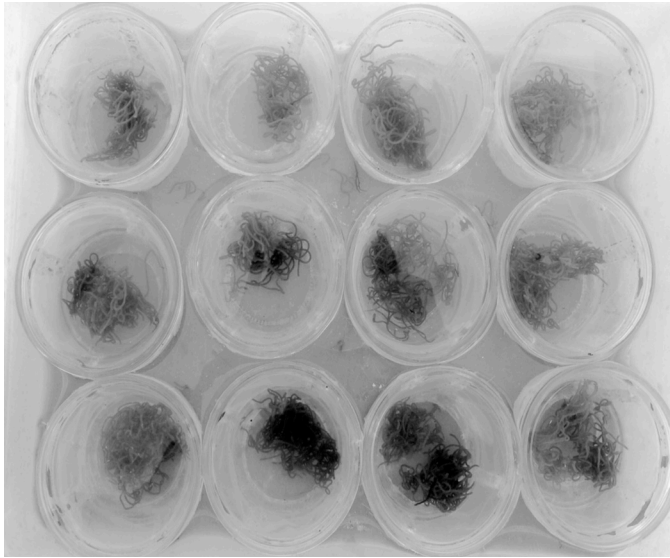


Figure 1. The three *A. dactylomela* egg masses were divided randomly and evenly and placed into 12 semi-permeable mesh containers prior to being placed in treatment tanks

dactylomela, a marine gastropod found on sandy substrates in coral reefs. Sea hares, like other opisthobranch gastropods, are simultaneous hermaphrodites that reproduce via internal fertilization and have the capability of sperm storage [7]. Sea hare eggs are laid in discreet capsules within a gelatinous string that is usually attached to algae or under a hard surface at a rate of 5.9 cm min^{-1} , which is approximately $41,000 \text{ eggs min}^{-1}$ [8]. Because survival to hatching and through the larval planktonic stage to sexual maturity is incredibly low, sea hares lay thousands of eggs to ensure that enough sea hares will survive to reproduce [8]. The sea hares develop a CaCO_3 shell during their embryonic development, which later becomes covered in tissue as they mature [9].

The spotted sea hare serves as a useful model animal because of its ecological importance and the available previous research on its embryonic development. The spotted sea hare plays an essential role in maintaining coral reef health by helping to control the macroalgal population. Without the sea hares and other key herbivores, the macroalgal population could overrun coral reefs and there could be a transition from reef dominated to algal dominated systems. Previous studies have been conducted regarding *A. dactylomela* embryonic development under the predicted changing oceanic conditions. One such study explored the effects of both lower pH and increased SST on the embryonic development over their incubation time prior to hatching [10]. However, the research did not quantitatively investigate survival to hatching rates or the condition of the hatched larvae. Understanding the survival rates of the larvae and their health upon hatching is critical to understanding how these sea hares, and many other marine calcifying organisms, will survive in the changing oceans of the future. The goal of this work was to quantify the effects of ocean acidification and sea surface temperature rise would have on the health and viability of *A. dactylomela* during embryonic development.



Figure 2. The ocean acidification tank set up on an outdoor deck at Heron Island Research Station. The large pH and heat-controlled sumps are located under the table.

METHODS

Study Site

This study was conducted on Heron Island ($23^\circ 26' 31.2'' \text{ S}$, $151^\circ 54' 50.4'' \text{ E}$), 80km east of Gladstone in Queensland, Australia. The island is located on the leeward edge of a 27 km² coral reef platform and has a deep lagoon with patches of coral. The surrounding reef is very high in diversity, supporting 72% of all the Great Barrier Reef coral species and a high proportion of endemic species [11].

Egg Mass Collection

Five sea hares (*Aplysia dactylomela*) were collected from the reef flat off the southern coast of Heron Island and kept in a tank with unfiltered water pumped in directly from the reef. They were fed the macroalgae *Laurencia* and allowed to copulate for several days. The sea hares were collected on 11 October 2014, and three egg masses were first sighted on 15 October 2014. The masses were randomly and evenly split into small semipermeable mesh containers and then allocated randomly to a treatment tank resulting in one mesh container per treatment tank (Figure 1). All of the eggs were placed in treatment tanks within 12 hours of being laid.

Experimental Design

Four conditions were used in this experiment: control, heated, acidic, and acidic plus heat. The control treatments used a flow-through system directly connected to the reef to keep the egg masses at ambient temperature and pH. The ambient temperature ranged from 22°C to 25°C , and the pH was approximately 8.1. The heated treatments simulated an increase in SST and were maintained at $+3^\circ\text{C}$ above ambient temperature. The temperatures in these tanks ranged from 25°C to 28°C . The acidic treatments were maintained at pH 7.6, the IPCC estimated oceanic pH for 2100, with ambient temperature. The tanks under acidic plus heated conditions were maintained at pH 7.6 as well as $+3^\circ\text{C}$. There were three replicates of each treatment leading to a total of

12 tanks. (Figure 2).

Ocean Acidification System

The ocean acidification system used for this experiment was designed to model the increased oceanic pH values based on IPCC predicted concentration of CO₂ in the atmosphere (ppm). Unfiltered seawater was pumped directly off the reef flat into large sumps then into the treatment tanks, which allowed for a constant flow of fresh water to the embryos and for all non-experimental parameters (such as dissolved oxygen) to remain the same as within the natural environment. The tanks in this system used as controls received the water from the reef flat with no alteration. The water for the acidic treated tanks was pumped into a large 200 L sump where an automated CO₂ injection system was used to elevate the dissolved CO₂ concentrations to reach a pH of 7.6 by injecting CO₂ into a diffuser and rapidly dissolving CO₂ into the seawater. A pH control unit (Aquatronic, AEB Technologies, Italy) was used to monitor and maintain the desired pH level in the sump. This control unit was connected to a pH probe in the sump and controlled an electronic solenoid connected to a cylinder of CO₂ that would be opened if the pH rose above 7.65 and closed when the pH returned to 7.6. The water from the sump was then pumped into the treatment tanks for both the acidic and acidic plus heat treatments.

Aquarium heaters were used to heat another large 200 L sump of unfiltered seawater from the reef flat to +3°C. This water was then pumped into the heated treatment tanks. The tanks that were both acidic and heated received the acidic water from the large pH controlled sump and had individual aquarium heaters within the tank to increase the temperature 3°C. Thermometers in the tanks allowed for temperature monitoring to ensure the temperature stayed within the desired range.

Image Analysis

A small segment (~2 cm) of each egg mass and larvae collected from within the mesh containers were viewed under an OPTIKA B-600 Biological Microscope with an attached camera. Images

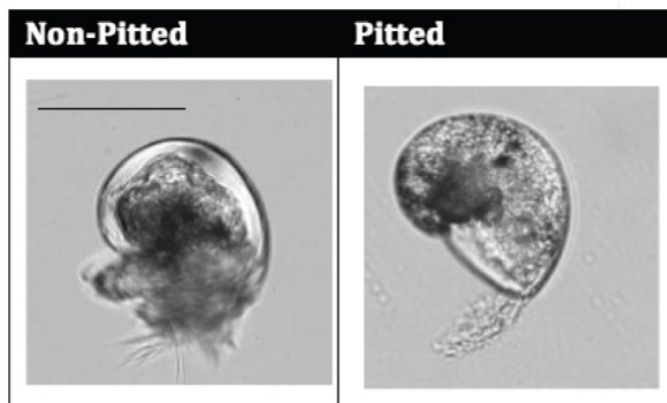


Figure 3. Images of *A. dactylomela* larvae taken under a compound microscope at 20 times magnification. The non-pitted sample with a visibly smooth shell was taken from the control treatment while the pitted image with noticeable deformities in the shell was taken from the acidic plus heat treatment. Both were taken on the day larvae hatched. Scale bar = 100µm.

Scale	Percentage
1	0-20%
2	20-40%
3	40-60%
4	60-80%
5	80-100%

Table 1. The five-point scale used to quantify survival and movement within the sea hare larvae.

were taken through the connected computer. The images were analyzed using ImageJ Software (version 1.48) to determine the shell length. Lengths were found in pixels then converted to millimeters using a calibrated scale from the microscope software (QC Pro), which was accomplished by taking pictures of a 1 mm scale under the same microscope. The longest length of the shell, which is visible when the organism is on its side, was used as the standard for shell length measurements. At least 3 images were taken from each tank and used to get 5 shell length measurements from each tank as a representative sample.

Data Analysis

The percent of larvae with pitted shells (Figure 3), survival rates, and degree of movement were quantified through two one-minute long observations under a microscope. Survival and movement rates were assigned a value using a five-point scale to most accurately quantify the entire egg mass from the representative sample taken (Table 1). Movement rate was based on the percent of larvae seen moving as well as the frequency and duration of movement observed. These observations were then fit to the five-point scale seen in Table 1.

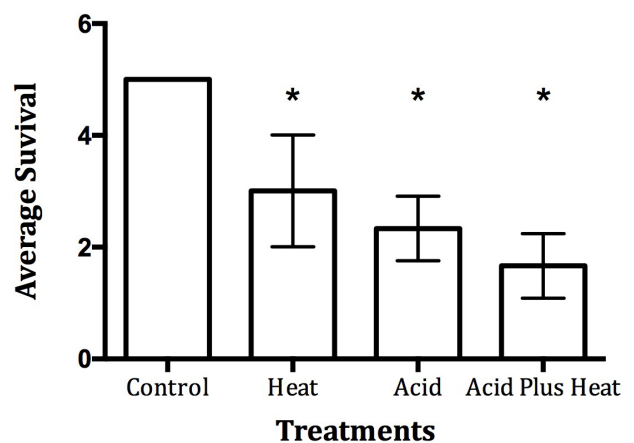


Figure 4. Average larvae survival to hatching with standard error as observed under a microscope and categorized on a five-point scale by treatment. Significant difference ($p < .05$, ANOVA) are indicated by a (*).

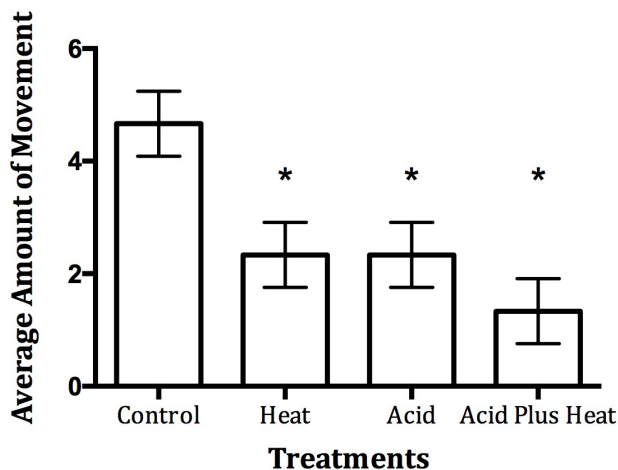


Figure 5. Average amount of movement observed under a microscope with standard error on the day the larvae hatched as categorized on a five-point scale by treatment. Significant difference ($p < .05$, ANOVA) are indicated by a (*).

A one-way analysis of variance (ANOVA) was used to analyze the means and variance in the data collected for survival and movement rates as well as shell-length to determine statistical significance. Due to the extreme heterogeneity of the treatments, an ANOVA could not be performed for the shell pitting data. Instead, a Cochran's Q Test was performed instead.

RESULTS

General Observations

There was a visible reduction in the overall health of the egg masses from the control sample to the samples treated in acid (pH 7.6). The gelatinous strings in the egg masses changed color, degraded, and began falling apart even before the larvae hatched. These visible signs of degradation were present to a lesser extent in the egg masses treated with heat. When the egg masses hatched, there was observable movement of larvae in the mesh containers in the control and heat-treated samples, while in the acid and acid plus heat-treated samples, there were visible numbers of dead larvae floating on the surface. The heat-treated egg masses, including the acid plus heat-treated egg masses, hatched first. The acid and control treatments all hatched the following day. No statistical analysis was done on the difference in condition between the day the larvae hatched and the following day because no second day observations could be taken for the control samples due to the limited time available at the research station. However, it was clear that the average percent of surviving larvae and the amount of visible movement decreased dramatically, particularly in the acid plus heat-treated larvae where the average survival dropped from category 3 (40-60% survival) to category 1 (0 – 20% survival).

Survival Rates

Significantly more larvae survived to hatching in the control treatments than the experimental treatments (Figure 4, ANOVA p -value = 0.0012). The differences between the control treatments and all the other treatments were statistically significant. There was no significant difference in the amount of survival between

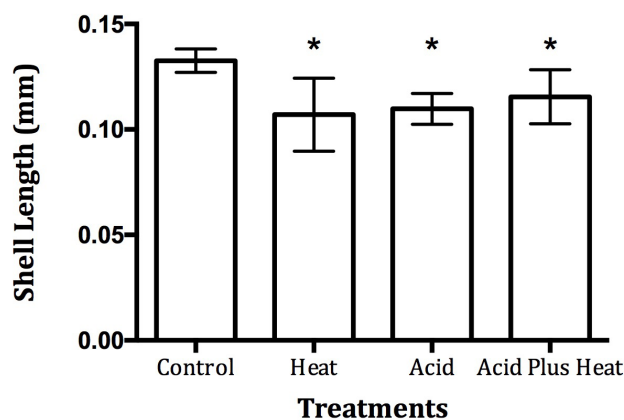


Figure 6. Average shell length measured along the longest length of the shell in mm by treatment with standard error. Shell length was quantified using ImageJ analysis on images taken from under a microscope on the day larvae hatched. Significant difference ($p < .05$, ANOVA) are indicated by a (*).

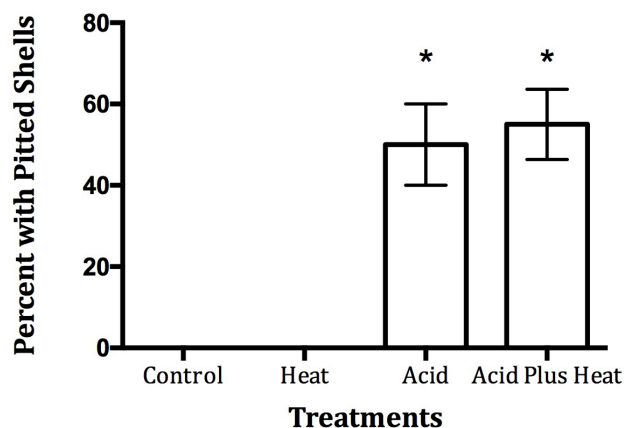


Figure 7. Average percent of larvae with pitted shells as observed under a microscope by treatment, with standard error (see Figure 3 for examples of pitted versus non-pitted shells). Significant difference ($p < .05$, Cochran's Q Test) are indicated by a (*).

acid and the acid plus heat treatments.

Amount of Movement

The amount of movement visible under a microscope differed significantly between the control treatment and all other treatments (Figure 5, ANOVA p -value = 0.0002). The amount of movement visible under the microscope corresponded to the amount of movement qualitatively observed within each of the mesh containers after hatching, with the highest percentage of movement seen in the control treatment. The larvae in the heat treatment had a higher rate of movement than the larvae in the acid plus heat treatment; however, this difference was not statistically significant.

Shell Length

Shell length was significantly longer in larvae that survived to hatching in the control treatment than all other treatments (Figure 6, ANOVA $p = 0.001$). The largest difference in shell length was observed between the control and heat treatments.

Percent of Larvae with Pitted Shells

Only two of the treatments (acid and acid plus heat) had larvae that showed signs of pitted shells. A Cochran's Q Test found a significant difference between the treatments with pH 7.6 and the treatments at ambient pH (p -value = 0.0143). There was no significant difference between the larvae in the acid treatment versus those in acid plus heat treatment.

DISCUSSION

The results of this study suggest that the increased SST and decreased pH in the oceans of the future will have serious negative impacts on the health and survival of the gastropod *A. dactylomela* and other marine calcifying organisms. Not all control treated egg masses survived to hatching, which is expected and normal for egg masses [8]. However, the control treated egg masses had significantly higher survival rates to hatching than all the other treated egg masses. The control egg masses looked healthier and showed fewer signs of discoloration and deterioration. The acid plus heat treatment had the lowest survival rates, which is expected as those larvae were subjected to the most stressful conditions. However, there was no significant difference in survival rates from the acid only treatment compared to the control treated egg masses.

These findings suggest that fewer larvae may survive to hatching as the ocean warms and acidifies from climate change. A higher larval mortality could ultimately cause the population to decline as fewer organisms survive to sexual maturity and particularly poor survival rates lead to population bottlenecks [4]. Other calcifying larvae such as other gastropods or arthropods that are dependent on high CaCO_3 saturation levels to build and maintain shells or exoskeletons necessary for survival could face similar challenges in the future.

The *A. dactylomela* larvae that did survive exhibited reduced movement in all non-control treatments. The amount of movement was studied as an indicator of general health and activity level of the larvae, with higher amounts of movement potentially indicating healthier larvae. The hypothesized existence of an organism-specific threshold temperature under which metabolism increases and beyond which activity rapidly declines may explain the difference between the results of this study and previous research, which found that heat-treated embryos showed an increased amounts of movement [4, 10]. Though heat-treated egg masses hatched first, which may suggest that development occurs faster in heated conditions, there was reduced movement in all experimentally treated larvae. This reveals that even the egg masses that survive to hatching may not be healthy enough to survive the planktonic larval phase. This is supported by the observed reduction in survival and amount of movement one day after hatching in the heat-treated egg masses. The decreased swimming mobility and overall fitness of the larvae could lead to reduced larval survival rates to sexual maturity.

Although natural variation in shell length is expected, the significant difference in shell length between the experimental and

control larvae suggests that larvae in the experimental treatments did not develop as fully as those in the control treatment. While the larvae in the heat treatment developed faster, as demonstrated by their earlier hatching time, they had the smallest average shell lengths of any treatment. This suggests that while higher temperatures may accelerate development rate, they may not develop to be as large. The reduced size of the shells in the acid and acid plus heat larvae suggests that the environmental stress of ocean acidification could also lead to smaller larvae due to decreases in development. Because smaller larvae may be more susceptible to predation and less able to compete for food, the smaller shells may also contribute to higher larval mortality rates.

Pits and other deformations in the CaCO_3 shells of larvae could affect sea hare survival to sexual maturity and the overall species population. The CaCO_3 shells developed as embryos are crucial for protection from predators and harmful UV, and thus damaging these shells could make the larvae more vulnerable to environmental factors [12]. The deformations were only present in egg masses in pH 7.6 treatments (acid and acid plus heat treatments), which suggests that ocean acidification may affect proper shell development and lead to deformities that may compromise the larvae's ability to survive. Studies of other calcifying larvae such as sand dollars and sea urchins similarly found a reduced calcification rate and increased deformities in larvae treated under acidic conditions [12]. The outer mantle of the sea hare protects its shell from further pitting and deformations from acidic conditions when the sea hares become mature adults. However, many other marine calcifying organisms, such as reef building corals, are affected by shell pitting throughout their life, and this study suggests that at the highly reduced pH expected in the ocean within the century, deformities and calcification difficulties would become increasingly present. This could result in population declines of many calcifying organisms, which could interrupt the food web on coral reefs as well as within other ecosystems.

Not only are oceans are becoming more acidic, but they are also warming. This adds a secondary stress to marine organisms [1]. However, there was no significant difference observed between the acid and the acid plus heat treatment in this study. This could be due to limited numbers of measurements and replicates in this study. Qualitatively, the egg masses in the acid plus heat treatments appeared more discolored and deteriorated than the egg masses in the acid treatment alone, which supports the observed lower survival rate in the acid plus heat treatment. There was no significant difference observed in any metric between the acid treatment and the heat treatment, with the exception of shell pitting, which suggests that both stressors are similarly detrimental. These results demonstrate a need for further investigation on the combined effects of acidification and temperature rise on the development of sea hare larvae and other marine organisms.

This study suggests that the increasingly acid and warm waters due to climate change may threaten the future survival of *A. dactylomela*. This has broader implications for coral reef ecosystems overall as sea hares play a critical role in maintaining the balance between macroalgae populations and corals. As coral reefs struggle to calcify in waters under-saturated with carbonate due to ocean acidification, macroalgae will have greater opportunity to grow on the damaged reefs. A decrease in the population of a key herbivore like *A. dactylomela* would become a

compounding problem as there would be reduced herbivory on the macroalgae. These two factors together could contribute to a phase shift from a coral to an algal dominated reef and the loss of valuable coral reef ecosystems around the world [13]. This shift could also disrupt entire reef ecosystems, as many other organisms that are dependent on the corals for food and shelter would also come under threat due to their loss of habitat.

Further research on the impacts of acidification and sea surface temperature rise is key to understanding how sea hares will react to the changing ocean environment. Continued observations of sea hare larvae throughout the planktonic larval stage, and perhaps over the course of their entire life, will offer insight into how acidification and heat affect the survival of sea hares to sexual maturity. Such observations could help predict how the sea hare populations will be impacted in the future. Understanding the threats to key herbivores like sea hares can inform population management, which is key to preventing coral reefs from shifting to algal dominated systems in the future. Assessing the impact of ocean acidification and sea surface temperature rise on sea hares and other marine calcifying larvae will be critical in understanding and protecting sensitive coral reef environments and other ecologically important calcifying organisms that are important in the biological food web as well as the marine carbon cycle.

ACKNOWLEDGEMENTS

I would like to thank the staff at Heron Island Research Station for their help in the field and Dr. Selina Ward for designing and providing the experimental ocean acidification system. I would like to thank Dr. Kevin Arrigo for his help with the project. I would also like to thank the International Programs Office of University of Queensland and the Bing Overseas Program of Stanford University for making this research possible.

REFERENCES

1. Pachauri RK *et al.*, "Climate change 2014: synthesis report to the 5th assessment report to the intergovernmental panel on climate change," Intergovernmental Panel on Climate Change, Geneva, Switz., 151, 2014.
2. Orr JC *et al.*, "Anthropogenic ocean acidification over the twenty-First Century and its impact on calcifying organisms," *Nature*, vol. 437, pp. 681-686, Sept. 2005.
3. Albright R and C Langdon, "Ocean acidification impacts multiple early life history processes of Caribbean coral *Porites asteroids*," *Global Change Bio.*, vol. 17, pp. 2478-2487, Mar. 2011.
4. Byrne M *et al.*, "Impacts of ocean warming and ocean acidification on marine invertebrate life history stages: vulnerabilities and potential for persistence in a changing ocean," *Ocean. and Marine Bio.: Annu. Review*, vol. 49, pp. 1-42, Jul., 2011.
5. Fabry VJ *et al.*, "Impacts of ocean acidification on marine fauna and ecosystem processes," *ICES J. Mar. Sci.*, vol. 65, pp. 414-432, Feb., 2008.
6. Crowley TJ, "Causes of climate change over the past 1000 years," *Science*, vol. 289, pp. 270-277, Jul., 2000.
7. Yusa Y, "Factors regulating sperm transfer in an hermaphroditic sea hare, *Aplysia parvula*," *J. Exp. Marine Bio. And Ecology*, vol. 181, pp. 213-221, Sep., 1994.
8. MacGinitie GE, "The egg-laying activities of the sea hare *Tethys californicus* (cooper)," *Bio. Bulletin*, vol. 67, pp. 300-303. Oct., 1934.
9. Blackenship JE and RE Coggeshall, "The abdominal ganglion of *Aplysia brasiliana*: a comparative morphological and electrophysiological study, with notes on *A. dactylomela*," *J. Neurobiology*, vol. 7, pp. 383-405. Sept., 1976.
10. Lindstrom A, "The combined impacts of ocean acidification and sea surface temperature rise on embryonic sea hare (*Aplysia dactylomela*) development," Targeted Research Project Report, Stanford in Australia, University of Queensland, 2012. Unpublished.
11. Mather P and I Bennett, *A Coral Reef Handbook*, 3rd ed. Brisbane, Australia: Aus. Coral Reef Society, 1993.
12. Chan KY *et al.*, "Effects of ocean-acidification-induced morphological changes on larval swimming and feeding," *J. Exp. Bio.*, vol. 214, pp. 3857-3867, Nov., 2011.
13. Hughes, TP *et al.*, "Phase shifts, herbivory, and the resilience of coral reefs to climate change," *Current Bio.*, vol. 17, pp. 360-365, Feb., 2007.



ANJA MALAWI BRANDON

Anja is a senior majoring in environmental systems engineering with a focus on the coastal environment. Originally from Seattle, Washington, Anja's drive to care for the marine environment started young from growing up in and around the Puget Sound. At Stanford, Anja does research in the environmental engineering department and pursues her other passion of rock climbing by both working at the Stanford climbing gym and being on the climbing team. She plans to continue her research on impacts on the ocean in graduate school pursuing a doctoral degree in environmental engineering studying plastic pollution. In her free time, Anja loves getting outside to go climbing, hiking, and exploring the world.

Vigilance in hand-reared agile wallabies (*Macropus agilis*) after release

Amalia Saladrigas

Stanford University

Considering the impact human development continues to have on wild Australian fauna such as agile wallabies (*Macropus agilis*), it is important to not only establish wildlife rehabilitation centers, but also study the effects this different upbringing may have had on the survival of hand-reared agile wallabies once they are released into the wild. Vigilance is a vital anti-predator behavior: in order to survive in the wild, wallabies must spend a substantial amount of time looking out for possible threats in their surroundings. Specifically, the study investigated whether time spent vigilant changed progressively after release, comparing wallabies that had been raised and released in Sheoak Ridge in 2011-2013 with those released in 2014. Wallabies released in 2011-2013 allocated 23.63% of their time to vigilance whereas wallabies released in 2014 allocated 22.67% of their time to vigilance. Overall, hand-reared wallabies dedicated 22.96% of their time to vigilance. There was no significant difference in the time spent vigilant between wallabies released in 2011-2013 and those released in 2014, nor between all hand-reared wallabies and the 21% average found by Stirrat for wild wallabies during dusk [1]. Vigilance in hand-reared agile wallabies did not change progressively after release, and was similar overall to the percentage of time allotted to vigilance by wild wallabies. These results suggest that wildlife rehabilitation efforts are successfully preparing hand-reared wallabies to survive in the wild, and can serve as groundwork that we can continue to improve upon in order to alleviate our impact on their natural communities.

Introduction

Agile wallabies (*Macropus agilis*) are the most common macropod species found in northern Australia [1]. Although the International Union for Conservation of Nature (IUCN) has given agile wallabies a conservation status of “least concern,” wallaby populations are experiencing increasing stress from introduced predators such as red foxes (*Vulpes vulpes*) and dingoes (*Canis lupus dingo*) [2, 3], native predators such as wedge-tail eagles (*Aquila audax*) and salt water crocodiles (*Crocodylus porosus*) [4], as well as encroaching human development [5].

Wallabies are herbivores and feed mainly on grass. During periods of grazing or foraging, the wallaby’s attention is diverted, making it more vulnerable to predator attacks. Previous studies have shown that in order to manage the risks associated with feeding, agile wallabies travel in mobs of about 10 individuals, with alternating periods of vigilance to provide mutual protection. Vigilance is defined as a period where the wallaby is actively alert or watchful, a crucial behavior to survival. In general, animals that have predators, especially terrestrial predators, will “trade off foraging with anti-predator vigilance” [4]. A study conducted by Stirrat found that in the dry season wild agile wallabies spend about 56% of their time foraging. Vigilance represented approximately 17% of their activity budget during the daytime, 16% at night, and 21% at dusk in the dry season [1].

Australia’s increasing urban sprawl has led to an increased number of wildlife-car collisions where native animals, like wallabies, are often the victims [6]. Joeys that do not die on

impact can survive up to five days inside the pouch of their dead mother. The increase in wildlife-car collisions has resulted in more orphaned joeys raised in captivity by trained wildlife caretakers [7]. While human contact is minimized, and caretakers try to emulate the circumstances the joey would experience in the wild, it is possible that developing in an artificial environment could impact the wallabies’ behavior. If development in captivity affects the wallabies’ vigilant behavior, such repercussions could be detrimental once the wallabies are released into the wild.

The wallabies used in this study were raised at Sheoak Ridge Nature Reserve in Far North Queensland. Sheoak Ridge is a reserve highly valued for its ecological diversity: its 165 acres are home to open eucalyptus woodlands, riparian rainforest, melaleuca wetlands, billabongs, and complex ecotones [8]. In Queensland, particularly in the coastal areas, such as where the reserve is located, the number of road kill accidents increases during sugarcane harvesting season between May and November. Traffic may intensify to six trucks an hour twenty-four hours a day on the main road located 13km from Sheoak Ridge [9].

Agile wallabies are considered ready to be released into the wild when they reach a weight between 6-7 kilograms, as long as their caretakers see no patterns or behaviors that would excessively endanger the animal, such as if a wallaby is considered to be too comfortable around humans, or if the animal is apathetic. At Sheoak Ridge wallabies go through a “staged pen release” where they live in pens of increasing size with decreasing amounts of human contact. The staged release allows the wallabies to familiarize themselves with other wallabies and with the predator sounds that would be natural in the wild. Finally, the wallabies are

released under “soft release.” The doors to the pen are opened, allowing the wallabies to leave to go into the wild, but they can return at any time to their pen for additional support, such as supplementary feeding. This leads to a much less stressful release than “hard release,” where wallabies are released into the wild with no support system. Hard release is known to have a much higher mortality rate [9].

This study aims to compare the general activity budget of recently released wallabies with those that have already lived in the wild between 1-3 years. In particular, we investigate the percentage of time spent vigilant between recently released and wallabies released 1 to 3 years ago. While previous studies have focused on activity budget allocation in wild wallabies [1,10] and the activity budget and vigilance in captive wallabies as compared to wild wallabies [11,12], there are no studies that assess how or whether vigilance in released, hand-reared wallabies changes progressively after release. Most research concerning captive wallabies studies their physiology and biology. In fact, there are no studies that focus exclusively on the behavior of captive agile wallabies, and there exists even less information about the behavior of rehabilitated and released wallabies. Most of the information on released wallabies is anecdotal because placing tracking devices on released animals is both invasive and expensive [9].



Figures 1 and 2. Sheoak Ridge – These are photos of the lawn where released wallabies may come back twice a day for supplementary feed.

While it may be difficult to accurately measure if the time that wallabies spend in human care has any repercussions on their general behavior after release, this study aims to provide a first approach at assessing the impact that time spent in care may have in the anti-predator behavior of released hand-reared agile wallabies.

METHODS

This study uses behavioral observations of wallabies taken into care as early as 2010 and released back into the wild as recently as September, 2014. Observations quantifying the general behavior of the animals yielded data on 17 different wallabies, 10 of who were released in 2014 and 7 who were released between 2011 and 2013.

Observations were taken on October 8th and 9th, 2014, towards the end of the dry season. For the purpose of this study, wallabies released before 2014 are considered to have spent more time in the wild, and those released in 2014 are considered recent releases. These observations were carried out at Sheoak Ridge, in North Queensland, Australia.

Observation Sessions

Observations were conducted between 6-8am in the morning and between 4-6pm in the afternoon. Observers sat on a couch facing an open lawn where the wallabies may come get supplemental food twice a day. The lawn is surrounded by rainforest on two sides, woodland on another, and a house behind the observers. There

Behavioral Category	Description	Position
Foraging	Anytime the wallaby was observed eating or searching for food. Regurgitation was classified under foraging while standing (F2).	(F1) = Crouching
		(F2) = Standing
		(F3) = Propped up
Vigilance	Anytime the wallaby was acting alert, whether the wallaby looked up from feeding, scanned his surroundings, or focused on a specific object or location.	(V1) = Crouching
		(V2) = Standing
		(V3) = Propped up
Locomotion	Anytime the wallaby moved from one place to another.	(LW) = Walking
		(LH) = Hopping
Other	Any behavior that did not fit the descriptions for the previous categories. Times when the animal exhibited alertness or foraged while lying down, or times when the wallaby did not engage in any particular activity, were considered as resting (OR).	(OS) = Socializing
		(OG) = Grooming
		(OR) = Resting

Table 1. Categories of Behavior – Description and characterization of the four categories used to classify the wallabies’ behavior.

Name	Sex	Stage of life when wallaby came into care	Date of arrival at Sheoak Ridge**	Date of release	Time since release	Weight at release
Harley	Male	Velvet	Oct. 18, 2010	Aug. 17, 2011	38 months	6kg
Wee Willy	Male	Velvet	Nov. 13, 2011	Jan. 27, 2012	34 months	~6kg
Lara	Female	Pinkie	Dec. 1, 2011	Apr. 3, 2012	30 months	~6kg
Barry	Male	Furred	Jan. 1, 2012	May 9, 2012	29 months	6kg
Tiffany	Female	Velvet	June 7, 2012	Sept. 9, 2012	25 months	~6kg
Holly	Female	Velvet	Feb. 15, 2013	Jul. 16, 2013	15 months	~6kg
Jojo	Female	Pinkie	Feb. 21, 2013	Jul. 16, 2013	15 months	~6kg
Poppy	Female	Pinkie	Jan. 9, 2014	May 25, 2014	5 months	~6kg
Archie	Male	Furred	Jan. 21, 2014	May 25, 2014	5 months	~6kg
Meika	Female	Furred	Jan. 21, 2014	May 25, 2014	5 months	~6kg
Jacob	Male	Furred	Feb. 7, 2014	Aug. 9, 2014	2 months	~6kg
Laura	Female	Velvet	May 15, 2014	Aug. 9, 2014	2 months	5kg
Joseph	Male	Pinkie	Apr. 27, 2014	Sept. 23, 2014	<1 month	4.7kg
Josephine	Female	Pinkie	Apr. 27, 2014	Sept. 23, 2014	<1 month	5kg
Midget	Female	Velvet	Apr. 27, 2014	Sept. 23, 2014	<1 month	5.5kg
Oscar	Male	Velvet	Jul. 27, 2014	Sept. 23, 2014	<1 month	5kg
Texas	Male	Furred	Jul. 27, 2014	Sept. 23, 2014	<1 month	4.5kg

Table 2. Individual histories for the 17 observed wallabies. **The date of arrival at Sheoak Ridge is not necessarily the date the wallabies came into care. The wallabies may have spent time at a veterinary clinic.

were 16 plates dispersed throughout the area for supplementary feeding at dusk and dawn. The primary wildlife caretaker at the site sat near the observers to provide the names and histories of the wallabies present.

Upon arrival at the property, a waiting period of 10 minutes was allotted before observations were taken in order to minimize disturbance from the vehicle's arrival and allow the wallabies to grow accustomed to human presence. The wallabies were then observed individually, in random order, for a period of five minutes each.

There was a distance of about 3-4m between the lawn and the observers so that, once the wallabies became accustomed to our presence, the interactions were minimal. Those wallabies that approached the human observers and their caretaker were not evaluated for the time that the interaction lasted.

Classification of Behavior

The wallabies' behavior was measured from the moment a behavior started to the moment that it stopped, and was classified into one of the four categories described in Table 1 [1, 5, 11, 12].

Analysis

The data was analyzed for the four main behavioral categories listed above, and was then used to calculate the percentage of time the wallabies spent on each of the four main categories per day, both individually and as a group according to their date of release (2011-2013 or 2014). All data is reported as each behavior's averages \pm standard error. After analyzing the data using a two-tailed t-test, differences are considered significant when the p-value is less than 0.05.

The statistics found for hand-reared wallabies in this study were compared with the values found in Stirrat's 2004 study on wild wallabies' behavior activity budgets during the dry season. The comparable value for vigilance between the two studies was

considered to be the percentage of time spent vigilant at dusk, which represented 21% of the wallabies' total activity budget [1].

Developmental Stages When Wallabies Came into Care

Joey's that are rescued from a dead or injured mother's pouch may be found at various stages of development [7]. These can be classified as the following:

Pinkie: The stage lasts up to when the joey is 3 – 3 ½ months old of age. Joey's rescued at this stage are often found still fused to the teat. They do not have fur and cannot thermo-regulate [9].

Velvet: This stage lasts until the joey is between 4-5 months old. The joey's skin gets darker as the beginnings of fur start to show. Joey's begin to poke their heads out of the pouch [9].

Furred: Agile joey's become properly furred at about 5 months of age. At this stage they will also start to come out of the pouch for short periods of time. Usually, at 6 months old they will spend most of the time outside the pouch, and at about 7-9 months of age, depending on the individual wallaby and the mother, they will be completely out of the pouch [9].

RESULTS

Individual History

There were 17 wallabies observed during the two-day observation period. The developmental stage when the wallaby entered care, date of arrival and release at Sheoak Ridge, time since release, and weight at release were recorded for each wallaby (Table 2). The wallabies released in September of 2014 had not yet reached the ideal weight for release (6-7kg), but were let out earlier due to research activities happening in October at Sheoak Ridge.

Vigilance Related to the Overall Behavior Budget

Vigilance was the second most frequent behavior following foraging in the wallabies released in 2011-2013 and those released in 2014 (Figure 3). The overall mean percentage of time spent vigilant for wallabies that have already spent some time in the wild (2011-2013) was 23.6%±3.2% whereas the mean percentage

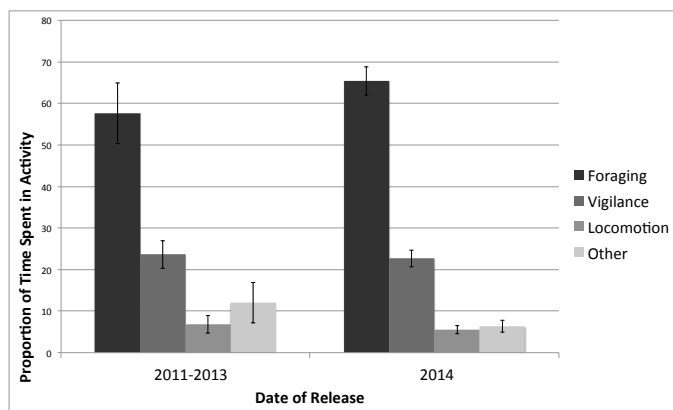


Figure 3. Mean Behavior Budget According to Date of Release – The graph displays the overall average time percentages spent on foraging, vigilance, locomotion, and other behaviors for wallabies released in 2011-2013 and wallabies released in 2014.

was 22.7%±2.0% for wallabies released in 2014. The t-test yielded a p-value of 0.59, making the difference not statistically significant. The overall average percentage of time spent foraging was 57.6%±7.4% for wallabies released in 2011-2013 and 65.4%±3.4% for wallabies released in 2014. While the percentage of time spent foraging was greater for recently released wallabies, the difference was found to be statistically insignificant (p=0.15).

Both groups of wallabies spent similar amounts of time in locomotion. Wallabies released in 2011-2013 spent on average 6.8%±2.1% of their time moving while those released in 2014 spent 5.6%±1.0% of their time moving on average. Wallabies released in 2011-2013 spent 12%±4.8% of their time grooming, socializing, or resting, while wallabies released in 2014 spent 6.4%±1.4% doing the same.

Vigilance Related to Stage of Life When the Wallabies Came into Care

Wallabies that were released in 2011-2013 and came into care furred, the last stage of a joey's development, appear to have spent a higher percentage of time being vigilant than wallabies that were released the same year but came into care as pinkies or velvets (Figure 4). Nonetheless, for this measurement only one individual was assessed and was found to spend 50.7% of his time being vigilant. Wallabies that came in as pinkies and were released in 2011-2013 exhibited a mean vigilance time of 19.9%±3.6%, and those that came in as velvets exhibited a mean of 22.4%±3.7%. The difference between these last two groups was not statistically significant (p=0.15).

Among the wallabies released in 2014, those that came in as velvets spent the most time being vigilant (Figure 5). The average percentage of time spent vigilant for wallabies that came into care as velvets was 24.3%±4.7%, while those that came in as pinkies spent 23.7%±2.6% of their time being vigilant, and those that came in furred allocated 19.8%±3.5% of their time to the same behavior. None of these differences were found to be statistically significant.

Those wallabies that came into care as pinkies and velvets spent similar amounts of time vigilant across the group released in 2011-2013 and the group released in 2014. There was a

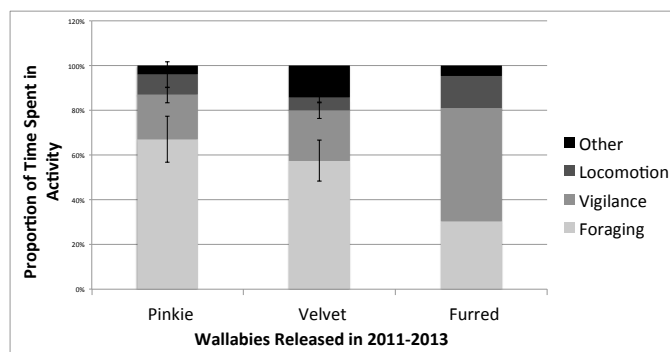


Figure 4. Mean Behavior Budget According to Stage of Life When Wallabies Came into Care (2011-2013) – This graph displays the overall average time percentages spent on each of the four previously described behavioral categories for wallabies that came into care as pinkies, velvets, or already furred, and were released in 2011-2013. The furred category was composed of a single individual for 2011-2013.

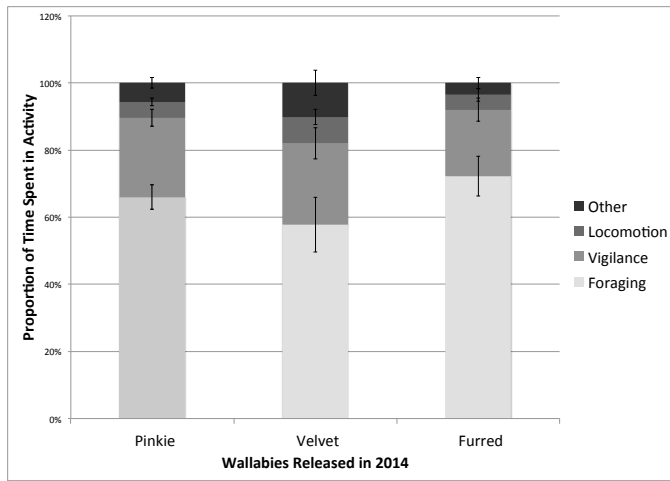


Figure 5. Mean Behavior Budget According to Stage of Life When Wallabies Came into Care (2014) – This graph displays the overall average time percentages spent on each of the four previously described behavioral categories for wallabies that came into care as pinkies, velvets, or already furred, and were released in 2014.

statistically significant difference ($p=0.02$) between wallabies that came in as pinkies in 2011-2013 and those released in 2014, which suggests that recently released wallabies spend more time alert than wallabies that have already spent some time in the wild. Wallabies that came in as velvets and were released in 2011-2013 and those recently released did not exhibit a significant difference in the amount of time spent vigilant. Although there was a great difference in mean vigilance between those wallabies that came in furred and have already spent some time in the wild as compared to those that have been released recently, it was not found to be statistically significant either.

DISCUSSION

Hand-reared wallabies released in 2011-2013 and those released in 2014 displayed a similar overall activity budget, including the percentage of time allocated to vigilance. Contrary to what may be expected, recently released wallabies do not devote more time to anti-predator behavior than wallabies that have already spent some time in the wild. In both groups, wallabies spent the greatest percentage of their time foraging. Allocating a considerably larger percentage of time to foraging rather than to vigilance may be due to the fact that wallabies are less vigilant in the dry season, when food availability decreases, than in the wet season [10]. Foraging comprises 56% of the activity budget of wild wallabies in the dry season and 42% in the wet season [1]. Wallabies in this study spent 57.6% (2011-2013 releases) and 65.4% (2014 releases) of their time foraging, which is higher in both cases than the percentage we could expect to see in the wet season.

Although the difference was not significant, wallabies that had been released recently spent a greater percentage of time foraging than those released in 2011-2013. Five of the wallabies had been released very recently, less than three weeks before this study was carried out, and they had not yet reached the ideal weight for release of 6-7kg. Wallabies released in 2014 in general, and the aforementioned five in particular, are younger than those released in 2011-2013, and therefore still growing, so it would be

presumed that they require larger quantities and more nutritious food [10].

Conversely, although the difference was not statistically significant, wallabies released in 2011-2013 spent more time engaging in behavior classified as “Other” than wallabies released in 2014, suggesting that wallabies that have already spent some time in the wild spend more time grooming, resting, and socializing. The difference is not unusual, for wallabies that were released earlier are older, and are becoming or have already become sexually mature. They are more likely to engage in interactive social behavior. Much like the younger wallabies need to spend more time foraging due to the demands of their developmental stage, sexual maturity can cause older wallabies to allocate time differently in their activity budget [9].

The stage of development at which the wallabies came into care (pinkies, velvets, and furred joeys) did not greatly influence the percentage of time allocated to vigilance. For the group of wallabies released in 2011-2013, Barry, the only individual observed that came in furred, spent 50.7% of the time being vigilant, while wallabies that came in as pinkies and velvets allocated 19.9% and 22.4% to vigilance, respectively. Barry is a mature male and had been observed persistently courting females prior to, as well as during, this study [9]. This makes him more likely to be alert, looking around in order to be aware of any potential mates in his surroundings; which explains the much greater portion of time dedicated to vigilance. There was no difference between the wallabies that came into care as pinkies and velvets and were released in 2011-2013.

Although the wallabies that came into care as velvets had similar activity budgets across both release groups, there was a small significant difference in the time spent vigilant between the wallabies that came into care as pinkies. This suggests that recently released wallabies that came in as pinkies spend a greater percentage of time being vigilant (23.67%) than those released in 2011-2013 (19.9%). This difference in vigilance may be due to the fact that the recently released wallabies are younger. In addition, some of the wallabies released in 2014 that came in as pinkies were released shortly before this study was conducted. The recently released wallabies may be more attentive towards human presence, which could erroneously inflate the amount of time they spend vigilant. The last wallaby released in 2011-2013 that came into care as pinkies was released more than a year before the last wallaby that came in as a pinkie was released in 2014.

Wallabies that came into care furred and were released in 2014 spent significantly less time being vigilant (19.8%) than the individual released in 2011-2013 (50.7%). Once again, this was probably greatly influenced by Barry’s inclination to be, as a sexually mature male, particularly alert towards the other wallabies in the group, rather than by the life stage at which the wallabies came into care.

In general, aside from the category of the wallabies that came in furred and were released in 2011-2013, all other groups show similar levels of vigilance. This result is not surprising, considering that the wallabies went through the same type of care and release regardless of the life stage at which they were brought into care.

A previous study found that wild wallabies forage for an average of 56% of their total activity budget in the dry season [1], while hand-reared wallabies in this study foraged for an average of $63\% \pm 3.3\%$ of their time. Nonetheless, this increase in time

spent foraging could be due to the supplementary feeding offered at the observation site that wallabies would not have access to in the wild. It is natural, especially in the dry season, for wallabies to broaden their diet “to include alternative, higher-quality foods in order to supplement their relatively nutrient-poor herbage diet” [1]. They are therefore more likely to take advantage of the supplementary feeding that is conveniently laid out for them.

In the dry season, vigilance represent approximately 21% of the wild wallabies’ activity budget at dusk [1], the time of day that compares best to the time at which observations for this study were taken [9]. Meanwhile, released hand-reared wallabies observed in this study allocated $22.9\% \pm 1.7\%$ to vigilance overall. Furthermore, in a concurrent study carried out on a population of wild wallabies at Trinity Beach, in Far North Queensland, vigilance represented 23.7% of the wallabies’ total activity budget (Yates L, unpublished Targeted Research Project, 2014). The difference between the results found in this study and the results found by Yates was statistically insignificant, and our values were very close to those established by Stirrat.

The lack of a significant difference in the percentage of time allotted to vigilance or anti-predator behavior between wild and hand-reared wallabies speaks well for the wildlife rehabilitation strategies implemented to date. Without awareness for potential threats, their survival in the wild could be compromised. This study aimed to assess the anti-predator behavior of released hand-reared wallabies, as well as if and how this behavior changes progressively after release: in spite of being raised by humans, hand-reared wallabies develop the same, natural perspicacity towards their surroundings as wild wallabies do.

There are, nonetheless, some limiting factors to this study. Due to the fact that the observations could only be recorded over a period of two days, the main concern is lack of data on all the wallabies living at Sheoak Ridge, which limits the generalizability of these results. Conducting the study over a longer period of time would likely allow for greater data collection, and may yield more reliable results. Repeating the study in the wet season, as well as at different times during the day, may similarly yield different results.

Furthermore, there was only data for one wallaby that came into care furred and were released in 2011-2013. This could have potentially skewed the results, making it appear that individuals that came in furred and were released in 2011-2013 spent more time being vigilant than those released in 2014. While we gathered data on 10 wallabies released in 2014, we gathered data on only 7 wallabies released in 2011-2013. Given that the observation site is also where wallabies may come twice a day for supplementary feeding, and therefore involves more exposure to humans, this may actually prove to be a good sign: wallabies released in 2011-2013 may be less present because they are more used to being in the wild and have developed a natural apprehension towards humans.

Moreover, audible noise from the sugar cane trucks and other activity in the neighboring plantation as well as occasional disturbance from other research studies being conducted at the property at the time may have affected the overall amount of time the wallabies spent alert. The familiarity of the observation site may have made the wallabies feel safer and less likely to be as vigilant as they normally are in the wild. They may have felt particularly comfortable around their caretaker: the wallabies were more likely to approach the primary wildlife caretaker than the observers, but these interactions happened at most twice per

observation session. Finally, it is also possible the results would vary if the gender of the wallabies was factored in, as well as if wallabies raised by different caretakers and under different circumstances were taken into account.

CONCLUSION

This is the first study that looked at how hand-reared wallabies’ activity budgets change progressively after release, particularly the percentage of time allotted to vigilance and anti-predator behavior. The similarity in vigilance between released hand-reared wallabies and the wild-reared wallabies in Stirrat’s 2004 study suggests that the anti-predator awareness in hand-reared wallabies is not impacted. The percentage of time that wallabies released in 2011-2013 spent vigilant (23.6%) did not differ significantly from the percentage of time that wallabies released in 2014 spent doing the same (22.7%), suggesting that the care provided in wildlife rehabilitation appropriately prepares hand-reared wallabies for the wild.

This study aims to contribute to the research being done on captive agile wallabies and rehabilitated wallabies, particularly after their release into the wild. The results provide an initial assessment on the effectiveness of rescue care and rehabilitation in agile wallabies, but it is crucial to continue studying these animals to ensure that the potential negative human impact humans may have on their population has feasible, healthy solutions. While tracking devices may be too intrusive for the animals, the sample size under observation can be expanded to include multiple wildlife rehabilitation facilities. Further studies are needed to assess whether differences in the type of care and release into the wild impact wallaby survival and quality of life once they are released back into the wild. Increasing our knowledge on the wallabies’ progress after release enables us to improve the management and care of captive agile wallabies. Further research will allow us to know whether we can, at least, suitably mitigate our negative impact on wildlife, and even sometimes redress it.

ACKNOWLEDGEMENTS

I would like to thank the Bing Overseas Studies Program in Australia for providing me with the opportunity to study in an incredible country and Dr. Claire Baker from the University of Queensland for advising me throughout this research project.

REFERENCES

1. Stirrat SC. Activity Budgets of the Agile Wallaby, *Macropus Agilis*. *Australian J Zoology* 2004; 52: 49-64.
2. Banks PB, Newsome AE, Dickman, CR. Predation by Red Foxes Limits Recruitment in Populations of Eastern Grey Kangaroos. *Austral Ecology* 2000; 25:283-291.
3. McKenzie NL, Burbidge AA, Baynes A et al. Analysis of Factors Implicated in the Recent Decline of Australia’s Mammal Fauna. *J Biogeography* 2007; 34:597-611.
4. Blumstein DT, Daniel JC, Sims RA. Group Size but Not Distance to Cover Influences Agile Wallaby (*Macropus Agilis*) Time Allocation. *J Mammalogy* 2003; 1.84:197-204.
5. Pest Risk Assessment: Agile wallaby (*Macropus agilis*). Invasive Species Branch, Department of Primary Industries, Parks, Water and Environment (DPIPWE), Hobart, Tasmania, 2013.
6. Rowden PJ, Steinhardt DA, Sheehan MC. Road crashes involving animals in Australia. *Accident Analysis and Prevention* 2008; 40(6):1865-1871.

7. Sadler S. Husbandry Guidelines for Agile Wallaby (*Macropus Agilis*). Certificate 3 Captive Animals, Western Sydney Institute of TAFE, Sydney. July 2009.
8. Sheoakridge.com, 'One man's Trash is another man's Treasure', 2014. [Online]. Available: <http://sheoakridge.com>. [Accessed: 10-Mar-2016].
9. Baker, Claire. "Wallabies in Wildlife Care." Personal interview. 8 Nov. 2014.
10. Stirrat, Simon C. "Foraging Ecology of the Agile Wallaby (*Macropus Agilis*) in the Wet-dry Tropics." *Wildlife Research* 29 (2002): 347-61. CSIRO Publishing. Web. 2 Oct. 2014.
11. Krumm, Justin. "Vigilance in Captive Reared Agile Wallabies (*Macropus agilis*) before soft release in the wet-dry tropics of Far North Queensland." Unpublished data: Targeted Research Project (2010). Web. 2 Oct. 2014.
12. Tran, Thuy-An. "Characterizing the Activity of Agile Wallabies in Rescue Care." Unpublished data: Targeted Research Project (2010). Web. 2 Oct. 2014.



AMALIA SALADRIGAS

Amalia Saladrigas-Malaret is a senior at Stanford University majoring in Biology with a focus on Marine Biology. She is a pre-veterinary student especially interested in large mammals. Raised in both Barcelona and Puerto Rico, she has always loved the ocean. Her honors thesis looks into the three-dimensional morphology of baleen across different species of baleen whales and its implications for their filter-feeding mechanisms. In her free time, Amalia enjoys traveling, reading, learning new languages, watching movies, and swimming.