## The Journal of the Utah Academy of Sciences, Arts, & Letters Volume 93 • 2016

Includes selected and refereed articles from the 2016 Annual Conference held at Westminster College March 10, 2016

**Editor** Kristin L. Kraus

Copyright 2017, Utah Academy of Sciences, Arts, & Letters.

Copyright reverts to authors after publication, Printed in the United States of America by Brigham Young University Academic Publishing, Provo, Utah. Neither the editors nor the sponsoring organization assumes responsibility for statements of fact or opinion made by the contributors.

ISBN-13: 978-0-9988268-0-6 All Rights Reserved

#### **Board of Editors**

Arts: Angie Banchero-Kelleher, Utah Valley University Biological Science: Laurie Mauger, Southern Utah University Business: Taowen Le, Weber State University Education: Laura Snelson, Utah Valley University Engineering: Doran Baker, Utah State University Exercise Science and Outdoor Recreation: L Nathan Thomas, Salt Lake Community College Letters—Humanities, Philosophy, Foreign Language: David Richter, Utah State University Letters-Language and Literature: Keith Lawrence, Brigham Young University Physical Sciences: Rico Del Sesto, Dixie State University Social Science: Wade M. Cole, University of Utah Posters: **R. Steven Turley,** Brigham Young University 2015-2016 Utah Academy of Sciences, Arts, & Letters Officers President: Jonathan Westover, Utah Valley University President-Elect: H. Laine Berghout, Weber State University Past President: Erin O'Brien, Dixie State University

Secretary: Susanna Garcia, Utah Valley University Treasurer: Karl Israelsen

Scholary Communications: **Peter Kraus**, University of Utah Member at Large: **Dwight Israelsen**, Utah State University Technology Officer: **Rachel Ramsey**, Dixie State University Institutional Representatives: **Rachel Keller**, Snow College

**Cover Photo:** Alessandro Perego, Cameron Olsen, J. Ryan Peterson, Kameron Hansen, John S. Colton, Richard K. Watt. "Fabrication of Dye-Sensitized Solar Cells Using Native and Non-Native Nanocrystals in Ferritin as the Dye," p. 183.

## Utah Academy of Sciences, Arts, and Letters

History: Founded 3 April 1908, the Utah Academy of Sciences was organized "to promote investigations and diffuse knowledge in all areas of science." Beginning in 1923, the Academy started publishing the papers presented in its annual meetings in *Proceedings*. In June 1933 at the annual meeting, the Academy was enlarged to include arts and letters, and the name was changed to the Utah Academy of Sciences, Arts, and Letters. Articles of incorporation and non-profit organization status were accepted by the Academy membership at the spring meeting in April 1959. In 1977, the name of the journal of the Academy was changed from Proceedings to Encyclia. It became a refereed journal at this time. In the mid 1980s, the scope of the Academy was expanded further to include (1) business, (2) education, (3) engineering, (4) library information and instruction, and (5) health, physical education, and recreation. Beginning with the 1998 issue, the journal became The Journal of the Utah Academy of Sciences, Arts, and Letters

**Annual Meeting:** The Academy's annual meetings are normally held in the spring on one of the Utah campuses of higher education. The plenary session is called the Tanner Lecture, endowed by Mr. O.C. Tanner in 1986.

**Best Paper Awards:** The best paper presented in every division is given a cash award, which is presented at the Academy's "Awards Evening" held the following fall.

**Distinguished Service Awards:** The Academy recognizes outstanding contributions to teaching and scholarship by means of annual Distinguished Service Awards, alternating every other year between disciplines.

**Membership:** When the Academy was founded in 1908, membership was by nomination, ratified by the Council, and elected by a "three-fourths votes of members present." Today, the Academy's membership is available by application.

**Institutional Members:** All Utah institutions of higher education are members of the Utah Academy. The Academy appreciates their patronage.

## **Publication Policy**

The Journal of the Utah Academy of Sciences, Arts, and Letters publishes works in all of the fields of study encompassed in the Academy's mission. Papers published in *The Journal of the Utah Academy of Sciences, Arts, and Letters* are drawn from papers presented by members in good standing at the annual conference of the Utah Academy. To qualify for publication, the papers must be recommended through a refereeing system.

Presenters are encouraged to publish their paper in *The Journal of the Utah Academy. The Journal's* criteria are that a submission is (1) fresh, meaningful scholarly insight on its subject; (2) readable and well written; and (3) of general interest for an academic readership beyond the author's field.

If you wish your paper to be considered for publication in The Journal, please submit a Microsoft Word document to the section editor of the appropriate section by the indicated deadline. Contact information for the section editors is available on the Utah Academy's website (www.utahacademy.org).

*The Journal of the Utah Academy* is a refereed journal. Editorial responses will be forthcoming after the resumption of school the following fall when referees have returned their comments to the division chairs.

Papers should be between ten and twenty double-spaced pages. Detailed instructions to authors are available at http://www.utahacademy.org/ Instructions\_for\_Authors.pdf.

Among the bibliographic services listing at Bowker Serials Bibliographies and The Standard Periodical Direction. Indexing and abstracting services that cite articles in the journal include Arts and Humanities Citation Index, Biosciences Information Services, Current Geographical Publication, Chemical Abstracts, Mathematical Reviews, MLA Biography, Sociological Abstracts, Excerpta Botanica, Social Planning, Policy and Development Abstracts, Language and Language Behavior Abstracts, Index to Scientific Technical Proceedings, and Index to Social Sciences, and Humanities Proceedings.

## The Journal of the Utah Academy of Sciences, Arts, & Letters—2016

## AWARDS

Distinguished Service Award	11
Academy Fellow 2017	12
O.C. Tanner Lecture	14
John and Olga Gardner Prize	15
Honorary Member	16
2016 Best Paper Awards	17

## ARTICLES

## **BIOLOGICAL SCIENCES**

Chronic Effects of Intermittent Sound Disturbance on	
Buenos Aires Tetra	19
Jeremy L. Arnt, Erica Neisler, Weber State University	
Student's t Test and Scientific Reproducibility	37
Brian Knaeble, <sup>1</sup> Julian Chan, <sup>2</sup> Russell Costa, <sup>3</sup> Katie McLean, <sup>3</sup>	
Douglas Getty, <sup>3</sup> Riddhi Rampeearee <sup>3</sup> , <sup>1</sup> Utah Valley University;	
<sup>2</sup> Weber State University; <sup>3</sup> Westminster College	
Integration of Produced Water with Microalgae	
Cultivation	55
Brian McNeil, <sup>1</sup> Derek Hess, <sup>1</sup> Eric Torres, <sup>1</sup> Ronald C. Sims, <sup>1</sup> Jason	
C. Quinn <sup>2</sup> , <sup>1</sup> Utah State University, <sup>2</sup> Colorado State University	

X-ray Analysis of Lead(II) Binding to <i>Haloferax volcanii</i> Malate Synthase	77
Michael J. Adams, Bruce R. Howard, Southern Utah University	
BUSINESS	
Exploring Shifting Global Labor Management Practices and Comparative Job and Life Satisfaction Bergen Eskildsen, Joe Light, Kaitlin Carlisle, and Jonathan H. Westover, Utah Valley University	97
<b>Corporate Payouts and Macroeconomic Influences</b> Lauren Lo Re, <sup>1</sup> Robert Patterson, <sup>1</sup> and Mahfuz Raihan <sup>2</sup> , <sup>1</sup> Westminster College and <sup>2</sup> University of Utah	119
A Framework for Teaching the Goal of the Firm in Introductory Business Classes: Shareholder Wealth Maximization Ethicality and Classical Philosophical Paradigms Rebekah Inez Brau, <i>Brigham Young University</i>	135
EDUCATION	
Evaluation of the Effect of Study Skills and Lifestyle Factors on Performance in Organic Chemistry Don R. Davies, Heather Root, Valerie Herzog, <i>Weber State</i> <i>University</i>	163
ENGINEERING	
Fabrication of Dye-Sensitized Solar Cells Using Native and Non-Native Nanocrystals in Ferritin as the Dye Alessandro Perego, Cameron Olsen, J. Ryan Peterson, Kameron Hansen, John S. Colton, Richard K. Watt, <i>Brigham Young Unive</i>	<b>183</b> trsity
A New Perspective to Antenna Design for Reconfigurable Wireless Networks Mehedi Hasan, Israfil Bahceci, and Bedri A. Cetiner, Utah State University	195

Throttled Launch-Assist Hybrid Rocket Motor for an Airborne NanoSat Launch Platform Zachary S. Spurrier, Sean D. Walker, Stephen L. Merkley, Steph A. Whitmore, <i>Utah State University</i>	<b>201</b> .en
LETTERS—FOREIGN LANGUAGE, HUMANITIES, AND PHILOSOPHY	
That's Rich: Survival at the Tudor Court Joshua Wiggins, Southern Utah University	225
<b>'Where Has Nabonidus Gone? Where Can He Be?' A</b> <b>Synthesis of the Nabonidus Controversy</b> Spencer C. Woolley, <i>University of Utah</i>	235
LETTERS—LANGUAGE AND LITERATURE	
<b>"Thy Placeless Power": Melville, Mobility, and the Poetics of Placelessness</b> Todd Goddard, <i>Utah Valley University</i>	247
Out-printing the Crown: Richard Carlile's "Leaking" of The Age of Reason during the Regency Period Andrew Doub, Brigham Young University	259
PHYSICAL SCIENCES	
Patterning Supported Lipid Bilayers Using Magnetic Tweezers Travis M. Bulloch, T. Charles Argyle, Madeline Gleave Parson, Christopher F. Monson, <i>Southern Utah University</i>	<b>269</b> and
How Much Less Is More? The Predictive Consequences of Overfitting Brian Knaeble, <sup>1</sup> Bill Bynum, <sup>2</sup> Gano Hasanbegovic, <sup>2</sup> Garret Wilce <sup>1</sup> Utah Valley University, <sup>2</sup> Westminster College	<b>283</b> ox <sup>2</sup> ,
Density Functional Theory Investigation of Polycyclical Peroxide Stability Charles Joseph Simon, Don R. Davies, and H. Laine Berghout, Weber State University	293

# Lyman-α Emission from an Infant Black Hole in the Early Universe

Brandon K. Wiggins,<sup>1,2</sup> Joseph M. Smidt,<sup>2</sup> Jarrett L Johnson<sup>2</sup>, <sup>1</sup>Southern Utah University, <sup>2</sup>Center for Theoretical Astrophysics, Los Alamos National Laboratory

#### SOCIAL SCIENCES

#### Positive Psychology Holistic Determinants, Testosterone Treatment, and Veteran Happiness

327

307

Hannah-Lee Brau, Brigham Young University

#### DISTINGUISHED SERVICE AWARD

The Distinguished Service Award is given to an academic professional for exceptional service to the higher education community in Utah.

#### **Richard Sadler**

Weber State University

Dr. Richard Sadler has been teaching in the history department at Weber State University since 1969, with his research and publications focusing on United States History, Utah History, and Western American History. Additionally, Dr. Sadler was Dean of the College of Social and Behavioral Sciences for 26 years (1985 through 2011) and has served on numerous local and state-wide boards during his long career, including his tenure as president of the Utah School Boards Association. In 2004, he was made a Fellow of the Utah State Historical Society for outstanding historical research and writing and he has received numerous other awards and recognitions for his teaching, research, and contributions to the state of Utah.

#### ACADEMY FELLOW 2017

#### L. Dwight Israelsen

Utah State University

Dr. L. Dwight Israelsen is Professor of Economics, Director of Academic Programs and Initiatives in China, and Director of International Academic Initiatives in the Jon M. Huntsman School of Business at Utah State University, where he has taught for 36 years. He earned a B.A. in Economics from USU, studied Slavic Languages at Harvard, and earned his Ph.D. in Economics from Massachusetts Institute of Technology. He has taught economics for more than 45 years since his first academic appointment at MIT. He also taught and conducted research at Harvard and BYU. Over his career, he has won more than two dozen awards for outstanding teaching and research, including 8 Best Paper Awards from Utah Academy of Sciences, Arts, and Letters. Dr. Israelsen's research covers a wide variety of topics in economics, including comparative economic systems, microeconomic theory, macroeconomic theory and policy, incentives and behavior in economic organizations, economic history of Russia and the American West (including Mormon economic history), econometric theory, and most recently, demographic economics. His research is represented by more than 300 articles, books, and presentations at professional meetings. He has published articles in the top-ranked American Economic Review and seminal articles on the theory of producer cooperatives (Journal of Comparative Economics) and income distribution measurement theory (Advances in Econometrics) among others.

At USU, Professor Israelsen has taught more than 40 courses that span the entire curriculum, from economic principles courses to Ph.D. courses in economic theory. His teaching awards at USU include 8 Teacher of the Year Awards from the Department of Economics and Finance, 4 Teacher of the Year Awards from the Jon M. Huntsman School of Business, and 4 Teaching Excellence Awards from USU. In addition, he received the Professor of the Year Award and the Hall of Fame Award from the President's Leadership Council. He has been very active in international education and has negotiated cooperative agreements between USU and universities in Russia, Ukraine, Thailand, and China. As Director of Academic Programs in China, he supervises USU's Bachelor's Degree in Economics with almost 1,000 students at 4 locations in Mainland China and Hong Kong. In addition to his academic work, Dr. Israelsen has been consultant to a variety of entities, ranging from private firms to municipalities, Indian tribes, non-profit groups, and government agencies.

Professor Israelsen has been active in a number of professional organizations and has found some of his greatest professional and personal satisfaction from his association with the Utah Academy of Sciences, Arts, and Letters, its officers and members. In the 40 years since 1976, he has presented 47 papers at UASAL meetings and has published 18 articles in the Utah Academy journal. In addition, he served as an officer (Chair of Business Division, President, and Member-at-Large) on the UASAL Board for the 20-year period beginning in 1997 and ending this year. Dr. Israelsen and his wife JillAnn have 6 children and 17 grandchildren. His hobbies include hunting, traveling, painting, and writing poetry.

#### O.C. TANNER LECTURE "MY LOVE AFFAIR WITH HISTORY"

### **Richard Sadler**

Weber State University

Richard Sadler is an active historian; he has coauthored books on Ogden City, Weber County, The Weber Basin water history and has written several articles published in academic journals, including the *Utah Historical Quarterly*. He also served as editor of *The Journal of Mormon History*. He is the author and editor of "Weber State College...A Centennial History," published in 1989, which chronicles WSU's first 100 years. In addition to teaching generations of Weber State students about history, he has lived it, and in the process, has added to the institution's story.

He was president of the Utah Academy of Sciences, Arts, & Letters from 1979 to 1981. His publications include: A History of Weber County (Salt Lake City: Utah State Historical Society; Ogden: Weber County Commission, 1997), coathored with Richard C. Roberts; Utah, The Right Place: The Official Centennial History (Salt Lake City: Gibbs Smith, 1996), by Thomas G. Alexander, Richard W. Sadler, general editor, Susan A. Whetstone, photographic editor, Jay M. Haymond, project director; The Weber River Basin: Grass Roots Democracy and Water Development (Logan, Utah: Utah State University Press, 1994), coathored with Richard C. Roberts; and Ogden: Junction City (Northridge, Calif.: Windsor Publications, 1985), coathored with Richard C. Roberts.

#### JOHN & OLGA GARDNER PRIZE

The Gardner Prize is awarded annually for exceptional achievement by an academic professional in Utah.

#### **Randy Silverman**

University of Utah

Randy Silverman, Head of Preservation at the University of Utah's J. Willard Marriott Library since 1993, has worked in the field of book conservation for 38 years. He holds a Masters degree in Library Science from BYU and teaches Preservation as an adjunct faculty member for Emporia State University (KS), Dominican University (IL), and University of Arizona (AZ). He is a founding member of the Western States and Territories Preservation Assistance Service (2007) and teaches disaster preparedness workshops throughout Arizona, Montana, Utah, and Wyoming. He has 86 professional publications, has been involved in disaster recovery since 1986 (including Hurricane Katrina), and has presented professional lectures and workshops in 27 states and 11 countries outside the U.S. He received the American Library Association's Banks Harris Award in 2013 for "outstanding achievement in the field of preservation" and was given a Fulbright Specialists award in 2014 to begin helping the National Library of Uzbekistan improve its preservation standards.

#### HONORARY MEMBER 2017

#### **Diana Major Spencer**

Director, Casino Star Theatre Foundation

With a 25-year academic career spanning forty years and three institutions, Diana Major Spencer always included public service and nonprofit board service among her activities. In 1975, she left the University of Utah to relocate to Sanpete County, unwittingly becoming unemployable because of a combination of both over- and underqualification. Yet her passion for teaching prompted several years of freelance teaching, offering credit through U of U Continuing Education and choosing Shakespeare as her topic to attract locals who might want to visit the Utah Shakespeare Festival (USF) in Cedar City. Diana's newspaper ad triggered an invitation from USF to write an article for the 1979 Festival, which initiated a now 38-year series of 2-3 annual articles and nearly two years of full-time employment (1989-90). Meanwhile, she has served on the following boards: South Sanpete School Board, Sevier Valley ATC, Utah Humanities Council (1 year as Acting Assistant Director during Cynthia Buckingham's 1989 leave), Utah Academy (Chair of Languages and Literature), Utah Alliance for Arts and Humanities Education, Utah Nonprofits Association, and Traditional Building Skills Institute, Snow College, She also served on grant panels for the Utah Arts Council, along with selection committees for the Governor's Awards in the Arts and Utah's Poet Laureate.

Dr. Spencer retired in 2005 from Snow College as Professor of English and Dean of Humanities Emerita after 15 years, during which she upgraded the Convocation program from a random, monthly, lowattendance program to a weekly, for-credit, academically worthy lecture/performance series with 300+ students enrolled. In 2004, she and her collaborator, Gunnison native Lori C. Nay, established a nonprofit to purchase, restore, and supplement the offerings of a 92year-old Beaux Arts-style theatre in Gunnison.

## 2016 BEST PAPER AWARDS

## **Biological Sciences**

# X-ray Analysis of Lead(II) Binding to *Haloferax volcanii* Malate Synthase

Michael Adams and Bruce Howard. Southern Utah University

Business

## **Exploring Shifting Global Labor Management Practices** and Comparative Job and Life Satisfaction

Bergen Eskildsen, Joe Light, Kaitlin Carlisle, and Jonathan H. Westover *Utah ValleyUniversity* 

## Engineering

## Fabrication of Dye Sensitized Solar Cells Using Native and Non-Native Nanocrystals in Ferritin as the Dye

Alessandro Perego, Cameron Olsen, J. Ryan Peterson, Kameron Hansen, John S. Colton, Richard K Watt. *Brigham Young University* 

Letters-Humanities, Philosophy, and Foreign Language

## That's Rich: Survival at the Tudor Court

Joshua Wiggins Southern Utah University

Letters-Language and Literature

## "Thy Placeless Power": Melville, Mobility, And The Poetics Of Placelessness

Todd Goddard Utah Valley University **Physical Sciences** 

# Lyman- $\alpha$ Emission from an Infant Black Hole in the Early Universe

Brandon K. Wiggins, Joseph M. Smidt, Jarrett L Johnson Southern Utah University

## **Chronic Effects of Intermittent Sound Disturbance on Buenos Aires Tetra**

Jeremy L. Arnt, Erica Neisler Weber State University

#### Abstract

The hearing-specialist Buenos Aires tetra (Hyphessobrycon anisitsi) was subjected to two, independent, five-week trials to determine the differences between growth and behavior in intermittently disturbed and quiet habitats. One group was subjected to intermittent acoustic disturbance (135±3 dB rel:  $V/\mu$ Pa) twice per day for one hour. The other was subjected to ambient sound disturbance (85±3 dB rel:  $V/\mu$ Pa). Growth in length, weight, and Fulton's condition factor were compared between groups using repeated-measures ANOVA. Frequency of observation for five categories of behavior was compared using one-tail Student's t tests. Linear regression was used to examine evidence for habituation or sensitization. We found no difference in growth between trials; however, strong differences between three of the five behaviors (nipping, darting, and normal swimming) were observed, showing sensitization (i.e., increased display of stress behaviors) over one-hour periods of disturbance exposure. Sensitization to the onset of disturbance was evident from increased darting over the course of the five-week trials. Subjects showed decreased behavioral responses to the initial stimulus in darting and normal swimming behaviors. These results suggest H. anisitsi is not sufficiently stressed by chronic disturbance to alter growth over a five-week period, but sound does induce behavioral changes on this timescale.

## Introduction

Some animals affected by sound disturbances experience negative fitness consequences (Wright et al. 2007) such as reduced growth, reduced foraging success, increased vulnerability to predators, auditory sensitivity shifts, reduced reproduction via disrupted mate signaling, auditory tissue damage, immunosuppression, and death by intolerable stress hormone levels (Popper and Hastings 2009). Exposure to specific acoustic intensities and frequencies can cause fish to perceive a stressor, initiating a stress response (Popper and Hastings 2009). A stressor is any stimulation that may induce the hypothalamic-pituitaryadrenal (HPA) axis to release stress hormones. These hormones are produced in the hypothalamic region of the brain, which releases signal hormones to the pituitary gland, which releases additional hormones to signal the adrenal gland to release cortisol along with other stress hormones. When subjected to intolerable levels of stress hormones for prolonged periods, the vital organ systems, such as the heart, can begin to shut down, leading to death (Idler et al. 1959). However, vertebrates can adapt to stressors by habituating to the stimuli. Vertebrates habituate by reducing the signal hormones in the hypothalamus of the HPA system reducing prolonged levels of stress hormones in the system. In the case of sensitization, the HPA complex can increase hormone levels, prolonging time spent in stressed conditions. Increased time spent in a stressed state increases the negative consequences of stress.

Fish can habituate to stimuli by varying degrees depending on their sensitivity to the sound stimulus generated and the duration of the disturbance. Fishes have a wide array of adaptations to detect sound. For example, some fishes have simplistic auditory anatomy and detect a limited range of frequencies and sound pressure levels (SPLs) (Smith et al. 2004b), whereas others have Weberian ossicles that amplify vibrations in gas bladders to increase detection sensitivities and ranges (Fay and Popper 2000). With such a varying degree of hearing in fishes, more data from many different species are needed to fully evaluate fish auditory capabilities. In addition, understanding the consequences of acoustic disturbance on fish species' auditory capabilities has become more important because of the increase of anthropogenic sound disturbance in aquatic environments, such as wind energy, pile driving, recreational fishing, and high human dependency on fish for food (Popper and Hastings 2009).

Hearing specialists, such as species with Weberian ossicles, are more sensitive to sound disturbances than hearing-generalists because they detect sounds across wider frequency ranges and can discern relatively weak sounds, which stimulate behavioral responses (Amoser and Ladich 2005). Most behavioral studies on hearing-specialists have focused on mating vocalization and long-range movement behaviors, post-disturbance, which are long-term effects of disturbance (Wardle et al. 2001; Morley et al. 2014). There are relatively few data on immediate stress responses during disturbance periods.

The frequency and duration of a sound can affect fish species differently (Popper and Hastings 2009). Three main categories of anthropogenic sound impact aquatic systems. First, continuous sound disturbances are moderate increases in background sound for long periods of time, such as heavy boat traffic or aquaculture-related sound disturbances from pumps (Slabbekoorn et al. 2010). These time intervals may last from a single day or months (Popper and Hastings 2009). Second, impulsive sounds last only a few seconds, but with high SPLs, such as sonar or mining explosions (Popper and Hastings 2009). Third, intermittent sound disturbance raise background sound levels occur at shorter intervals than constant sounds, but longer than impulsive sounds (Smith et al. 2004a). Some examples of intermittent sound disturbances are harbors where boats arrive and depart discontinuously, with little to no temporal overlap. The SPLs between the three acoustic classifications can range from 84 dB re 1 µPa for ambient intensity levels, to 152 dB re 1 µPa in more intense acoustic soundscapes (Haxel et al. 2013). Unfortunately, intermittent sound is less studied, but is the more likely scenario in the wild, demonstrating a need for further research (Smith et al. 2004a).

Intermittent acoustic disturbance habitats may not allow hearing specialists to physiologically or behaviorally habituate as easily as the other categories of sound and may induce a cycle of stress with alternating periods of disturbance (high sound intensity periods) and physiological recuperation (low sound intensity periods). Typically, animals in stress states do not undergo negative consequences unless the perceived stress is severe or maintained for prolonged periods (Barton 2002). In a habitat disturbed by intermittent sound, hearing-specialists may not have the time to reduce stress hormones to tolerable levels. This cycle could put the fish in a constant cycle of stress-

induced damage and healing, which may chronically alter their behavior or growth.

We performed lab experiments on the hearing-specialist fish Buenos Aires tetra (Hvphessobrvcon anisitsi) in an intermittent sound disturbance environment to determine chronic (i.e., growth and health differences) and acute behavioral changes (frequency of stress behaviors). H. anisitsi was selected for our experiments because of its relatively short life history, acoustic anatomical structures, and individual and schooling behaviors. H. anisitsi are native to the Parana River in South America, which has motorized boat traffic that we are attempting to simulate. They also possess some morphological features associated with a hearing-specialist species, such as Weberian ossicles and have relatively short life cycles of four to five years. In addition, this species shows schooling behaviors that allowed us more statistical data and enabled more behaviors to be observed. We used four behaviors (darting, air gulping, nipping, and freezing) as indicators of stress (Schreck et al. 1997). We hypothesized that H. anisitsi would show increased stress behaviors and reduced growth because fo an inability to habituate. If our hypothesis is supported, intermittent sound disturbance causes: (1) a decrease in non-stressed behavior (normal swimming), (2) an increase in stressed behaviors (darting, nipping, freezing, and air gulping) chronically and throughout sound disturbance periods, and (3) a reduction in growth in the Buenos Aires Tetra (H. anisitsi).

## **Materials and Methods**

## Husbandry/experimental set-up

Two sets of trials were performed: an ambient control trial  $(85\pm3)$  dB rel: 1 V/µPa) and a high intensity treatment trial  $(135\pm3)$  dB rel: V/µPa). The duration of each trial was 5 weeks. Test subjects varying in age from 6 to 9 months were purchased from Petco and brought directly to the laboratory. Males and females were randomly assorted into tanks to eliminate effects of sexual dimorphism in growth. Prior to each trial start date; fish had 24 hours to acclimate to their tanks, which gave ample time for test subjects to acclimatize (Artigas et al. 2005). Low-and high-intensity trial runs were conducted using 30 test subjects divided equally among five tanks. After the first 30 control test subjects had finished the trial, 30 different fish were purchased for the experimental trial and, again, divided equally among tanks. Six individuals were used in each tank to observe additional behaviors such as mate nipping, which is not observable without multiple individuals. In the wild, *H. anisitsi* exhibits schooling behavior in the wild. To eliminate

potential stress from lack of school mates, we provided six test subjects per tank

Five, 20-gallon tanks were set for a 12:12 hour day/night cycle with fluorescent lighting lasting from 0700 to 1900. All tanks were kept at 25-27°C with a submerged 50-watt tank heater. Tanks were filtered and pumped using an Aqeon Quietflow 10 system to reduce sound and clean the tanks. Filters were changed on the first day of week 3 for both trials. Each tank had gravel substrate and was visually blocked on three sides to reduce external visual stimulation. Tanks were insulated from reverberation with foam padding along the base of the tanks. In addition, each tank had an individual exterior speaker set to the back side for sound projection. The speaker was placed in the same position in the control and treatment trials; however, the control trial did not proiect sound.

Test subjects were observed twice daily for behaviors. Two tanks were manually observed daily, 0700-0800 and 1700-1800. To reduce the effect of the observer on the test subject, a single view port on the front of the tanks was used for data collection. Commercial flake food was provided 15 to 20 minutes prior to observation periods in 0.15  $\pm$ 0.001 g portions per tank. The reason for providing food at this time was to reduce the occurrence of aggressive or stress behaviors associated with feeding. For behavioral analysis, one random tank was selected for observation each observational period with a random number generator (www.random.org). During each observational period, the following behaviors were examined at two-minute intervals: normal swimming, darting, nipping, freezing, and air gulping (Altmann 1974; Schreck et al. 1997; Table 1).

tetras (Hyphessobrycon anisitsi) every two minutes during 60-			
minute observation periods.			
Behavior	Description		
Normal Swimming	All fish swim < 3 body lengths $\cdot$ s <sup>-1</sup> for $\geq$ 5 s		
Darting	Fish swim > 3 body lengths $\cdot$ s <sup>-1</sup> for > 1 s,		
	but not at another fish		
Nipping	Fish accelerate toward and physically		
	contacted and/or bite another individual.		
Freezing	Fish shows no operculum or fin movement		
	for $> 1$ s		
Air Gulping	Fish move within 1 cm of water surface		
	with repetitious mouth-sucking motion for		
	> 1 s		

Once per week, mass and fork length of each fish were measured in all five tanks, after anesthetizing fish in a clove oil solution (IACUC protocol number 15-01; Mitjana et al. 2014). All 30 test subjects were weighed and measured on day 7 of each week after the 1700–1800 observation periods. To minimize the stress effects of handling, we provided each tank a commercial stress coat water conditioner after all fish were measured (API Stress Coat Aquarium Water Conditioner). Mean growth of individuals from all tanks combined was calculated for each week for each treatment. Fulton condition factor  $K = 100 {W \choose 15}$  was used to determine average condition of individuals each week (Fulton 1902;

Ricker 1975).

#### Sound projection and measurement

The audio recording used for the simulation of boating sound was a 1-hour digital recording of a QSLA9 Cummings diesel engine (281– 405 hp) on a 28' single, round bottom, fiberglass hull vessel cycling and idling on a hydrophone in Lake Conroe, TX. Sound frequencies ranged from 0 Hz to 12 KHz (Motorola R2600 Communication System Analyzer). These sounds were downloaded and stored in a computer, then projected through a 5.1 speaker system (VM Audio EXMS581W 1000 Watt system) for 1 hour during high-intensity trial observation periods. Speakers were set centered on the back side of each tank 5 cm away from the glass to reduce vibrations on the tank itself. Each speaker was capable of producing frequency ranges from 35Hz to 20 KHz, with a total power output of 370 RMS watts for the entire system.

Prior to introduction of fish into the experimental tanks, sound measurements were made using the H1a hydrophone (Aquarian Audio Products, Inc., Anacortes, WA, sensitivity = -190dB re:  $1V/\mu$ Pa, frequency range = 1 Hz to 100 KHz). Mean, maximum, and minimum SPL for the tanks were measured 2 cm above the bottom four corners of the tank, as well as the center and 2 cm below the water line of the top of the tank (Table 2). Measurements were made using a RMS voltmeter measurement on the hydrophone and calculations were made according to broadband spectrum analysis

$$\left(\text{SPL} \left(\text{dB re 1 } \mu\text{Pa RMS}\right) = 20 \log \left[10\left(\left(\frac{X \times 103}{\text{HCV}}\right) \times 106\right)\right)\right]$$

(Davidson et al. 2007). In this equation, SPL = Sound pressure level, RMS = Root of mean squared, X = RMS reading, HCV = Hydrophone calibration value.

Table 2. Sound pressure levels in dB rel: 1 V/µPa of the hydrophone measurements prior to the introduction of H. anisitsi in both the high-intensity (HI) and low-intensity (LI) trials (HI) Tank Mean ± SD Minimum Maximum Number (SPL) (SPL) (SPL)  $114.88 \pm 4.96$ 122.45 1 108.01 2  $115.00 \pm 4.03$ 106.33 122.36 3  $115.75 \pm 5.19$ 109.48 125.84 4  $114.88 \pm 4.40$ 107.93 125.23 5  $115.75 \pm 4.12$ 111.26 124.33 Mean ± SD (LI) Tank Minimum Maximum (SPL) Number (SPL) (SPL)  $85.21 \pm 0.61$ 1 84.32 86.38 2  $85.30 \pm 0.50$ 84.62 86.27 3  $85.31 \pm 0.623$ 84.32 86.64 4  $85.47 \pm 0.72$ 84.55 86.01 5  $85.19 \pm 0.82$ 83.92 87.08

## Statistical analyses

Analyses of behavioral observations were conducted using a single-tailed, independent *t*-test to compare mean percentage of each behavioral category documented at each two-minute interval. Simple linear regressions for behavior categories were used to examine the effect of time, in minutes within the observation period, on the percentage of time individual behavior was displayed (Dupont and Plummer 1998).

ANOVA with repeated measures (week as a covariate) was used to compare fork length, weight, and Fulton condition factor between treatments (Meekan et al. 2003). Simple linear regressions of slope versus the trial number and of the y-intercept of within-trial frequency versus time were used to assess habituation/sensitization within and among trials, respectively, over the trial period. Y-intercepts were used to represent the initial reaction to the acoustic stimulus showing an increase, decrease, or static maintenance over the trial period to determine sensitization, habituation, or no effect. All statistical analyses were performed in the SPSS (v 23, Manufacturer IBM) statistics program. All *p*-values were based on one-tailed tests with an alpha level of 0.05.

## Results

#### Survival

Both trials had several deaths. Three of the 30 test subjects in the low-intensity group died after infection with *Ichthyophthirius multi-filiis*, a common aquarium parasite, within a single tank. In noisy trials, three tanks had deaths. Within the three tanks, four individuals had their caudal fins nipped off. All fish corpses were found prior to morning observations and removed immediately after discovery.

## Length, weight, and Fulton's condition factor

Fork length differed between low-intensity (mean  $\pm$  SD 4.1  $\pm$  0.44 cm) and high-intensity (3.6  $\pm$  0.33 cm) trials, but duration period had no effect on this comparison (Table 3). Similarly, weight differed between low- (1.3  $\pm$  0.35 g) and high-intensity (1.0  $\pm$  0.22 g) trials, but duration period had no effect. Fulton condition factor did not differ between low- (2.0  $\pm$  0.21) and high-intensity (2.0  $\pm$  0.22) trials.

Table 3. Results of ANOVA with repeated measures of fork length (FL), weight, and Fulton's condition factor (K) of <i>H</i> . <i>anicitei</i> subjected to low intensity versus high intensity habitats						
Source df MS F p						
FL	1	1.00	17	< 0.001*		
$FL \times Week$	1	< 0.01	< 0.01	0.979		
Between-subject	28	0.12				
error						
Within-subject error	28	0.06				
Weight	1	0.74	17	<0.001*		
Weight × Week	1	< 0.01	0.02	0.896		
Between-subject	28	0.10				
error						
Within-subject error	28	0.05				
K	1	< 0.01	0.24	0.631		
$K \times Week$	1	0.01	0.34	0.567		
Between-subject	28	0.05				
error						
Within-subject error	28	0.03				

Measurements were conducted weekly on each individual fish (30 fish per treatment).

## Behavior frequencies

Mean frequency of observation varied between treatments for three of the five behaviors: nipping, darting, and normal swimming (Table 4). Nipping (p<0.001) and darting (p<0.001) were more frequent in noisy trials and increased linearly in frequency during the time exposed, whereas there was little nipping in low-intensity trials (Figure 1). Accordingly, normal swimming (p<0.001) was significantly reduced during observation periods in high-intensity trials, whereas in low-intensity trials, normal swimming was the predominant behavior at all periods (Figure 1). Rates of air gulping and freezing behaviors did were not significantly different between the noisy and control trial during each observation period (Table 4).

Table 4. Results of single-tailed sample independent t-tests for
mean frequency and standard of deviation (SD) of five observed
behaviors measured at two-minute intervals during low-
intensity (LI) versus high-intensity (HI) treatments

intensity (EI) versus ingli-intensity (III) ir eatments					
Behavior	Mean ± SD	df	Р	Т	
(LI) Normal	$69 \pm 0.5$	58	< 0.001	12.0	
Swimming					
(HI) Normal	$46 \pm 10.0$				
Swimming					
(LI) Darting	$9 \pm 2.7$	58	< 0.001	-29.0	
(HI) Darting	$59 \pm 9.0$				
(LI) Nipping	$6 \pm 2.2$	58	< 0.001	-15.0	
(HI) Nipping	$50 \pm 15.0$				
(LI) Air Gulping	$1 \pm 1.0$	58	0.260	-1.1	
(HI) Air Gulping	$2 \pm 0.8$				
(LI) Freezing	$1 \pm 1.9$	58	0.050	-2.0	
(HI) Freezing	$2 \pm 1.3$				

#### Behavior habituation

The increase of regression line steepness of positive within-trial frequencies of nipping and darting behaviors over time, with concurrent decreases in frequencies of normal swimming, suggested individuals were increasingly sensitized to the sound disturbance (Figure 2). The y-intercepts regressions showed a decrease of reaction to initial behavioral stimulus to normal swimming and darting behaviors, suggesting habituation, with no significant change in nipping behavior (Figure 3).



**Figure 1.** Simple linear regression of cumulative (all samples combined) frequency of three behaviors darting (A), nipping (B), and normal swimming (C) observed in Buenos Aires tetras (*H. anisitsi*) (n=30) every two minutes during 60-minute observation periods during low-intensity (LI) and high-intensity (HI) treatments.



Sound Disturbance Effects in Tetra Fish 29

**Figure 2.** Slope of regression line values of only individual high intensity trials throughout the five-week study showing three behaviors of Buenos Aires tetras (*Hyphessobrycon anisitsi*): darting (A), nipping (B), and normal swimming (C). Results of simple linear regression are provided. N=70



**Figure 3.** Simple linear regression of Y-intercept values of each trials regression line over the entire five-week high-intensity trials showing three behaviors of Buenos Aires tetras (*H. anisitsi*) (n=70): darting (A), nipping (B), and normal swimming (C). Results of simple linear regression are provided. Y-intercept values were used to illustrate the stimulus level of test subjects according to the onset of sound stimuli.

### Discussion

Comparisons of length, weight, and Fulton condition factors between treatments indicated that the intermittent sound disturbance did not affect *H. anisitsi* growth (Table 3). Differences of fork length and weight can be attributed to initial differences in fish size between trials. The absence of growth effect may indicate an ability to heal sound damage between disturbances or, more simply, the sound had no effect. SPL may not have been high enough to induce physiological damage. In this situation, the typical level of a single diesel engine's boating sound (135 dB rel  $1V/\mu$ Pa) may not be intense enough to damage the hearing tissue (Wysocki and Ladich 2005). However, the Parana River, to which *H. anisitsi* is native, may encounter multiple engine boats or larger boats with more powerful engines, which creates higher SPL (Paini et al. 2009).

*H. anisitsi* growth may be affected by increased stress hormone levels exacerbated by louder sound. Metabolic process such as oxygen transport and utilization, glycogen transport, and immunosuppression are affected if an individual is exposed to higher acoustic intensities, especially if the exposure period is prolonged (Bonga 1997). If metabolic processes are further affected by more intense sound, then resources normally used for growth may be repurposed to mitigate sound effects and not utilized for growth.

Feeding behavior of H. anisitsi may have played a role in growth of the test subjects (Selye 1946). Our fish were fed *ad lib* and may have compensated for metabolic stress affects by consuming more calories. However, outside of the laboratory, fish are not necessarily able to forage enough to compensate for stress responses. In the wild, if an individual cannot forage optimally to compensate for stress, it would likely show negative growth consequences of sound exposure.

Another potential reason for the lack of growth consequences may be the sound frequencies projected by the diesel engine. Our study used a broad-spectrum analysis of many frequencies to determine SPL. Broad-spectrum sound detection takes all frequencies within the hydrophone's range to determine the SPL; however, specific frequencies can have individual SPL within a broad spectrum. These individual frequencies may not harm *H. anisitsi* at the intensities created by boats (Wysocki and Ladich 2005). Other hearing specialists are stimulated by specific sound frequencies, such as *Carassius auratus* in the range of 0.1–4.0 KHz and *Clupea harengus* above 4.0 KHz (Yan et al. 2000; Kastelein et al. 2008). These frequency ranges are small compared with those tested in our broad-spectrum analyses (0–25 KHz). In addition, fish can adapt to acoustic environments by shifting or modifying their ranges to move outside of those ranges that are overstimulated (Popper and Hastings 2009). However, the hearing range and generated sound did overlap because we did observe a behavioral response to increased noise stimulus, indicating at least a partial overlap of ranges. To further understand this aspect of hearing in *H. anisitsi*, future research should seek to describe hearing frequency ranges and narrow spectrum SPL of tests (Kastelein et al. 2008).

A pronounced behavioral response to sound from H. anisitsi suggested increased sensitivity over the five-week period; however, the stimulus levels of the onset of disturbance of behavior to normal swimming and darting showed a habituation with a decrease of initial darting as the experiment progressed. There was no change in nipping, which shows that as the experiment progressed, the initial level of nipping at the onset of sound projection remained strong while nipping throughout each trial was increasingly severe. In other words, fishes appeared to be increasingly sensitive to the intermittent sound, causing increased aggression (i.e., nipping). The increase of nipping over the single hour is particularly important when considering mortalities. Mortalities in the low-intensity group appeared to be due to an I. multifiliis infection in a single tank. This common aquarium parasite can cause death in healthy test subjects (Matthews 2005). In contrast, mortalities in the high-intensity group appeared to be associated with increased, sound-induced nipping behavior because the caudal fins were nipped down to the peduncle, which may have hindered feeding and increased social inferiority, while elevating individual stress levels. If this is the case, then H. anisitsi may be more vulnerable to parasites like I. multifiliis, when present, because they invade hosts through epithelial cells to invade mucous cells (Buchmann and Nielsen 1999). If these cells were damaged in some way by nipping they may be more susceptible to infection due to stress-induced immunosuppression or removal of physical barriers such as cell membrane (Ewing et al. 1985).

Contrary to laboratory test subjects, wild individuals may have the option to move away from stressful sounds. Moving away from the sound disturbance may mitigate the effects of sound stress; however, if optimal habitats are restricted to sound-disturbed waters, fish will be forced to either contend with stress or occupy suboptimal habitats.

Evidence for sensitization over time supports our initial hypothesis that intermittent sound may not allow for habituation. There may be several possibilities to explain reasons for the increase of stressed behavior over time. Fish may not be able to habituate to their acoustic stimuli because they do not have enough time between intermittent disturbances to do so. Additionally, the aggressive nipping behavior may increase stress levels. Some fish within the school would be the victim of nipping, which may have an additive effect to the victims. Although the increase of nipping in our study did not affect growth rates, the duration of our individual trials was only five weeks, which may not have been sufficient to see growth changes. Alternatively, death occurs prior to onset of growth effects.

The most puzzling question derived from our study is the apparent change in behavior without effects on growth. This is especially perplexing considering the intricate cause-and-effect relationships of physiology and behavior (Schreck et al. 1997). One explanation may be Hans Selye's General Adaptation Syndrome (GAS) (Selye 1946). If the body stays in the alarm stage of Selye's GAS before returning to homeostasis, this would result in a physiological response that may change behavior but not alter growth rates. The alarm stage is characterized by releasing hormones such as adrenaline, noradrenaline, and cortisol. These hormones affect behavior in the short term, but can be removed from the body fairly quickly, from minutes to hours (Schreck et al. 1997). If the sound disturbance only progresses to the alarm stage in *H. anisitsi*, but not to resistance or exhaustion stages, it would not be expected to affect growth.

If the GAS is the reason for strong behavioral and no growth response in our study, the ability of *H. anisitsi* to alter stress hormone effects becomes increasingly important. This study shows an increasing sensitivity of stress behavior over time. Since behavior and stress physiology are so interconnected, it is reasonable to postulate that stress hormone levels would increase along with the increasing stress behaviors. Future studies should seek to evaluate hormone levels to see whether they correlate with the behavioral responses.

#### Acknowledgments

I am grateful for the support by the Ralph Nye Charitable Foundation. In addition, I would like to thank Nicholas DeYoung and Justin Easley for their assistance with data collection. Finally, I appreciated the guidance of Christopher Hoagstrom, my undergraduate thesis advisor, and Aaron Ashley and John Mull, my thesis committee members. Experiments all complied with Weber State University's Animal Care and Use Committee's guidelines.

#### References

Altmann, J., (1974). Observational study of behavior: sampling methods. *Behaviour* 49, 227-267.

Amoser, S., Ladich, F., (2005). Are hearing sensitivities of freshwater fish adapted to the ambient noise in their habitat? *J. Exp. Biol.* 208, 3533-3542.

Artigas, M.L., Skjaeraasen, J.E., Utne-Palm, A.C., Nilsen, T., (2005) Recovery from handling stress in *Gadus morhua*. J. Fish Biol. 67 (2) 384-391.

Barton, B.A., (2002). Stress in fishes: a diversity of responses with particular reference to changes in circulating corticosteroids. *Integr. Comp. Biol.* 42(3), 517-525.

Bonga, W.S.E., (1997). The stress response in fish. *Physiol. Rev.* 77 (3) 591-625.

Buchmann, K., Nielsen, M.E., (1999). Chemoattraction of *Ichthyophtirius multifiliis* (Ciliophora) to host molecules. *Int. J. Parasitol.* 29, 1415-1423.

Davidson, J., Frankel. A.S., Ellison, W.T., Summerfelt, S. Popper, A.N., Mazik, P., Bebak, J., (2007). Minimizing noise in fiberglass aquaculture tanks: Noise reduction potential of various retrofits. *Aquacult. Eng.* 37, 125-131.

Dupont, W.D., Plummer, W. D., (1998). Power and sample size calculations for studies involving linear regression. *Control. Clin. Trials*, 19 (6), 589-601.

Ewing, M.S., Kocan, K.M., Ewing, S.A., (1985). *Ichthyophthirius multifiliis*, (Ciliophora) invasion of gill epithelium. *J. Protozool.* 32, 305-310.

Fay, R.R., Popper, A.N., (2000). Evolution of hearing in vertebrates: the inner ears and processing. *Hear. Res.* 149, 1-10.

Fulton, T., (1902). Rate of growth of seas fishes. Sci. Invest. Fish. Div. Scot. Rept. 20: 1035-1039.

Haxel, J.H., Dziak, R.P., & Matsumoto, H., (2013). Observations of shallow water marine ambient sound: the low frequency underwater soundscape of the central Oregon coast. *J. Acoust. Soc. Am.* 133(5), 2586-2596.

Idler, D.R., Ronald, A.P., & Schmidt, P.J., (1959). Biochemical Studies On Sockeye Salmon During Spawning Migration: Vii. Steroid Hormones In Plasma. *Biochem. Cell Biol.* 37 (10) 227-238.

Kastelein, R.A., Heul, S.V.D., Verboom, W.C., Jennings, N., Veen, J.V.D., Haan, D.D., (2008). Startle response of captive North Sea fish species to underwater tones between 0.1 and 64 kHz, *Mar. Environ. Res.* 65(5), 369-377.

Matthews, R.A., (2005). *Ichthyophthirius multifiliis* Fouquet and ichthyophthiriosis in Freshwater teleosts. *Adv. Parasitol.* 59, 159-241.

Meekan, M.G., Carlton, J.H., McKinnon, A.D., Flynn, K., Furnas, M., (2003) What determines the growth of tropical reef fish larvae in the plankton: food or temperature? *Mar. Ecol. Prog. Ser.*, 256, 193-204

Mitjana, O., Bonastre, C., Insua, D., Falceto, M.V., Esteban, J., Josa, A., Espinosa, E., (2014) The efficacy and effect of repeated exposure to 2-phenoxyethanol, clove oil and tricaine methanesulphonate anesthetic agents on juvenile Angelfish (*Pterophyllum scalare*). *Aquaculture*, 433, 491-495.

Morley, E.L., Jones, G., Radford, A.N., (2014). The importance of invertebrates when considering the impacts of anthropogenic noise. *Proc. Roy. Soc.* B, 81, 2013-2683.

Paini, M.C., Morata, T.C., Corteletti, L.J., Albizu, E., Marques, J.M., Santos, L., (2009). Audiologial findings among workers from Brazilian small-scale fisheries. *Ear Hearing*, 30(1) 8-15.

Popper, A.N., Hastings, M.C., (2009). The effects of human-generated sound on fish. *Integr. Zool.* 4, 43-52.

Ricker, W.E. (1975). Computation and interpretation of biological statistics of fish populations. *Bull. Fish. Res. Board Can.* 191, 1-382.

Schreck, C.B., Olla, B.L., Davis, M.W. (1997). Behavioral responses to stress. *Fish Stress Health Aquacult*, 62, 145-170.

Selye, H. (1946). The general adaptation syndrome and the diseases of adaptation 1. J. Clin. Endocrinol. Metab. 6(2), 117-230.

36 Biological Sciences

Slabbekoorn, H., Bouton, N., Opzeeland, I.V., Coers, A., Cate, C.T., Popper, A.N., (2010). A noisy spring: The impacts of globally rising underwater sound levels on fish. *Trends Ecol. Evol.* 25, 419-427.

Smith, M.E., Kane, A.S., Popper, A.N. (2004a). Noise induced stress response and hearing loss in goldfish (*Carassius auratus*) J. Exp. Biol. 207, 427-435.

Smith, M.E., Kane, A.S., Popper, A.N. (2004b). Acoustical stress and hearing sensitivity in fishes: does the linear threshold shift hypothesis hold water? *J. Exp. Biol.* 207(20), 3591-3602.

Wardle, C.S., Carter, T.J., Urquhart, G.C., Johnstone, A.D.F., Ziolkowski, A.M., Hampson, G., Mackie, D., (2001). Effects of seismic air guns on marine fish. *Cont. Shelf Res.* 21, 1005-27.

Wright A.J, Soto N.A, Baldwin M., Beale C.M. et al. (2007) Anthropogenic noise is a stressor in animals: a multidisciplinary perspective. *Int J. Comp. Physiol.* 20:250-273.

Wysocki, L.E., & Ladich F. (2005). Hearing in fishes under noise conditions. *JARO* 6(1), 28-36.

Yan, H.Y., Fine, L., Horn, N.S., Colon, W.E., (2000). Variability in the role of the gasbladder in fish audition. *J. Comp. Physiol.* 186, 435-445.
# Student's t test and scientific reproducibility

Brian Knaeble,<sup>1</sup> Julian Chan,<sup>2</sup> Russell Costa,<sup>3</sup> Katie McLean,<sup>3</sup> Douglas Getty,<sup>3</sup> Riddhi Rampeearee<sup>3</sup> <sup>1</sup>Utah Valley University; <sup>2</sup>Weber State University;

<sup>3</sup>Westminster College

#### Abstract

Many scientific results cannot be reproduced. Although there are many reasons for irreproducibility in science, here we focus on the influence of a statistical procedure that is less than optimal. For simple inference about a population mean, the one-sided t test is less powerful than the optimal likelihood-ratio test, assuming independent and normally distributed observations. Here we conduct simulations to determine whether this power discrepancy is of practical significance. We estimate false positive rates for both tests under various assumptions for alternative parameters and use Bayes' rule to evaluate implications for scientific reproducibility. We assess the sensitivity of our results to departures from independence and normality. Results are tabulated for reference by scientists concerned about reproducible inference for a population mean.

# Introduction

The philosopher of science Popper (1934, p. 66) has written "nonreproducible single occurrences are of no significance to science... We shall take a theory as falsified only if we discover a reproducible effect which refutes the theory." Scientific experiments are often repeated, and when there is agreement between tests or measurements by the same observer in the same laboratory under the same conditions the results are said to be repeatable (Vaux, Fidler, & Cumming, 2012), while they are said to be reproducible when there is agreement between independent experiments performed at different times or places (Casadevall & Fang, 2010). A perceived crisis of reproducibility has been of recent concern to researchers in the biomedical and behavioral sciences. For example, Begley and Ellis (2012) describe a situation where 47 of 53 medical research papers focusing on pharmacological cancer research were not reproducible, and similar problems have been argued to exist more generally in pharmacology (e.g., Prinz, Schlange, & Asadullah, 2011) and other biomedical research areas (e.g., Collins & Tabek, 2014), as well as in psychology (e.g. Bakker, van Dijk, & Wicherts, 2012; Open Science Collaboration, 2015; Pashler & Wagenmakers, 2012) and other related disciplines.

Increasingly, scientists are calling for analyses of and remedies for this perceived replication crisis to avert their fields of inquiry from being undermined (e.g., Ioannidis et al., 2001; Ioannidis, 2005; Schooler, 2014; Steckler, 2015). Some explanations for the difficulty of replication focus on questionable research practices (see Pashler & Wagenmakers, 2012). While scientific misconduct such as plagiarism and fabrication are part of the problem, there are other less blatantly fraudulent practices in the research, review, and publication process such as the neglect of negative results, inappropriate use of research design and statistics, peer review bias, careless research practices, and honest error that also likely contribute. Here we investigate the extent to which routine (Gigerenzer, 2004) use of common statistical tests with less than optimal power contributes to irreproducibility.

We focus exclusively on one-sample, one-sided tests for means. Assuming normality, the null and alternative hypotheses take the form  $H_0 \mid \sigma = \sigma_0$  and  $H_a \mid \sigma = \sigma_a$ . Without loss of generality, we can as-

 $\mu_0 = \sigma_0$  and  $\mu_0 = \sigma_a$ . Without loss of generality, we can assume that  $\mu_0 = 0$  and  $\sigma_0 = 1$  (see Appendix, Lemma 1). When  $\mu_0 = 1$  and  $\sigma_0 = 1$ , the positive one-sided z test is most powerful (Hogg, McKean, & Craig, 2013, Example 8.1.2). Given  $\mu_0 \neq 0$ , we can select an appropriate one-sided z test that is most powerful (see Appendix, Theorem 1) as long as  $\sigma_0 = 1$ , and when  $\sigma_a \neq 1$ , at test

can be used, but remarkably the t test is not most powerful. For background reading on the t test see Moore, Notz, and Fligner (2015). The rejection region of the t test is a conical subset of the sample space. For any fixed ( $\mu_a, \sigma_a \neq 1$ ) the Neyman-Pearson lemma (Wackerly, Mendenhall & Scheaffer, p 542) specifies an optimal rejection region bounded by a sphere centered at (**b**, **b**, ..., **b**) where  $\mathbf{b} = \frac{\sigma_0(\mu_0 - \mu_a)}{\sigma_a^2 - \sigma_0^2}$ , and we call this most-powerful test the ball test (see Appendix, Theorem 1). In the appendix, we derive the ball test from the likelihood-ratio test. We seek to determine how much power is gained when switching from the t test to the ball test in various situations, while assessing implications for scientific reproducibility. The details of our methods are described in Section 2. The results of simulations are described in Section 3. Section 4 describes the results of sensitivity analyses, and a discussion occurs in Section 5.

#### **Materials and Methods**

Size refers to the probability of rejecting the null hypothesis given that the null hypothesis is indeed true. Size is the predetermined risk level for a type 1 error. We write  $\mathbf{a}_t$  for the size of t tests and  $\mathbf{a}_b$  for the size of ball tests. The radius of the ball is denoted with r, and  $\mathbf{a}_b$  is a one-to-one, monotonic function of r. Power refers to the probability of rejecting the null hypothesis given that the alternative hypothesis is true. One minus power gives the probability for a type 2 error. We write  $\mathbf{a}_t$  for the power of a t test and  $\mathbf{a}_b$  for the power of a ball test. We write p for the prior probability of  $\mathbf{H}_0$ . Note that the prior probability of  $\mathbf{H}_a$  is 1-p. Using Bayes's rule (see Figure 1), we can compute posterior probabilities for the false positive rate

$$P(H_0|roj) = \frac{p\alpha}{p\alpha + (1-p)m},$$

the false negative rate

$$P(H_{g}|\text{nrej}) = \frac{(1-p)(1-\pi)}{p(1-\alpha) + (1-p)(1-\pi)},$$

the alternative reproducibility rate

$$P(H_{a} | rej) = \frac{(1-p)\pi}{p\alpha + (1-p)\pi},$$

and the null reproducibility rate

$$P(H_0|nrej) = \frac{p(1-\alpha)}{p(1-\alpha) + (1-p)(1-\pi)}.$$



**Figure 3.** The posterior probabilities are obtained from the prior probabilities p and (1 - p) along with the conditional probabilities  $(1 - \alpha)$ ,  $\alpha$ ,  $(1 - \pi)$ , and  $\pi$  using Bayes' rule.

Simulations were carried out with R (in the appendix we note a connection to the noncentral chi-squared distribution). The computer code is provided in the appendix for those who wish to run similar simulations with different parameter values. We provide seed values for those who wish to validate our results. Seed values were selected using a random algorithm based on the letters of last names of the authors. Iterative use of Monte Carlo approximation was used to find r values for ball tests of a specified size. Given a ball test or t test, Monte Carlo methods were used to approximate  $\boldsymbol{a}_{\boldsymbol{b}}, \boldsymbol{\pi}_{\boldsymbol{b}}$ , and  $\boldsymbol{\pi}_{\boldsymbol{t}}$ . In Section 3, are was set by R. For sensitivity analysis in Section 4, nominal test size was .05, simulating a situation where researchers are proceeding under the (mistaken) assumption that the size of their test is .05. Monte Carlo methods were used to determine actual values for  $\alpha_t$  and  $\alpha_b$ , and these values are displayed in the tables of Section 4. In Section 3, the results are tabulated by sample size (n=10,25,100), alternative mean ( $\mu_a=3$ , 1,1/3,1/10), and alternative standard deviation ( $\sigma_{\alpha}=2,1/2$ ). The sample size is set at 25 throughout Section 4. We speak of  $\mu_a$  as the effect size.

#### Results

Large differences in power between the ball test and the t test were observed for small effect sizes. Corresponding improvement in reproducibility can be expected for sample sizes up to 100. The greatest improvement occurs with smaller samples. The improvement is slightly greater when the alternative has larger variance. Tables 1–4 indicate that there is more room for improvement with null reproducibility  $P(H_0|nre|)$  than with alternative reproducibility  $P(H_{a}|rej)$ . For large effect sizes (relative to the null standard deviation), the t test and ball test perform comparably.

Table	Table 1: Simulation results for small effect size of 1/10									
Simulation Parameters			Pov	wer	$P(H_{a})$	rej)	$P(H_0   nre )$			
Seed	r	n	σα	π	$\pi_{b}$	t	b	t	b	
1519	4.28	10	2	6.72	91.9	57.3	94.8	50.5	92.1	
38	2	10	.5	14.5	90.0	74.3	94.7	52.6	90.4	
6.14	513	25	2	8.00	99.8	61.6	95.2	50.8	99.8	
1411	3.85	25	.5	25.0	100	83.3	95.2	55.9	100	
1521	11.2	100	2	12.5	100	71.5	95.2	52.1	100	
1821	8.91	100	.5	63.4	100	96.3	95.2	72.2	100	

Table	Table 2: Simulation results for moderately small effect size of 1/3									
<b>Simulation Parameters</b>				Power		$P(H_o   ref)$		$P(H_0   nre )$		
Seed	r	n	$\sigma_a$	π	$\pi_{b}$	t	b	t	b	
123	4.31	10	2	12.4	92.7	71.3	94.9	52.0	92.8	
114	2.18	10	.5	61.8	94.8	92.5	95.0	71.3	94.8	
209	6.17	25	2	20.2	99.8	80.1	95.2	54.3	99.8	
91	4.2	25	.5	94.4	100	95.0	95.2	94.5	100	
144	11.2	100	2	50.4	100	91.0	95.2	65.7	100	
125	9.68	100	.5	100	100	95.2	95.2	100	100	

Table	Table 3: Simulation results for medium effect size of 1									
Simulation Parameters			Power		P(H <sub>a</sub>	$P(\mathbf{H}_a   r \epsilon j)$		nrej)		
Seed	R	n	σα	$\pi_t$	$\pi_{b}$	t	b	t	b	
512	4.51	10	2	42.8	97.2	89.5	95.1	62.4	97.1	
129	3.72	10	.5	100	100	95.2	95.2	100	100	
111	6.46	25	2	78.3	100	94.0	95.2	81.4	100	
218	6.8	25	.5	100	100	95.2	95.2	100	100	
51	11.8	100	2	100	100	95.2	95.2	100	100	
52	15.2	100	.5	100	100	95.2	95.2	100	100	

Table 4: Simulation results for moderately large effect size of 3									
Simulation Parameters			Power		$P(\mathbf{H}_{a} rej)$		$P(H_0   nrej)$		
Seed	R	n	σa	$\pi_t$	$\pi_{b}$	t	b	t	b
1919	5.84	10	2	99.7	100	95.2	95.2	99.6	100
1021	11.4	10	.5	100	100	95.2	95.2	100	100
201	8.46	25	2	100	100	95.2	95.2	100	100
115	19.0	25	.5	100	100	95.2	95.2	100	100
312	15.6	100	2	100	100	95.2	95.2	100	100
141	39.6	100	.5	100	100	95.2	95.2	100	100

#### Sensitivity analysis

The results of the previous section show that power and therefore reproducibility can be improved by using the ball test in place of the t test, but the results depend on the assumption of normality and precise

For this reason, we have assessed the sensitivity of these results to departures from normality, misspecification of  $\mu_{\alpha}$ , and misspecification of  $\sigma_{\alpha}$ . In place of normality, we use scaled and shifted t distributions having three degrees of freedom:  $\frac{\iota(\mathfrak{A})}{\sqrt{\mathfrak{A}}}$  under the null hypothesis and  $\mu_{\mathfrak{a}} + \sigma_{\mathfrak{a}} \frac{\mathfrak{t}(\mathfrak{A})}{\sqrt{\mathfrak{A}}}$  under the alternative hypothesis. Tables 5–9 show how with mistaken assumptions the ball test (despite sensitivity of its actual size to departures from normality) still usually outperforms the t test, especially when the effect size  $\mu_{\mathfrak{A}}$  is small.

Table 5: Results of sensitivity analysis with n=25 when the distributions are actually scaled and shifted t distributions										
Sim Para	on ers	Actual Size		Actual Power		$P(\mathbb{H}_a rej)$		P(H <sub>0</sub>  nrej)		
Seed	μ	σα	α <sub>t</sub>	α <sub>b</sub>	π <sub>t</sub>	$\pi_b$	t	b	t	b
189	1/3	2	.05	.11	25.2	93.1	56.0	92.7	84.1	89.1
520	1/3	1/2	.05	.27	94.0	96.8	94.1	95.8	95.1	77.9
2025	3	2	.05	.09	100	100	99.7	100	95.5	91.9
77	3	1/2	.05	.04	100	100	100	100	95.4	96.3

Table 6: Sensitivity analysis with n=25 when the alternative
mean is actually 10% larger than the hypothesized alternative
mean

Simulation Parameters		Actual Size		Actual Power		P(Ha <b>rej</b> )		$P(H_{\alpha} nre )$		
Seed	μ <sub>a</sub>	σα	ατ	α <sub>b</sub>	π	π <sub>b</sub>	t	b	t	b
1316	1/3	2	.05	.05	22.5	99.9	81.8	95.2	55.1	99.8
89	1/3	1/2	.05	.05	97.3	100	95.1	95.3	97.2	100
181	3	2	.05	.05	100	100	95.2	95.2	100	100
44	3	1/2	.05	.05	100	100	95.2	95.3	100	100

Table 7: Sensitivity analysis with n=25 when the alternative mean is actually 10% smaller than the hypothesized alternative mean

Simulation Parameters		Actual Size		Actual Power		P(H <sub>o</sub> ref)		$P(H_{\alpha} nre )$		
Seed	$\mu_{a}$	$\sigma_a$	α	α <sub>b</sub>	π	$\pi_b$	t	b	t	b
1920	1/3	2	.05	.05	18	99.8	78.3	95.2	53.7	99.8
118	1/3	1/2	.05	.05	89.7	100	94.7	95.3	90.2	100
56	3	2	.05	.05	100	100	95.2	95.2	100	100
55	3	1/2	.05	.05	100	100	95.2	95.3	100	100

Table 8: Sensitivity analysis with n=25 when the alternative standard deviation is actually 10% larger than the hypothesized alternative standard deviation

Simulation Parameters		Actual Size		Actual Power		P(H <sub>a</sub>  ref)		$P(H_{\alpha} nre )$		
Seed	μ	$\sigma_a$	α	α <sub>b</sub>	$\pi_t$	$\pi_b$	t	b	t	b
519	1/3	2	.05	.05	18.2	100	78.5	95.2	53.7	100
514	1/3	1/2	.04	.05	90.3	100	94.8	95.3	90.8	100
2020	3	2	.05	.05	100	100	95.2	95.2	100	100
214	3	1/2	.04	.05	100	100	95.2	95.3	100	100

Table 9: Sensitivity analysis with n=25 when the alternative standard deviation is actually 10% smaller than the hypothesized alternative standard deviation

Simulation Parameters		Actual Size		Actual Power		<b>P</b> (H	(rej)	$P(H_{\alpha} nre )$		
See		$\sigma_{\alpha}$					t b		t	b
d						, v				
95	1/3	2	.05	.05	22.8	99.2	82.0	95.2	55.2	99.2
145	1/3	1/2	.05	.05	97.4	100	95.1	95.3	97.4	100
193	3	2	.05	.05	100	100	95.2	95.2	100	100
201	3	1/2	.05	.05	100	100	95.2	95.3	100	100

# Application

Hypothesized variance under the null may differ from hypothesized variance under the alternative when studying heterogeneous treatment effects. This situation occurs in many biomedical and psychological research scenarios. To give some examples, heteroscedasticity occurs in drug studies (Howell, 1997), studies of risk taking (Watson, 1985), mental health testing (Feingold, 1992), studies of eating disorders (Gross, 1985), and studies of brain injury (Wilde, Boake, and Sherer, 1995), along with numerous studies in psychological and educational research (Keppel, 1991). See Grissom (2000) for a definition of heteroscedasticity and a discussion of the problem in clinical research. To provide a more detailed example here, we consider data from a recent study of functional electrical stimulation on horses presenting with sore backs.

Ravara et al. (2015) observed 2,004 equine muscle fibers, 1,461 before treatment with functional electrical stimulation and 1,643 after treatment to assess the effects of treatment on fiber diameters. The average post-treatment diameter was 52.57  $\mu$ m, while the average pretreatment diameter was 59.38  $\mu$ m, and the difference in means, -6.81  $\mu$ m, was statistically significant (p<.001) based on a two-sample Welch's t test. The authors argued that treatment acted independently on distinct fibers even within the same horse. All fibers were obtained from just six horses.

The question arises as to whether the treatment should be applied to horses presenting with sore backs. Subtracting each horse's post-treatment average fiber diameter from its pre-treatment average fiber diameter yields a sample of six differences: 3.1, -11.26, 9.43, -15.1, -14.9, and -11.1  $\mu$ m. Fiber diameters increased in two horses and decreased in four horses. Ravara et al. (2015, p. 113, 115, 116) acknowledge that different horses respond differently to treatment, i.e., the treatment effect is not constant across horses. See Rosenbaum (2009, Section 2.4) for a discussion of heterogeneous treatment effects.

Heterogeneous treatment effects that are independent of background variation lead to increased variance. The standard deviation under the alternative hypothesis (specifying an effective but heterogeneous treatment effect) can be larger than the standard deviation under a null hypothesis of ineffective treatment. Suppose under the null hypothesis that post minus pre differences for each horse are normally distributed with a mean of 0  $\mu$ m and a standard deviation of 4  $\mu$ m (the sample standard deviation across the six horses, pre-treatment, is 3.8  $\mu$ m), and under the alternative we hypothesize a mean of -6.8  $\mu$ m (the effect size observed on thousands of pooled fibers mentioned previously) and a standard deviation between 5 and 20  $\mu$ m.

Regardless of the specified standard deviation, the ball test indicates p<.05 (consistent with the pooled conclusion) while the t test results in an inconclusive p=.08. The ball test results were not sensitive to small variation in the other parameters, and as shown in our simulation section, power is not expected to be overly sensitive to departures from assumptions, including normality. We suggest that the ball test be considered only in situations when the parameters can be specified with a high degree of certainty, and when the variance under the alternative hypothesis differs from the variance under the null hypothesis.

#### Discussion

Our results indicate that routine use of t tests can contribute to scientific irreproducibility, especially when the parameters can be specified under the assumption of normality. Our sensitivity analyses indicate also that the ball test can improve reproducibility even when assumptions are mildly violated. Greatest improvement is likely to be found when the effect size  $\mu_{a}$  is small. It should be noted, however, that a strength of the t test is its rejection region that is invariant under large misspecifications for the alternative parameters, and the ball test should only be used with caution when the alternative parameters are far from certain. If the data are normally distributed and the test parameters are accurately and precisely defined, then the ball test should be considered.

Note that our reproducibility rates depend on the subjective assignment of  $p(H_0) = 0.5$ . The simulations can be repeated for other values of p and other simulation parameters as well. For those concerned about subjective reasoning, we point out that the traditional selection of a one- or two-sided, one-sample, mean test is itself a (coarse) act in Bayesian (subjective) reasoning. A more refined approach may involve more precise (fine) specification of the parameters, as we have done with the ball test. More generally, subjective distributions may be defined for the parameters, and the generalized likelihood ratio (Bain and Engelhardt, 1992, Chapter 12.8) can be replaced with Bayes's factors (Johnson, 2013a). Johnson (2013b) has introduced non-subjective, uniformly most powerful Bayesian tests, and for classical tests he has suggested lowering alpha for more consistency with Bayesian approaches.

The American Statistical Association has recently released a statement (ASA News) regarding "statistical significance and p-values." The statement contains six principles, two of which are relevant to the concerns of this paper: a) "decisions should not be based only on whether a p-value passes a specific threshold," and b) "p-values do not measure the probability that the studied hypothesis is true." Here we have addressed these issues by focusing on quantities such as  $P(H_a|ref)$  and  $P(H_0|nref)$ . We have revealed situations where reproducibility rates with the t test are as low as 50% (similar to a coin flip), whereas in these same situations reproducibility rates with the ball test are consistently over 90%. We recommend that researchers concerned about reproducibility consider the ball test when the data are

46 Biological Sciences

normally distributed and the test parameters are specified fairly well. This advice could have a positive impact on reproducibility in scientific fields where hypothesis tests are conducted to detect small effects.

# References

ASA News. (2016). American Statistical Association Releases Statement on Statistical Significance and P-Values. Alexandria, VA: American Statistical Association.

Bain, L.J., & Engelhardt, M. (1992). *Introduction to Probability and Mathematical Statistics*. Pacific Grove, CA: Duxbury/Thomson Learning.

Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science*, 7(6), 543–554.

Begley, C. G., & Ellis, L. M. (2012). Drug development: Raise standards for preclinical cancer research. *Nature*, 483(7391), 531–533.

Casadevall, A., & Fang, F. C. (2010). Reproducible science. *Infection and Immunity*, 78(12), 4972–4975.

Collins, F. S., & Tabak, L. A. (2014). NIH plans to enhance reproducibility. *Nature*, 505(7485), 612.

Feingold, A. (1992). Sex differences in variability in intellectual abilities: A new look at an old controversy. *Review of Educational Research*, 62, 61–84.

Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics*, 33(5), 587–606.

Grisson, R. (2000). Heterogeneity of variance in clinical data. *Journal of Consulting and Clinical Psychology*, 68(1), 155–165.

Gross, J. S. (1985). Weight modification and eating disorders in adolescent boys and girls. Unpublished doctoral dissertation, University of Vermont, Burlington.

Hogg, R.V., McKean, J.W., & Craig, A.T. (2013). *Introduction to Mathematical Statistics*. Boston: Pearson.

Howell, D. C. (1997). *Statistical Methods for Psychology* (4th ed.). Boston: Duxbury.

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Med*, 2(8), 124.

Ioannidis, J. P., Ntzani, E. E., Trikalinos, T. A., & Contopoulos-Ioannidis, D. G. (2001). Replication validity of genetic association studies. *Nature Genetics*, 29(3), 306–309.

Johnson, V.E. (2013a) Revised standards for statistical evidence. *Proceedings of the National Academy of Sciences of the United States of America*, 110(48), 19313–19317.

Johnson, V.E. (2013b). Uniformly most powerful Bayesian tests. *The Annals of Statistics*, 41(4), 1716–1741.

Keppel, G. (1991). *Design and Analysis: A Researcher's Handbook* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.

Moore, D., Notz, W., & Fligner, M. (2015). *The Basic Practice of Statistics*. New York: W.H. Freeman.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716.

Pashler, H., & Wagenmakers, E. J. (2012). Editors' introduction to the special section on replicability in psychological science a crisis of confidence? *Perspectives on Psychological Science*, 7(6), 528–530.

Popper, K. R. (1934/1959). *The Logic of Scientific Discovery*. London: Hutchinson.

Prinz, F., Schlange, T., & Asadullah, K. (2011). Believe it or not: how much can we rely on published data on potential drug targets? Nature Reviews Drug Discovery, 10(9), 712–712.

48 Biological Sciences

Ravara, B., Gobbo, V., Carraro, U., Gelbmann, L., Pribyl, J., & Schils, S. (2015). Functional electrical stimulation as a safe and effective treatment for equine epaxial muscle spasms: Clinical evaluations and histochemical morphometry of mitochondria in muscle biopsies. *European Journal of Translational Myology*, 25(2), 109–120.

Rosenbaum, P. (2009). *Design of Observational Studies*. New York: Springer.

Schooler, J. W. (2014). Metascience could rescue the replication crisis. *Nature*, 515, 9.

Steckler, T. (2015). Preclinical data reproducibility for R&D—the challenge for neuroscience. *Psychopharmacology*, 232(2), 317–320.

Wackerly, D., Mendenhall, W., & Scheaffer, R. (2008). Mathematical Statistics with Application (7th ed.). Belmont, CA: Brooks/Cole.

Watson, J. S. (1985). Volunteer and risk-taking groups are more homogeneous on measures of sensation seeking than control groups. *Perceptual and Motor Skills*, 61, 471–475.

Wilde, M. C., Boake, C., & Sherer, M. (1995). Do recognition-free recall discrepancies detect retrieval deficits in closed-head injury? An exploratory analysis with the California Verbal Learning Test. *Journal of Clinical and Experimental Neuropsychology*, 17, 849–855.

Vaux, D.L., Fidler, F., & Cumming, G. (2012). Replicates and repeats—what is the difference and is it significant? A brief discussion of statistics and experimental design. EMBO Reports, 13(4): 291–296.

# **Appendix 1 (Proofs)**

All observations are assumed to be i.i.d. normal.

# Lemma 1

The likelihood ratio test for the hypothesis  $H_{1}: \begin{matrix} \mu - \mu_{1} \\ \sigma - \sigma_{1} \end{matrix}$   $H_{2}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{1} \end{matrix}$   $H_{1}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{1} \end{matrix}$   $H_{2}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{1} \end{matrix}$   $H_{3}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{4}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{2}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{3}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{4}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \sigma_{2} \end{matrix}$   $H_{5}: \begin{matrix} \mu - \mu_{2} \\ \sigma - \mu_{2} \end{matrix}$   $H_{$ 

Proof: The likelihood ratio tests yields the following equations:

$$\Lambda(\mu_{0},\sigma_{0},\mu_{a},\sigma_{a},X) = \frac{L(\mu_{0},\sigma_{0}|X)}{L(\mu_{a},\sigma_{a}|X)}$$
$$= \frac{(2\pi\sigma_{0})^{\frac{1}{2}}e^{\frac{1}{2}\left(\frac{x_{1}-\mu_{0}}{\sigma_{0}}\right)^{2}}\cdots(2\pi\sigma_{0})^{\frac{1}{2}}e^{\frac{1}{2}\left(\frac{x_{n}-\mu_{0}}{\sigma_{0}}\right)^{2}}}{(2\pi\sigma_{a})^{\frac{1}{2}}e^{\frac{1}{2}\left(\frac{x_{1}-\mu_{0}}{\sigma_{a}}\right)^{2}}\cdots(2\pi\sigma_{a})^{\frac{1}{2}}e^{\frac{1}{2}\left(\frac{x_{n}-\mu_{0}}{\sigma_{a}}\right)^{2}}$$

$$=\frac{(2\pi\sigma_{0})^{-\frac{n}{2}}e^{-\frac{1}{2}\sum_{l=2}^{n}\left(\frac{x_{l}-\mu_{0}}{\sigma_{0}}\right)^{2}}}{(2\pi\sigma_{0})^{-\frac{n}{2}}e^{-\frac{1}{2}\sum_{l=2}^{n}\left(\frac{x_{l}-\mu_{0}}{\sigma_{0}}\right)^{2}}}$$

and with the substitution  $\mathbb{Z}_{i} = \frac{x_{i} - \mu_{0}}{\sigma_{0}}$  we get

$$\Lambda(\mu_0, \sigma_0, \mu_a, \sigma_a X) = \frac{(2\pi\sigma_0)^{\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n} (Z_i)^2}}{(2\pi\sigma_a)^{\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n} \left(\frac{\sigma_0 Z_i - (\mu_0 - \mu_a)}{\sigma_a}\right)^2}}$$

$$\frac{(2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n} (Z_i)^2}}{\left(2\pi \frac{\sigma_a}{\sigma_0}\right)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n} \left(\frac{Z_i - \left(\frac{\mu_a - \mu_0}{\sigma_0}\right)}{\sigma_0}\right)^2}} = \Lambda\left(0, 1, \frac{\mu_a - \mu_0}{\sigma_0}, \frac{\sigma_a}{\sigma_0}, Z\right),$$

which proves the lemma.

50 Biological Sciences

#### Theorem 1

When  $\mathbf{v}_{\mathbf{0}} \neq \mathbf{v}_{\mathbf{a}}$ , the boundary of the rejection region for the likelihood ratio test of  $\mathbf{H}_{0}: \mathbf{\sigma} = \mathbf{\sigma}_{0}$  vs.  $\mathbf{H}_{\mathbf{A}}: \mathbf{\sigma} = \mathbf{\sigma}_{\mathbf{a}}$  is a sphere centered at  $\mathbf{\sigma}_{0}(\mu_{0} - \mu_{a})$ (b, b, ..., b) where  $\mathbf{b} = \frac{\mathbf{\sigma}_{0}(\mu_{0} - \mu_{a})}{\mathbf{\sigma}_{a}^{2} - \mathbf{\sigma}_{0}^{2}}$ . When  $\mathbf{\sigma}_{0} > \mathbf{\sigma}_{\mathbf{a}}$ , the region is the interior of that sphere. When  $\mathbf{\sigma}_{0} < \mathbf{\sigma}_{\mathbf{a}}$ , the region is the exterior of that sphere. When  $\mathbf{\sigma}_{0} = \mathbf{\sigma}_{\mathbf{a}}$ , the likelihood ratio test reduces to the z test.

#### Proof

Because of lemma 1, we may assume without loss of generality that  $\mu_0 = 0$  and  $\sigma_0 = 1$ . By the results of the Neyman-Pearson, lemma ensures that the likelihood ratio test is most powerful. The likelihood-ratio test rejects the null hypothesis when the likelihood ratio is less than some constant K, and in our case we have

$$K > \frac{L(\mu = 0, \sigma = 1, |X)}{L(\mu_{g}, \sigma_{a}|X)} - \frac{(2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n} (X_{i})^{2}}}{(2\pi\sigma_{a})^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n} (\frac{X_{i} - \mu_{g}}{\sigma_{a}})^{2}}} = \sigma_{a}^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n} X_{i}^{2} - (\frac{X_{i} - \mu_{g}}{\sigma_{a}})^{2}}).$$

We solve for the region of the sample space (x variables) satisfying the inequality:

$$\ln\left(K\sigma_{a}^{-\frac{n}{2}}\right) > -\frac{1}{2}\left(\sum_{i=1}^{n} X_{i}^{2} - (\frac{X_{i} - \mu_{a}}{\sigma_{a}})^{2}\right)$$
$$-2\sigma_{a}\ln\left(K\sigma_{a}^{-\frac{n}{2}}\right) < \sum_{i=1}^{n} \sigma_{a}^{2}X_{i}^{2} - (X_{i} - \mu_{a})^{2}$$
$$-2\sigma_{a}^{2}\ln\left(K\sigma_{a}^{-\frac{n}{2}}\right) < \sum_{i=1}^{n} X_{i}^{2}(\sigma_{a}^{2} - 1) + 2\mu_{a}X_{i} - \mu_{a}^{2}$$
$$\frac{\mu_{a}^{2}\sigma_{a}^{2}}{\sigma_{a}^{2} - 1} - 2\sigma_{a}^{2}\ln\left(K\sigma_{a}^{-\frac{n}{2}}\right) < (\sigma_{a}^{2} - 1)\sum_{i=1}^{n} \left(X_{i} + \frac{\mu_{a}}{\sigma_{a}^{2} - 1}\right)^{2}.$$

Student's t-test 51

Let 
$$A = \frac{\prod_{i=1}^{n} \frac{\sigma_{i}^{2} - 1}{\sigma_{i}^{2} - 1}}{\sigma_{i}^{2} - 1} \text{ and } B = \frac{\prod_{i=1}^{n} \frac{\sigma_{i}^{2}}{\sigma_{i}^{2} - 1}}{\prod_{i=1}^{n} \sigma_{i}^{2} - 1} \text{ and } A < \sum_{i=1}^{n} (X_{i} + B)^{2} \text{ .}$$
If  $\sigma_{a}^{2} > 1$ , then we reject the null hypothesis when  $A > \sum_{i=1}^{n} (X_{i} + B)^{2}$ . (Note that under the null hypothesis the distribution of  $\sum_{i=1}^{n} (X_{i} + B)^{2}$  . (Note that under the null hypothesis the distribution  $\sum_{i=1}^{n} (X_{i} + B)^{2}$  is a noncentral chi-squared distribution with noncentrality parameter  $\lambda = \sum_{i=1}^{n} B^{2}$  and under the alternative hypothesis  $\sum_{i=1}^{n} (X_{i} + B)^{2}$  is a noncentral chi-squared distribution with noncentrality parameter  $\lambda = \sum_{i=1}^{n} (\frac{\mu_{a} + B}{\sigma_{a}})^{2}$ . The ball test could be carried out using the noncentral chi-squared distribution.) Setting b=-B completes the proof of the first three claims. In original terms. we have  $B = \frac{\sigma_{0}(\mu_{a} - \mu_{0})}{\sigma_{a}^{2} - \sigma_{0}^{2}}$  and  $A = (\frac{\sigma_{a}^{2}(\mu_{a} - \mu_{0})^{2}}{(\sigma_{a}^{2} - \sigma_{0}^{2})^{2}} - \frac{2\sigma_{a}^{2}\ln((\kappa(\frac{\sigma_{a}}{\sigma_{0}})^{\frac{n}{2}})}{\sigma_{a}^{2} - \sigma_{0}^{2}}$  With  $\sigma_{a} = \sigma_{0}$  we have  $K > \frac{L(\mu - 0, \sigma - 1, [X)}{L(\mu_{a}, \sigma_{a} - 1[X)} = \frac{(2\pi)^{-\frac{n}{2}}e^{-\frac{1}{2}\sum_{i=1}^{n}(X_{i} - \mu_{a})^{2}}{(2\pi)^{-\frac{n}{2}}e^{-\frac{1}{2}\sum_{i=1}^{n}(X_{i} - \mu_{a})^{2}}$ 

$$= e^{\frac{1}{2}(\sum_{i=1}^{n} X_i^2 \cdot (X_i \cdot \mu_a)^2)} = e^{\frac{1}{2}(\sum_{i=1}^{n} 2\mu_a X_i \cdot n\mu_a^2)}.$$

Therefore, when  $\mu_a > 0$ , we reject the null hypothesis when  $\sum_{i=1}^{n} X_i$  is greater than some critical value. When  $\mu_a < 0$ , we reject the null hypothesis if  $\sum_{i=1}^{n} X_i$  is less than some critical value. This rejection

null hypothesis if f is less than some critical value. This rejection decision corresponds with the traditional rejection criteria. This completes the proof of the fourth claim and the whole theorem.

# Appendix 2 (Seeding & Code for simulations)

###Researchers should change the parameters below for their research needs

set.seed(555) ### replace 555 with your own seed mu0=0 ### null population mean s0=1 ### null population standard deviation mu1=1/10 ### alternative population mean s1=2 ### alternative population standard deviation n=10 ### sample size P=.5 ### P(H\_0)

m=1000000 ### for repititions B=-(mu1/(s1^2-1)) ### center of ball f=function(x) sqrt(sum((x-B)^2)) ### measures distance from B distances=numeric(m) ### m blank entries

for(i in 1:m) {distances[i]=f(rnorm(n,mu0,s0))} ### fills in blanks alpha=0 ### seed for alpha r=20 ### seed for r

```
### computes ball power
reject=numeric(m)
for(i in 1:m){reject[i]=f(rnorm(n,mu1,s1))}
if(s1<1){power=length(which(reject<r))/length(reject)}
if(s1>1){power=length(which(reject>r))/length(reject)}
```

```
### computes t power
t=numeric(m)
for (i in 1:m){x=rnorm(n,mu1,s1)
t[i]=mean(x)/(sd(x)/sqrt(n))}
```

```
tpower=length(which(t>qt(.95,n-1)))/length(t)
```

```
### computes P(H_0|nrej) and P(H_a|rej) for ball and t
bph0=(P-P*alpha)/(P-P*alpha+(1-power)-(1-power)*P)
alphat=.05
tph0=(P-P*alphat)/(P-P*alphat+(1-tpower)-(1-tpower)*P)
bpha=((1-P)*power)/((1-P)*power+P*alpha)
tpha= ((1-P)*tpower)/((1-P)*tpower+P*alphat)
```

##Print results r;alpha;tpower;bph0;tph0;bpha;tpha

# **Integration of Produced Water with Microalgae Cultivation**

Brian McNeil,<sup>1</sup> Derek Hess,<sup>1</sup> Eric Torres,<sup>1</sup> Ronald C. Sims,<sup>1</sup> Jason C. Quinn<sup>2</sup>

<sup>1</sup>Utah State University, <sup>2</sup>Colorado State University

#### Abstract

Microalgae represents a promising biofuel feedstock as it can be cultivated on low-quality land and integrated with a variety of waste streams. Produced water is the largest waste stream generated in the oil and gas industry. The integration of the two systems is synergistic as algal cultivation requires large volumes of water and there is a need for remediation of produced water. Produced water contains inorganic and organic contaminants that could inhibit growth and lipid production of microalgae directly impacting the yield of the system. In this study, produced water from the Uintah Basin, Utah, oil and gas industry was substituted for water in the growth media for microalgae, Nannochloropsis salina. Experiments utilized produced water as the base for media preparation and evaluated microalgae productivity at different dilutions with deionized water, i.e., 0% (control), 25%, 50%, 75%, and 100% produced water, as well as the recycling of produced water based growth media. Results show that produced water severely inhib-

its the growth of microalgae, with a 102% decrease in the 100% produced water media. Recycling of growth media has the potential to positively impact productivity results at concentrations of produced water less than 25%. Discussion focuses on the importance of integration studies, the impact of reduced productivity with higher concentration of produced water, and the need for further experimental growth studies.

# Introduction

Increased global energy demand combined with the negative impacts associated with the consumption of fossil based resources has motivated the development of renewable fuel sources (Nejat et al., 2015). Corn and soybeans are current first-generation feedstocks that are utilized for domestic biofuels; however, concerns over converting feed-based crops to fuel have limited future growth in food-based biofuels (Mohr & Raman, 2013; Naik et al., 2010). Other energy-based crops such as jatropha, switchgrass, and miscanthus are being investigated as alternatives to feed-based feedstocks but still require agricultural land for production (Damartzis & Zabaniotou, 2011; Heaton et al., 2008; Sims et al., 2010). Microalgae represents a promising biofuel feedstock as it has a high photosynthetic efficiency, has potentially high lipid content, can be cultivated on low-quality land, and can utilize various waste streams (Benemann et al., 1982; Quinn & Davis, 2015a; Wijffels & Barbosa, 2010). These advantages have led researchers to investigate various aspects of the microalgae biorefinery concept with technical challenges identified in the various processing steps. A current critical barrier regarding commercialization is the difficulty of developing an effective, economically viable cultivation system that efficiently utilizes available resources.

Microalgae-based biofuel production systems are being critically evaluated through techno-economic assessment, life-cycle assessment, and resource assessment. Results from studies indicate microalgae can potentially produce significant amounts of biofuel (ANL;NREL;PNNL, 2012; Beal et al., 2012; Bennion et al., 2015; Davis et al., 2011; Moody et al., 2014; Quinn & Davis, 2015b; Venteris et al., 2013). Current assessments of microalgae production systems assume the use of various wastes to enhance growth or reduce cost. The use of these waste streams is synergistic as the algae can utilize nutrients in the wastes stream while helping to reduce the environmental impact of the wastes. Seamless integration has been demonstrated for a number of waste streams including wastewater treatment, carbon from industrial processes like coal power plants, and contaminated waterways (Huang et al., 2016; Mehrabadi et al., 2016; Mortensen & Gislerod, 2016; Rickman et al., 2013; Singh et al., 2016; Stockenreiter et al., 2016; Wuang et al., 2016; Zhan et al., 2016).

Experimental work has highlighted the importance of understanding how wastes affect algae growth. Napan et al. (2015) show coal flue gas has the potential to negatively impact microalgae productivity based on heavy metal contamination in the flue gases. The majority of techno-economic assessments, life-cycle assessments, and resource assessments assume colocation with point source flue gas without experimentally validating whether such colocation is feasible (ANL;NREL;PNNL, 2012; Davis et al., 2011; Ouinn et al., 2014; Sun et al., 2011). Multiple studies have shown that the recycling of nutrientrich waste from hydrothermal liquefaction has the potential to degrade microalgae growth as there are other nondesirable components being introduced (Alba et al., 2013; Barreiro et al., 2015; Wang et al., 2016). However, a number of sustainability assessments assume the seamless integration of the aqueous phase from hydrothermal liquefaction with cultivation as it represents a requirement for economic viability and sustainability of the overall process (ANL;NREL;PNNL,2012; Bennion et al., 2015; Davis et al., 2014; Gerber et al., 2016; Liu et al., 2013; Venteris et al., 2014b).

It has been speculated that microalgae cultivation can be coupled with produced water from the oil and gas industry. This is a synergistic activity as microalgal cultivation is a water-intensive activity and there is a need to remediate the produced water (Batan et al., 2013; Farooq et al., 2015; Gerbens-Leenes et al., 2014; Gross et al., 2015; Guieysse et al., 2013; Tu et al., 2016; Venteris et al., 2014a). Further, oil and gas-based produced water is typically being extracted in areas with poor land quality where traditional agriculture cannot be supported but microalgae cultivation can be developed.

Based on a survey of the literature, there exists the need to experimentally evaluate the potential of integrating produced water with microalgae cultivation. This study experimentally evaluates the impact of media substitution with produced water on the productivity of *Nannochloropsis salina*, a promising strain for biofuel production. Media was created using produced water acquired from the Uintah Basin in Utah at varying concentrations, from 100% produced water to 25% produced water diluted with deionized water. A control made with 100% deionized water on microalgae productivity and lipid content. Microalgae was cultivated in a closed photobioreactor system at a light intensity comparable to summertime intensities and harvested after

approximately 1 week. Work was extended to evaluate the impact of media recycle. A two-week experiment was completed in which the media was either recycled or replaced at the end of the first week to evaluate the potential for growth-inhibiting compounds to be sorbed or volatilized. Each growth experiment was divided into groups of three photobioreactors (PBRs). For each condition, PBRs were operated in triplicate with the biomass productivity and lipid yield determined. Discussion focuses on the importance of experimentally evaluating the use of produced water and the potential impact on the economic viability and sustainability of integrating microalgae cultivation with oil and gas produced water.

#### Methods

This section outlines the system and inoculation procedure for the microalgae cultivation system, integration of and source of produced water, analysis techniques used for the characterization of the produced water, and techniques for monitoring and determining algal culture density and lipid content. The experimental work is divided into two phases: 1) evaluation of the effects of produced water on microalgal productivity, and 2) the impact of media recycle on productivity. All experimental work was completed in triplicate biological reactors.

#### Cultivation

Experimental work was performed in closed tubular PBRs. The inoculum for the system was grown from a starter culture. Nannochloropsis salina microalgae were obtained from the Culture Collection of Algae at the University of Texas at Austin for use in this study. Initially, the culture was grown in sterile petri dishes at constant low light on solid, nutrient-rich media. Colonies were transferred into Erlenmever flasks with 200 mL of liquid media and placed on a shaker table illuminated with low light. Cultures were finally transferred to 1.1 L of media in borosilicate glass PBRs. PBRs were 4.5 cm in diameter and 80 cm in length. The PBRs were illuminated at 200  $\mu$ mol m<sup>-2</sup> s<sup>-1</sup> for 16 hours per day and submersed in a temperature-controlled water bath maintained at  $23 \pm 1^{\circ}$ C. Cultures were mixed using CO<sub>2</sub>-enriched sparge air at a flow of 2.5 L min<sup>-1</sup> and CO<sub>2</sub> flow of 25 ml min<sup>-1</sup>. The media consisted of NaCl (761.2 mM), CaCl<sub>2</sub>·2H<sub>2</sub>O (1.0 mM), KCl (6.4 mM), Na<sub>2</sub>SiO<sub>3</sub>·9H<sub>2</sub>O (0.2 mM), MgSO<sub>4</sub>·H<sub>2</sub>O (6.0 mM), KNO<sub>3</sub> (10.1 mM), KH<sub>2</sub>PO<sub>4</sub> (0.5 mM), ammonium ferric citrate (2.0×10<sup>-2</sup> mM), H<sub>3</sub>BO<sub>3</sub> (1.5×10<sup>-2</sup> mM), Na<sub>2</sub>MoO<sub>4</sub> ·2H<sub>2</sub>O (5.0×10<sup>-5</sup> mM), MnCl<sub>2</sub> ·4H<sub>2</sub>O  $(1.5 \times 10^{-3} \text{ mM})$ , ZnSO<sub>4</sub> ·7H<sub>2</sub>O (2.1×10<sup>-4</sup> mM), CuSO<sub>4</sub> ·5H<sub>2</sub>O (8.0×10<sup>-5</sup>) mM) dissolved into deionized water. All reagents used were analytical grade. Prior to inoculation, media was autoclaved at 120°C for 30 minutes, and after it returned to room temperature, sterile biotin, vitamin B12, and thiamine solutions were added. The media was autoclaved to minimize contamination.

The experimental system utilized the same-sized PBRs as the inoculation system. The cultivation system consisted of a thermal water bath maintained at  $23 \pm 1^{\circ}$ C in which the PBRs were 90% submerged. The PBRs were inoculated with a volume of 1.1 L of microalgal culture at a density of approximately 2 g  $L^{-1}$ . This system was designed to be representative of outdoor light intensities and was illuminated at a photosynthetically active radiation level of 984 µmol m<sup>-2</sup> s<sup>-1</sup> using T5 fluorescent lights. The 12 reactors were divided into groups of three, with a pH monitor in one reactor from each group. The pH of the system was maintained through feedback control with CO<sub>2</sub> added to the sparge air only when the pH rose above 7.0. This maintained the pH value of the reactors at  $7.0 \pm 0.1$ . Sparge air was provided to three PBRs at a rate of 1.5 L min<sup>-1</sup> with CO<sub>2</sub> supplied at 50 ml min<sup>-1</sup> when needed. Previous experimental work validated controlling the pH of three triplicate reactors though the monitoring of one of the PBRs. Algal cultures were harvested with a batch centrifuge operated at 4303 g for 30 min. For analytical work, harvested biomass was placed in a -80°C freezer and then lyophilized.

#### Produced water

Because of the wide variety of geological and geographical environments in which oil and gas, and consequently produced water, are found, there is an inherent variability in levels and types of contamination of produced water. For this experimental work, produced water was acquired from the oil and gas industry in the Uintah Basin, Utah. Inorganic and semi-volatile analysis of the produced water was performed by Chemtech-Ford Laboratories, Salt Lake City, UT. Trace metal analysis was performed as detailed below.

For the experimental growth portion of this study, produced water was used in place of deionized water to prepare media. All of the salts and nutrients outlined above were added to the produced water media, with the only difference being the substitution of deionized water with produced water. The additional nutrients that come with the produced water are not considered in the media recipe. The produced water media was not autoclaved as this is too expensive when considering largescale cultivation. Multiple dilutions were tested, inoculating the system with produced water content of 0% (control), 25%, 50%, 75%, and 100%. A second set of experiments was performed that focused on understanding the potential to recycle growth media. This set of experiments cultivated algae for the first week on nutrient-rich media that was 0% (control), 50%, and 100% produced water. The second week included harvesting the algae and reinoculating in fresh standard nutrient-rich media (control), fresh nutrient-rich media containing 50% produced water, existing 50% produced water media with fresh nitrogen and phosphorus, and existing 100% produced water with fresh nitrogen and phosphorus.

#### Growth lipid measurement

Algal growth was measured daily using optical density (OD) at 750 nm using a GENESYS 5 spectrophotometer. The OD value was correlated to total suspended solids in g  $L^{-1}$  based on an ash-free dry weight analysis. The correlation coefficient was previously determined with a  $R^2$ =0.987.

Lipid analysis was performed based on the methods of Mishra et al. (2014). Quantifying the lipid content is done through a sulfophospho-vanillin (SPV) colorimetric method. Dried biomass was diluted to 2.33 g L<sup>-1</sup> with deionized water. One hundred microliters of microalgae culture was digested in 2 ml of sulfuric acid at 100°C for 10 minutes and then cooled in an ice bath for 5 minutes. Five milliliters of the reagent (vanillin, ethanol, deionized water, and phosphoric acid) was added to the sample and incubated at 37°C for 15 minutes. The OD of the digested sample was then measured at 530 nm. Lipid content was then calculated based on the mass of microalgae added and a standard curve. The standard curve was linear with a correlation coefficient of R<sup>2</sup>=0.999.

#### Trace contaminant analysis

The produced water was analyzed for trace contaminants through inductive coupled plasma mass spectrometry (ICP-MS). A 10-ml sample of produced water was digested in trace metal grade nitric acid (70%) at 105°C. The digestion took place over approximately seven days and was considered completed once the solution was visibly clear. The digestion was performed in a fume hood with periodic addition of nitric acid (approximately 1 ml) and deionized water such that the volume remained constant. The sample was periodically vortexed to ensure complete digestion. After digestion was complete, the sample was transferred to a 10-ml Luer-Lok tip syringe and filtered through 0.45µm filters into a 25-ml volumetric vial. Deionized water was run though the syringe and filter to capture any residuals and added to the volumetric vial with a final volume of 25 ml. A sample volume of 10 ml was then transferred to a capped test tube and stored in a refrigerator until analysis. ICP-MS analysis was performed by the Utah Water Research Laboratory.

#### Results

Experimental results include two data sets: 1) results from weeklong growth and lipid trials with an evaluation of various produced water media dilutions, and 2) results from week-long trials followed by reinoculation in the same media to evaluate the potential to recycle produced water media. Growth results indicate produced water media negatively impacts productivity with minimal improvements resulting from media recycle. Analytical results on the composition of the produced water are presented and discussed in relation to growth inhibition.

#### Impact of produced water on productivity

The potential to use oil and gas produced water for *Nannochloropsis salina* microalgae cultivation was explored. Standard deionized water base media was replaced with produced water at various dilutions. The effect of produced water at various produced water dilutions on the growth of *N. salina* (n=3 for each dilution) was directly compared with the control (n=6) over a one-week growth period (Figure 1).



**Figure 1.** Growth study results for *Nannochloropsis salina* comparing the control with multiple dilutions of produced water (100%, 75%, 50%, 25%, and 0%) with deionized water. Results are the average of triplicate with standard deviations represented by error bars.

Results show a statistically significant decrease in biomass productivity in the 100% produced water starting on day 3. The biomass productivity of the system was dramatically impacted with control cultures showing a production of  $3.78 \pm 0.38$  g L<sup>-1</sup> wk<sup>-1</sup> and the 100% produced water cultures showing  $-0.09 \pm 0.38$  g L<sup>-1</sup> wk<sup>-1</sup>, a 102% decrease in biomass productivity. A follow-up experiment was conducted that included various media compositions including a control (0% produced water), 25% produced water, 50% produced water, and 75% produced water. The experimental work was conducted in triplicate with results presented in Figure 1. Results show dilution of produced water can improve the performance of the system. A dilution of 75% produced water shows a negative impact on biomass productivity with a 33% reduction in growth over the 1-week trial. Decreasing the produced water to 50% and 25% of the media did not statistically impact the biomass productivity.

Harvested biomass was characterized for lipid content with results shown in Figure 2. Lipid accumulation in microalgae can be triggered by a number of stressors, which include but are not limited to, nutrient depletion, light, temperature, and salinity. Based on previous experimental work, the system is operated in batch mode to induce nutrient deplete lipid accumulation. The lipid content in the control experiments is expected and similar to those reported in literature (Quinn et al., 2012). Complete replacement of deionized water with produced water in the media recipe results in a 40% reduction in lipid content at the end of the experiment. This result is expected as the culture does not appear



**Figure 2.** Lipid content of biomass after harvest for the 0% (Cont.) and produced-water growths in 25%, 50%, and 75% produced-water media. Results are averages of triplicate samples with standard deviation presented with error bars.

to increase in density over the week-long growth period. Lipid content is negatively impacted at the 75% and 50% dilutions with a 31.0% and 13.4% decrease, respectively, compared with the control. At the 25% dilution, there is a 4.3% increase in lipid content compared with the control. The increase in lipid content could be due to an increase in cell stress at a level that does not yet impact cell function. The 50% dilution does not significantly impact biomass productivity but does negatively impact lipid yield.

*N. salina* represents a promising feedstock for biofuel production. By combining the biomass productivity results with lipid content, a more holistic evaluation of the impact of produced water content can be evaluated. A dose-response curve illustrating the impact of produced water content on lipid productivity is presented in Figure 3, which combines the results from the two experiments presented above. The results show initially a positive impact on lipid productivity with low dilutions, 25% of produced water. Increasing the dilution over 25% results in a negative impact on lipid productivity. At the full replacement of media with produced water, the lipid productivity is decreased by 68.9%.

The negative impacts associated with integrating oil and gas produced water with microalgae cultivation will be detrimental to the sustainability of a large-scale production facility at concentrations of



**Figure 3.** Combined impact of produced water on *N. salina* biomass growth and lipid production in 0%, 25%, 50%, 75%, and 100% produced-water media. Standard deviations within 3 replicates are represented by error bars.

produced water higher than 25%. Initial results show the produced water contains growth-inhibiting constituents. At minimal dilutions of 25% or less, the growth inhibitors do not significantly impact the culture system.

#### Produced-water composition

The produced water was analyzed to determine potential growth inhibitors. The analytical work included evaluating the concentration of inorganic contaminants, semi-volatile compounds, and trace metal contaminants. A summary of the inorganics and volatile compounds detected are presented in Table 1. The results show two critical nutrients to be present: nitrogen and phosphorus. The produced water experiments thus had an excess of these nutrients. Strictly considering nutrients, the biomass productivity of the system would be expected to increase compared with the control, with the lipid yield at the end of a 5- to 7-day growth to decrease. The decreased lipid yield may be the result of not entering into a nutrient-deprived stressing regime. However, results show a negative impact on biomass productivity. Other constituents of the produced water are dominating the impacts. Further, the nutrients detected could plausibly not be bioavailable.

Table 1. Analytical results for detectible compounds in oil andgas produced water									
Inorganic	Amount	Unit							
Ammonia as N	7.7	mg/L							
Biochemical oxygen demand	272	mg/L							
Phosphate, ortho-	4.5	mg/L							
Phosphorus, total	16	mg/L							
Total dissolved solids	65	g/L							
Total suspended solids	222	mg/L							
Volatile suspended solids	72	mg/L							

The composition of the produced water includes compounds such as the volatile suspended solids. Analytical work included the detection of some semi-volatile compounds. Specifically, 2-methylphenol (10  $\mu$ g/L), acetone (34.5  $\mu$ g/L), carbon disulfide (1.3  $\mu$ g/L), methyl ethyl ketone (14.8  $\mu$ g/L), and diesel range organics (15.7 mg/L) were all detected in analysis. The direct impact of these individual compounds on microalgae growth has yet to be directly quantified. The composition of

the produced water used in this study is compared to other produced water compositional studies. Benko and Drewes (2008) surveyed produced water from a variety of western U.S. locations, with results showing total suspended solids ranging between 1000 and 4000 mg L<sup>-1</sup>. The produced water in this study was much lower at 222 mg L<sup>-1</sup>. Strømgren et al. (1995) studied produced water generated in the North Sea and included a heavy metal analysis. One of the three produced water sites had significant Zn levels, 10.5 mg L<sup>-1</sup>, with the other two sites having minimal levels. Inorganic or trace metal analysis was also completed in this study with results presented in Figure 4.



**Figure 4.** Heavy metals content: This graph represents the concentration of 14 metal contaminants present in the produced water.

The majority of the trace metal contaminants were not present in significant quantities with the exception of As, Mn, Se, and Zn. Previous studies have specifically investigated the potential of N. salina to bioremediate Zn, with results showing a high removal efficiency of 95.77%; however, the impact on growth was not studied (Hala & Sjahrul, 2013). Dong et al. (2014) evaluated the growth effects of Zn and found 2.64 mg  $L^{-1}$  would reduce growth by 50%. The results from this study have a high amount of Zn, which could be contributing to the decrease in biomass productivity at the high produced water concentrations. Other ongoing experimental work has shown a small decrease (approximately 10%) in biomass productivity when similar levels of Zn, As, and Se are introduced into the growth media of N. salina. The amount of As present in the produced water is somewhat alarming. Ni and Cu have previously been shown to negatively impact microalgae productivity (Bartlett et al., 1974; Gibson, 1972; Martínez-Ruiz & Martínez-Jerónimo, 2015; Murray-Gulde et al., 2002). However, the

amount of these contaminant detected in the produced water are small in comparison to previous studies. The levels of Ni are slightly higher than that of Cu. The total amount of detected metals is directly compared with the World Health Organization (WHO) drinking water guidelines (Figure 5).



**Figure 5.** Concentration of heavy metals in relation to the World Health Organization drinking water guidelines. Arsenic is not shown because of extreme levels, at over 4200%, and there is no guideline for Cobalt or Tin.

The produced water contains a variety of compounds that limit use. Specifically, the levels of Ni, Se, and As exceed the recommendations from the WHO. Limited algal growth studies have been performed in the presence of some of the metals detected. Other studies have focused on the end fate of metals and have not included results related to the impact on biomass productivity.

#### Recycling produced water

Previous techno-economic and life-cycle modeling have highlighted the importance of growth media recycle (Batan et al., 2010; Davis et al., 2011; Quinn & Davis, 2015a). Microalgae cultures are dynamic in nature and could adapt to the produced water media (Buchanan et al., 2013). Experimental work was performed to evaluate the potential of produced water media recycle. The experimental work started with a 1-week-long growth period on media that was 50% produced water, 100% produced water, and a control (0% produced water). At the end of the 1-week growth period, the algae was harvested and the system was reinoculated as follows: The control harvested algae was reinoculated into fresh standard media, the "<sup>1</sup>/<sub>2</sub>" harvested algae was reinoculated in fresh 50% produced water media, the "recycle" harvested algae was reinoculated in the same 100% produced water media, and the "replace" harvested algae was reinoculated in 100% fresh produced water media. Biomass productivity and ending lipid content results are presented in Figure 6 and Figure 7, respectively.



**Figure 6.** Two-week growth study results: (Top) Week 1 growth results representing control, ½ produced water, and two sets of full produced water media (labeled recycle and replace). Standard deviations within 3 replicates are represented by error bars. (Bottom) Week 2 results representing growth in recycled control, ½ produced water, and full produced water media, as well as replaced produced water media. The average of triplicates is presented with error bars representing 1 standard deviation.



**Figure 7.** Two-week study lipid results: (Left) Lipid content after first week in control, <sup>1</sup>/<sub>2</sub> produced water, and two sets of full produced water media (labeled recycle and replace). Standard deviations within 3 replicates are represented by error bars. (Right) Lipid content after second week in recycled control, <sup>1</sup>/<sub>2</sub> produced water, and full produced water media, as well as replaced produced water media. Standard deviations within 3 replicates are represented by error bars.

Lipid productivity results are similar to the previous growth results. The growth of microalgae on produced water media of 50% and 100% negatively impacts productivity. The biomass growth in the 100% produced water media for week 1 is slightly higher than the previous initial experiment. The biomass productivity of the culture in the 100% produced water in the first week of growth was reduced by 73% compared with the 102% found in the first experimental work. The growth results in week 2 illustrate that the microalgae appears to be adapting to the produced water. However, similar yields in the recycled and the fresh 100% produced water media show the growth inhibitor is not being negated through media recycle. Results show the lipid content of the ending biomass is improved in the 100% produced water recycle. This result shows the microalgae culture could be adapting to the produced water growth media with standard nutrient deplete lipid induction achieved. Further investigation with the microalgae cultures, in terms of lipid productivity show an improved yield with media recycle. However, the general trend shows a negative impact from produced water cultivation (Figure 8).



**Figure 8.** Comparison of lipid productivity in first week vs. second week when the media was recycled in all except for the second set using produced-water media, which was replaced. Standard deviations within 3 replicates are represented by error bars.

The integration of produced water from oil and gas with microalgae cultivation has the potential to negatively impact microalgae productivity. The experimental results of this study show the assumption that integrating oil and gas produced water with algal cultivation in sustainability assessments is not a seamless activity. Growth on 100% produced water media dramatically decreases *N. salina* biomass productivity and end lipid content. Results from media-recycle experimentation showed slight improvements but still a significant decrease in yield. Analysis of the produced water shows a variety of compounds and trace metal contaminants, but limited study of microalgae growth in environments with these individual contaminants makes identifying a specific growth inhibitor challenging. Results show the need for integration of experimental data with engineering system modeling to support sustainability modeling of algal-based fuel production systems.

#### Acknowledgments

The produced water industry collaboration and produced water sample and chemical analysis that were used in this study were obtained through the financial support of the State of Utah Governor's Office of Economic Development, Award Number 160723. The authors acknowledge financial support from the National Science Foundation (award number 1335550). 70 Biological Sciences

# References

Alba, L.G., Torri, C., Fabbri, D., Kersten, S.R.A., Brilman, D.W.F. 2013. Microalgae growth on the aqueous phase from Hydrothermal Liquefaction of the same microalgae. *Chemical Engineering Journal*, **228**, 214-223.

ANL, NREL, PNNL. June 2012. Renewable Diesel from Algal Lipids: An Integrated Baseline for Cost, Emissions, and Resource Potential from a Harmonized Model. US Department of Energy Biomass Program.

Barreiro, D.L., Bauer, M., Hornung, U., Posten, C., Kruse, A., Prins, W. 2015. Cultivation of microalgae with recovered nutrients after hydrothermal liquefaction. *Algal Research-Biomass Biofuels and Bioproducts*, **9**, 99-106.

Bartlett, L., Rabe, F.W., Funk, W.H. 1974. Effects of copper, zinc and cadmium on *Selanastrum capricornutum*. *Water Research*, **8**(3), 179-185.

Batan, L., Quinn, J., Willson, B., Bradley, T. 2010. Net energy and greenhouse gas emission evaluation of biodiesel derived from microalgae. *Environmental Science & Technology*, **44**, 7975-7980.

Batan, L., Quinn, J.C., Bradley, T.H. 2013. Analysis of water footprint of a photobioreactor microalgae biofuel production system from blue, green and lifecycle perspectives. *Algal Research*, **2**(3), 196-203.

Beal, C.M., Hebner, R.E., Webber, M.E., Ruoff, R.S., Seibert, A.F. 2012. The energy return on investment for algal biocrude: results for a research production facility. *BioEnergy Research*, **5**(2), 341-362.

Benemann, J.R., Goebel, R.P., Weissman, J.C., Augenstein, D.C. 1982. Microalgae as a source of liquid fuels. *Final Technical Report, US Department of Energy, Office of Research*, DOE/ER/30014-TR.

Benko, K.L., Drewes, J.E. 2008. Produced water in the Western United States: geographical distribution, occurrence, and composition. *Environmental Engineering Science*, **25**(2), 239-246.

Bennion, E.P., Ginosar, D.M., Moses, J., Agblevor, F., Quinn, J.C. 2015. Lifecycle assessment of microalgae to biofuel: Comparison of thermochemical processing pathways. *Applied Energy*, **154**, 1062-1071.

Buchanan, J.J., Slater, F.R., Bai, X., Pratt, S. 2013. Algal growth and community structure in a mixed-culture system using coal seam gas water as the water source. *Environmental Technology*, **34**(6), 695-701.

Damartzis, T., Zabaniotou, A. 2011. Thermochemical conversion of biomass to second generation biofuels through integrated process design—A review. *Renewable and Sustainable Energy Reviews*, **15**(1), 366-378.

Davis, R., Aden, A., Pienkos, P.T. 2011. Techno-economic analysis of autotrophic microalgae for fuel production. *Applied Energy*, **88**(10), 3524-3531.

Davis, R.E., Fishman, D.B., Frank, E.D., Johnson, M.C., Jones, S.B., Kinchin, C.M., Skaggs, R.L., Venteris, E.R., Wigmosta, M.S. 2014. Integrated evaluation of cost, emissions, and resource potential for algal biofuels at the national scale. *Environmental Science & Technology*, **48**(10), 6035-6042.

Dong, B., Ho, N., Ogden, K.L., Arnold, R.G. 2014. Cultivation of *Nannochloropsis salina* in municipal wastewater or digester centrate. *Ecotoxicology and Environmental Safety*, **103**, 45-53.

Farooq, W., Suh, W.I., Park, M.S., Yang, J.W. 2015. Water use and its recycling in microalgae cultivation for biofuel application. *Bioresource Technology*, **184**, 73-81.

Gerbens-Leenes, P.W., Xu, L., de Vries, G.J., Hoekstra, A.Y. 2014. The blue water footprint and land use of biofuels from algae. *Water Resources Research*, **50**(11), 8549-8563.

Gerber, L.N., Tester, J.W., Beal, C.M., Huntley, M.E., Sills, D.L. 2016. Target Cultivation and Financing Parameters for Sustainable Production of Fuel and Feed from Microalgae. *Environmental Science* & *Technology*, **50**(7), 3333-3341. Gibson, C. 1972. The algicidal effect of copper on a green and a bluegreen alga and some ecological implications. *Journal of Applied Ecology*, 9(2), 513-518.

Gross, M., Mascarenhas, V., Wen, Z.Y. 2015. Evaluating algal growth performance and water use efficiency of pilot-scale revolving algal biofilm (RAB) culture systems. *Biotechnology and Bioengineering*, **112**(10), 2040-2050.

Guieysse, B., Béchet, Q., Shilton, A. 2013. Variability and uncertainty in water demand and water footprint assessments of fresh algae cultivation based on case studies from five climatic regions. *Bioresoure Technology*, **128**, 317-323.

Hala, Y., Sjahrul, M. 2013. Biosorption of Zn2+ and Cd2+ in a twometal system by *Nannochloropsis salina*. *European Chemical Bulletin*, **2**(5), 238-241.

Heaton, E.A., Dohleman, F.G., Long, S.P. 2008. Meeting US biofuel goals with less land: the potential of Miscanthus. *Global Change Biology*, **14**(9), 2000-2014.

Huang, G.H., Chen, F., Kuang, Y.L., He, H., Qin, A. 2016. Current techniques of growing algae using flue gas from exhaust gas industry: a review. *Applied Biochemistry and Biotechnology*, **178**(6), 1220-1238.

Liu, X., Saydah, B., Eranki, P., Colosi, L.M., Greg Mitchell, B., Rhodes, J., Clarens, A.F. 2013. Pilot-scale data provide enhanced estimates of the life cycle energy and emissions profile of algae biofuels produced via hydrothermal liquefaction. *Bioresource Technology*, **148**, 163-171.

Martínez-Ruiz, E.B., Martínez-Jerónimo, F. 2015. Nickel has biochemical, physiological, and structural effects on the green microalga *Ankistrodesmus falcatus*: An integrative study. *Aquatic Toxicology*, **169**, 27-36.

Mehrabadi, A., Farid, M.M., Craggs, R. 2016. Variation of biomass energy yield in wastewater treatment high rate algal ponds. *Algal Research-Biomass Biofuels and Bioproducts*, **15**, 143-151.
Mishra, S.K., Suh, W.I., Farooq, W., Moon, M., Shrivastav, A., Park, M.S., Yang, J.-W. 2014. Rapid quantification of microalgal lipids in aqueous medium by a simple colorimetric method. *Bioresource technology*, **155**, 330-333.

Mohr, A., Raman, S. 2013. Lessons from first generation biofuels and implications for the sustainability appraisal of second generation biofuels. *Energy Policy*, **63**, 114-122.

Moody, J.W., McGinty, C.M., Quinn, J.C. 2014. Global evaluation of biofuel potential from microalgae. *Proceedings of the National Academy of Sciences*, **111**(23), 8691-8696.

Mortensen, L.M., Gislerod, H.R. 2016. The growth of *Chlorella* sorokiniana as influenced by CO2, light, and flue gases. *Journal of* Applied Phycology, **28**(2), 813-820.

Murray-Gulde, C., Heatley, J., Schwartzman, A., Rodgers Jr, J. 2002. Algicidal effectiveness of clearigate, cutrine-plus, and copper sulfate and margins of safety associated with their use. *Archives of Environmental Contamination and Toxicology*, **43**(1), 19-27.

Naik, S.N., Goud, V.V., Rout, P.K., Dalai, A.K. 2010. Production of first and second generation biofuels: A comprehensive review. *Renewable and Sustainable Energy Reviews*, **14**(2), 578-597.

Napan, K., Teng, L.H., Quinn, J.C., Wood, B.D. 2015. Impact of heavy metals from flue gas integration with microalgae production. *Algal Research-Biomass Biofuels and Bioproducts*, **8**, 83-88.

Nejat, P., Jomehzadeh, F., Taheri, M.M., Gohari, M., Abd. Majid, M.Z. 2015. A global review of energy consumption, CO2 emissions and policy in the residential sector (with an overview of the top ten CO2 emitting countries). *Renewable and Sustainable Energy Reviews*, **43**, 843-862.

Quinn, J.C., Yates, T., Douglas, N., Weyer, K., Butler, J., Bradley, T.H., Lammers, P.J. 2012. *Nannochloropsis* production metrics in a scalable outdoor photobioreactor for commercial applications. *Bioresoure Technology*, **117**, 164-71.

Quinn, J.C., Davis, R. 2015a. The potentials and challenges of algae based biofuels: A review of the techno-economic, life cycle, and resource assessment modeling. *Bioresource Technology*, **184**, 444-452.

Quinn, J.C., Davis, R. 2015b. The potentials and challenges of algae based biofuels: A review of the techno-economic, life cycle, and resource assessment modeling. *Bioresource Technology*, **184**, 444-452.

Quinn, J.C., Smith, T.G., Downes, C.M., Quinn, C. 2014. Microalgae to biofuels lifecycle assessment-multiple pathway evaluation. *Algal Research*, **4**, 116-122.

Rickman, M., Pellegrino, J., Hock, J., Shaw, S., Freeman, B. 2013. Life-cycle and techno-economic analysis of utility-connected algae systems. *Algal Research-Biomass Biofuels and Bioproducts*, **2**(1), 59-65.

Sims, R.E., Mabee, W., Saddler, J.N., Taylor, M. 2010. An overview of second generation biofuel technologies. *Bioresource Technology*, **101**(6), 1570-1580.

Singh, V., Tiwari, A., Das, M. 2016. Phyco-remediation of industrial waste-water and flue gases with algal-diesel engenderment from micro-algae: A review. *Fuel*, **173**, 90-97.

Stockenreiter, M., Haupt, F., Seppala, J., Tamminen, T., Spilling, K. 2016. Nutrient uptake and lipid yield in diverse microalgal communities grown in wastewater. *Algal Research-Biomass Biofuels and Bioproducts*, **15**, 77-82.

Strømgren, T., Sørstrøm, S.E., Schou, L., Kaarstad, I., Aunaas, T., Brakstad, O.G., Johansen, Ø. 1995. Acute toxic effects of produced water in relation to chemical composition and dispersion. *Marine Environmental Research*, **40**(2), 147-169.

Sun, A., Davis, R., Starbuck, M., Ben-Amotz, A., Pate, R., Pienkos, P.T. 2011. Comparative cost analysis of algal oil production for biofuels. *Energy*, **36**(8), 5169-5179.

Tu, Q.S., Lu, M.M., Thiansathit, W., Keener, T.C. 2016. Review of water consumption and water conservation technologies in the algal biofuel process. *Water Environment Research*, **88**(1), 21-28.

Venteris, E.R., Skaggs, R.L., Coleman, A.M., Wigmosta, M.S. 2013. A GIS cost model to assess the availability of freshwater, seawater, and saline groundwater for algal biofuel production in the United States. *Environmental Science & Technology*, **47**(9), 4840-4849.

Venteris, E.R., McBride, R.C., Coleman, A.M., Skaggs, R.L., Wigmosta, M.S. 2014a. Siting algae cultivation facilities for biofuel production in the United States: trade-offs between growth rate, site constructability, water availability, and infrastructure. *Environmental Science & Technology*, **48**(6), 3559-3566.

Venteris, E.R., Skaggs, R.L., Wigmosta, M.S., Coleman, A.M. 2014b. A national-scale comparison of resource and nutrient demands for algae-based biofuel production by lipid extraction and hydrothermal liquefaction. *Biomass & Bioenergy*, **64**, 276-290.

Wang, S.Y., Shi, X.G., Palenik, B. 2016. Characterization of *Picochlorum sp* use of wastewater generated from hydrothermal liquefaction as a nitrogen source. *Algal Research-Biomass Biofuels and Bioproducts*, **13**, 311-317.

Wijffels, R.H., Barbosa, M.J. 2010. An outlook on microalgal biofuels. *Science*, **329**(5993), 796-799.

Wuang, S.C., Khin, M.C., Qiang, P., Chua, D., Luo, Y.D. 2016. Use of *Spirulina* biomass produced from treatment of aquaculture wastewater as agricultural fertilizers. *Algal Research-Biomass Biofuels and Bioproducts*, **15**, 59-64.

Zhan, J.J., Zhang, Q., Qin, M.M., Hong, Y. 2016. Selection and characterization of eight freshwater green algae strains for synchronous water purification and lipid production. *Frontiers of Environmental Science & Engineering*, **10**(3), 548-558.

# X-ray Analysis of Lead(II) Binding to *Haloferax volcanii* Malate Synthase

Michael J. Adams and Bruce R. Howard

Southern Utah University

### Abstract

Elucidation of protein structures at the atomic level using X-ray crystallography is an effective technique for studying metabolic enzymes and defining binding interactions of small molecules and ions. Our research centers on the structure of the malate synthase isoform H (MSH) from Haloferax volcanii, which is a key enzyme in the glyoxylate pathway of cellular metabolism and allows this organism to integrate two carbon compounds for anabolic biosynthetic reactions. We have collected X-ray diffraction data from a protein crystal soaked in a solution containing lead(II) acetate. This heavy atom derivative provided phasing information using the single isomorphous replacement with anomalous scattering (SIRAS) method to determine the structure of the native enzyme, but this structure has not previously been analyzed and fully refined. Here we report the iterative model-building and refinement of this structure at 2.1Å resolution to an overall R-value of 0.1831 and an R-free of 0.2167. This structure allows a detailed analysis of lead(II) ion binding to the protein with implications for lead toxicity and inhibition of the enzyme. In addition to the displacement of the required magnesium ion and accompanying distortions in the local vicinity of the active site, we find three additional binding sites for lead ions. Strong peaks are observed at these lead binding sites in anomalous difference Fourier maps, and there are very high electron-density peaks in the 2Fo-Fc map at these four locations. Lead binding at intersubunit contacts may explain the increased resolution of X-ray diffraction from this derivative versus the native protein.

### Introduction

The enzyme malate synthase isoform H (MSH) is found in the organism *Haloferax volcanii*, a halophile of the domain Archaea. It allows *H. volcanii* to utilize two-carbon compounds such as acetate for anabolic biosynthetic reactions [1]. This is achieved through the glyoxylate pathway of cellular metabolism, which is a variation of the citric acid cycle and has not been confirmed to be found in mammals. MSH is able to function in highly saline conditions and has a relatively high composition of acidic amino acid residues, specifically aspartate and glutamate [2]. This enzyme is found in the native state as a trimer/hexamer equilibrium exhibiting  $C_3$  and  $D_3$  symmetry, respectively. An essential magnesium ion is bound within the active site of the native enzyme, and its functional role is to bind and polarize the glyoxylate substrate for catalysis [2].

Lead  $(Pb^{2+})$  is able to readily substitute for divalent cations such as  $Ca^{2+}$ ,  $Mg^{2+}$ , and  $Fe^{2+}$ [3], and this is thought to be an important factor in mediating its ionic mechanism of toxicity [4]. This substitution can have a wide variety of effects on biological processes depending on the specific context of the lead-binding protein involved [4,5,6,7]. In the case of replacing catalytically essential divalent ions, the effect is to render the enzyme inactive. This is pertinent to our model because of the required catalytic magnesium ion ( $Mg^{2+}$ ) in the active site.

The coordination of lead is a complex issue, and a variety of coordination polyhedra are formed, including tetrahedral, trigonal bipyramidal, octahedral, or pentagonal bipyramidal with either axial or equatorial vacancies [8]. A search of the Protein Data Bank (PDB) [9] reveals that lead binding to proteins seems to be highly variable regarding ligand-sphere symmetry but seems to be driven primarily by ionic interactions with key acidic residues and solvent ions. With the inherent structural disorder in protein structures compared with small-molecule structures, these examples display varying degrees of modeling accuracy. In general, the higher the resolution, the better, but even in highresolution crystal structures, there are regions within proteins that exhibit higher degrees of thermal motion or static disorder. This can be compounded by partial occupancy of ion-binding sites, which can only be fully modeled as contributions of bound and unbound structures. Given the limitations in the resolution of x-ray data from protein crystals, defining such partial contributions is challenging. To help address the accurate modeling of metal ion coordination to proteins, the Metalloprotein Database and Browser (MDB) has been developed at the Scripps Research Institute [10]. This software package can check metal-bound structures for validity to help address modeling issues and to hopefully improve the quality of these protein structures.

Assigning atomic identities to ions in an X-ray crystal structure can often be problematic. Electron density levels depend not just on the identity of a particular atom at a given location within the crystal, but also strongly on how often a particular site is occupied throughout the crystal (occupancy) and on how well ordered that particular atom is. The level of order of a particular atom is modeled using a temperature factor (B-factor). The higher the B-factor, the more disordered the atom is. However, the B-factor of an atom and its occupancy within the crystal lattice are highly correlated. Therefore, the B-factor must be interpreted in the context of the actual occupancy for a given site. Atoms within the core of a protein are typically the best ordered because they are tightly held in position by surrounding, closely packed residues, whereas those near the surface tend to be less constrained as side chains become more flexible and interact with solvent in different ways. Therefore, atoms at or near the surface tend to have higher B-factors compared with those in the core. As the atomic B-factors increase, their electron density is spread over a larger volume, resulting in peaks that are lower and broader, making it more difficult to position them with confidence. In extreme cases of flexibility and disorder, the electron density becomes uninterpretable or drops to such low levels that it essentially disappears from the maps. For these reasons, more specific information is often required than just the level of electron density at a particular position to assign a particular atom's identity with confidence.

In the case of our lead derivative data, which was collected using copper K-alpha radiation, we are fortunate because lead atoms diffract anomalously at this wavelength (1.54 Å). This means that there is important information present in reflections that are symmetrically related through the origin by inversion (h, k, l and -h, -k, -l; known as Friedel pairs or Bijvoet pairs). The difference in diffraction intensities and the corresponding amplitudes of the structure factors in a given pair result

from the atoms in the crystal that diffract anomalously. Therefore, if a map is calculated using structure factor amplitudes calculated from these differences and using the phases from a reliable model, peaks are built up at positions in the map that correspond to these anomalous atoms. The higher the anomalous scattering for a given species, the higher the peak height in the anomalous difference map. Of course, the same caveats discussed above, which relate to occupancies and Bfactors of atoms and the resulting effects on electron density levels, still apply. (For an excellent discussion of macromolecular crystallography, see reference 11). At the X-ray wavelength we used, lead scatters with about 8.5 anomalous electrons, while potassium scatters with only about 1.1 electrons, chloride with about 0.7 electrons, sulfur with about 0.57 electrons, and magnesium with less than 0.2 electrons [12]. Thus, lead atoms within the structure should stand out high above every other element in the crystal by forming high peaks in an anomalous difference map, allowing their positions to be identified (and their identities to be assigned unambiguously).

### **Materials and Methods**

### Cell Culture and Protein Purification

Cell cultures of *Haloferax volcanii* were grown in a chemically defined medium as previously described with acetate as the sole source of carbon to increase expression of MSH [2,13]. Restricting the carbon source to two-carbon compounds makes the glyoxylate cycle, and therefore malate synthase, indispensable for carbon assimilation. These conditions ensured that malate synthase was maximally expressed in the native strain. The enzyme was purified using reverse-phase, ion-exchange, and size-exclusion chromatography as previously described [2,13]. These serial chromatographic steps resulted in a sample of malate synthase that was of sufficient purity for single crystal growth. The final product was approximately 90% pure, and the yield was approximately 0.5 mg per liter of cell culture.

### Assay of Enzyme Activity

MSH activity was measured by following the removal of the acetyl-Coenzyme A substrate with a spectrophotometer as the reaction proceeded as previously described [2,13]. Each enzyme reaction was monitored at a wavelength of 232 nm to measure the rate of disappearance of the thioester absorption as the malyl-CoA intermediate was hydrolyzed in the second step of the enzymatic reaction.

### Crystallization

Crystals of MSH were grown in sitting drops using the technique of vapor diffusion at room temperature as previously described [2,13]. In each experiment, 2  $\mu$ l of protein solution containing 7 mg/mL of MSH, 13 mM MgCl<sub>2</sub>, 3 mM glyoxylate, 2 M KCl, and 50 mM Tris buffer pH 8.0 were mixed with 2  $\mu$ l of well solution containing 170 mM ammonium acetate, 24.5–27% w/v polyethylene glycol 4500, 15% glycerol, and 85 mM sodium acetate pH 4.4-5.0. This mixture was sealed in vapor contact with a reservoir containing 1 mL of well buffer and allowed to equilibrate at room temperature to drive crystal growth. Crystals typically grew over a period of two weeks producing diffraction-quality crystals of maximum dimensions 0.2 mm × 0.15 mm × 0.1 mm.

### Lead Soak

The lead(II) derivative was produced as previously described [2,13]. Briefly, 0.4  $\mu$ l of a saturated lead(II) acetate solution was added to the mother liquor containing the crystals and then resealed and allowed to equilibrate for one week at room temperature.

### X-ray Diffraction, Data Processing, and Structure Solution

A lead-soaked crystal was suspended in a nylon loop and flashcooled in liquid nitrogen as previously described [2,13]. Data were collected at a temperature of 100 K using an R-axis IV image plate and a Rigaku 007 HF rotating copper anode generator with Rigaku Varimax-HR X-ray optics (See Table 1 below). Data from the lead-soaked crystal were compared with data from a native crystal without lead. The technique of single isomorphous replacement with anomalous scattering (SIRAS) was used to solve the structure of the native protein as previously described [2,13], and an atomic model was manually built into the experimental electron density map. The model phases of the native protein were then used to phase the diffraction data from the lead-bound structure. Iterative rounds of model building using COOT [14] to position protein residues, small molecules, and ions into the electron density maps were alternated with molecular refinement using Refmac 5 [15] in CCP4 [16] to complete the structural model of the lead-bound protein reported here. All figures were generated using PyMOL [17]. The refined coordinates and structure factors have been deposited in the Protein Data Bank (PDB) [9] with accession ID 5TAO.

## **Results and Discussion**

### The Structure

MSH crystallized in the R32 rhombohedral space group, which is sometimes represented in the hexagonal setting (H32) for software packages and databases such as the PDB to enable computational convenience. The crystal contains one molecule per asymmetric unit. The enzyme is expected to exist as a biological trimer and hexamer; therefore, crystallographic symmetry operations around a threefold rotation axis result in a trimer with C<sub>3</sub> symmetry. A perpendicular twofold rotation axis produces a hexamer ( $D_3$  symmetry) from two back-to-back trimers. The refined model is comprised of residues 5-22, 25-282, 324-357, 387-432, one acetate molecule, three glyoxylate molecules, four lead ions, three potassium ions, four chloride ions, one sodium ion, and 254 water molecules. Two significant loops of residues 283-323 and 358-386 are excluded in the refined model as they are not visible in the electron density maps, presumably because of their flexibility and static disorder within the crystal. This is a common occurrence in X-ray crystal structures of proteins when surface loops or flexible domains are present. A comparison of our structure with the native protein structure in the absence of lead indicates that the same regions of the protein structure are disordered in both cases. Model statistics are shown in Table 1

### Evidence for Lead-Binding Sites

Strong peaks of electron density are shown at all four leadbinding sites in both the anomalous difference Fourier and 2Fo-Fc maps. The magnitudes of anomalous electron density at these sites are clearly distinguishable from other types of atoms. Table 2 notes the peak heights in the anomalous difference map in standard deviations in the calculated map. Lower-level peaks are observed at potassium binding sites (the best-ordered potassium is only 8.8 sigma peak height), some chloride (best ordered at 4.6 sigma), and fewer sulfur sites (best ordered at 5.0 sigma). The lower peak heights for Pb 503 and Pb 504 are due to their higher B-factors and/or lower occupancies, and they are therefore not as well defined in terms of solvent ligands. All four sites. however, indicate very clearly which amino acid residues contribute their side chains or backbone atoms to each lead ligand-sphere as discussed in detail below. Because it is a lead-soaked crystal, it is possible that the lead-binding sites are not fully occupied. To estimate the actual occupancy of each binding site, the occupancy of each site was adjusted to match the temperature factor (B-factor) of each lead(II) ion to

those of the protein ligands involved in the chelation of each ion. These estimated occupancies are listed in Table 3. These values indicate why Pb 501 and 502 provided strong phasing information in comparison to sites 503 and 504. Pb 501 exhibits one of the lowest B-factors and the highest occupancy. Although Pb 502 and 504 appear to have similar levels of occupancy, the B-factor of site 504 is double that of site 502. To maintain proper bond distances within the lead-binding sites in our final refined structure, the site occupancy of each was set to 1 (100%). This prevented side chains from being stretched into the electron density peak of the lead ion, as occurred when partial occupancies were set. Therefore, the partial occupancies are modeled as increased B-factors from the expected values listed in Table 3. This also allows for a more direct comparison of scattering intensity from each lead(II) ion, with the lower B-factors indicating a proportionally stronger contribution.

Table 1. Data collection and refinement statistics			
Crystallographic data			
Space Group	R32 (H32)		
Unit cell dimensions a, b, c (Å)	155.4, 155.4, 139.4		
α, β, γ (°)	90, 90, 120		
Resolution (Å)	30-2.10 (2.18-2.10)		
Total	271,479		
Unique	73,241 <sup>a</sup>		
Redundancy	3.7 (3.6) <sup>a</sup>		
Complete (%)	99.5 (100.0)		
R <sub>sym</sub> <sup>b</sup>	0.103 (0.374)		
$< I/\sigma(I)>$	11.5 (3.3)		
Model refinement statistics			
R <sub>work</sub> <sup>c</sup>	0.1831		
R <sub>free</sub> <sup>d</sup>	0.2167		
Average B-factor	49		
Total number of non-hydrogen atoms	3129		
R.m.s. deviations from ideal			
Bond lengths (Å)	0.0199		
Bond angles (°)	1.9156		

Values in parentheses are for the high-resolution shell.

<sup>a</sup>Friedel mates treated as independent reflections for anomalous phasing. <sup>b</sup>Rsym =  $\Sigma$ hkl |I -  $\langle I \rangle$ / $\Sigma$ hkl (I), where I is the observed intensity, and  $\langle I \rangle$  is the average intensity for multiple observations of symmetry related reflections. <sup>c</sup>Rwork is the Rfactor for 95% of data used during refinement, where Rfactor =  $\Sigma$ hkl ||Fo|-|Fc|| $\Sigma$ hkl|Fo|.

<sup>d</sup>Rfree is the Rfactor for 5% of the data not used in refinement.

Table 2. Anomalous difference map peak heights		
Atom	Magnitude (Sigma)	
Lead 501 (subunit interface-pair)	53.0	
Lead 502 (active site)	40.9	
Lead 503 (subunit interface-pair)	17.7	
Lead 504 (periphery)	17.7	

Table 3. Occupancy Estimates and Temperature Factors			
Atom	Estimated site- occupancy	<b-factor> of protein ligands</b-factor>	
Lead 501 (subunit interface-pair)	0.75	38.6	
Lead 502 (active site)	0.55	35.3	
Lead 503 (subunit interface-pair)	0.33	51.3	
Lead 504 (periphery)	0.54	72.7	

# Analysis of Lead Binding at the Active Site

Pb 502 is the lead ion that occupies the magnesium binding site in the enzyme active site. Upon binding, the lead ion induces a dramatic rearrangement of nearby residues, resulting in a distorted active site that binds the glyoxylate substrate in a reversed orientation and in a position that is expected to completely eliminate catalytic activity of this enzyme (Figures 1 and 2). Both Asp 192 and Glu 158, which coordinate the magnesium ion in the native enzyme, also bind tightly to



**Figure 1.** A comparison of magnesium vs. lead binding on active site side chains. (a)  $Mg^{2+}$ -bound structure. (b)  $Pb^{2+}$ -bound structure. The 35-sigma lead anomalous difference peak is shown in both panels. The  $Mg^{2+}$ -bound protein structure was superimposed onto the  $Pb^{2+}$ -bound structure using SSM in Coot. Carbon atoms are shown in light grey, oxygen in medium grey, and nitrogen in dark grey.



**Figure 2.** Lead binding at the active site showing the anomalous difference electron density map contoured at both 35- and 15-sigma levels. Carbons are light grey and oxygens are a darker grey.

Pb 502. Glu 51 in the magnesium-bound structure interacts orthogonally to the guanidine group of Arg 84 at a distance of 3.22 Å. In the lead-bound structure, it now forms an interaction with Pb 502 (2.98 Å). Although Glu 123 is too far away to interact with the bound magnesium ion, it shifts to coordinate Pb502 (2.60 Å). A glyoxylate molecule binds with magnesium in the native structure, but is flipped to a new orientation to bind Pb 502 (2.90 Å). Additionally, a second glyoxylate molecule coordinates Pb 502 that is not present in the native structure. In total, there are eight close interactions in the Pb 502 coordination sphere. These are summarized in Table 4.

Table 4. Binding specifics of Pb 502			
Amino acid/ligand	Distance from Pb 502 (Å)		
Aspartate 52	2.82		
Aspartate 192	2.18		
Glutamate 51	2.98		
Glutamate 123	2.60		
Glutamate 158 (bidentate)	2.98 and 2.98		
Glyoxylate 491	2.90		
Glyoxylate 493	1.98		

### Lead Binding at Subunit Interface

Pb 501 and Pb 503 comprise the double lead binding at the subunit interface. Pb 501 is the strongest lead-binding site of the four, whereas Pb 503 is quite weak based on both temperature factors and occupancies (see discussion above regarding site occupancies and Bfactors), as well as peak heights in the anomalous difference electron



**Figure 3.** Lead binding at the subunit interface. The anomalous difference map is contoured at 45-sigma (dark, bold), 35-sigma (dark, stippled), and 15-sigma levels. The peak at the Pb 503 site is only just visible in the 15-sigma map. Atom types are shaded as in Figure 2 but with carbon atoms of the neighboring subunit shown at the top in white.

density map (Figure 3, Tables 2 and 3). This site bridges two symmetry mates within the trimer and also stabilizes the surface that interacts to form the hexamer. Pb 501 bridges a symmetry molecule through Ser 221 with multiple conformations. There are nine close interactions in the Pb 501 coordination sphere, including the bidentate binding of gly-oxylate 492, which is also shared with Pb 503. Pb 503 features eight weaker interactions within its coordination sphere and does not bind directly to other symmetry-related molecules. The binding specifics of both Pb 501 and Pb 503 are summarized in Table 5.

Table 5. Binding specifics of Pb 501 and Pb 503			
Amino acid/ligand	Distance from Pb 501 (Å)		
Aspartate 196 (bidentate)	2.55 and 3.03		
Glutamate 161 (bidentate)	2.66 and 2.69		
Glyoxylate 492 (bidentate, bridging)	2.65 and 2.78		
Serine 221 (sym. mate) side chain O	3.03		
Serine 221 (sym. mate) carbonyl O	2.96		
Water 836	2.84		
Amino acid/ligand	Distance from Pb 503 (Å)		
Aspartate 90 (bidentate)	2.74 and 3.15		
Glutamate 161 (bridging)	3.44		
Glyoxylate 492 (bidentate, bridging)	3.13 and 3.23		
Water 870	3.20		
Water 1011	3.31		
Water 1013	3.13		

### Peripheral Lead-Binding Site

Pb 504 is a single, weak (See Tables 2 and 3) lead-binding site at the subunit periphery that is also just visible in an anomalous difference map contoured at 15 sigma (Figure 4). All but two of the seven interactions for the Pb 504 coordination sphere (Table 6) span normal distances. Water 997 is barely near van der Waals contact at a distance of 3.54 Å, although another water (996, at distance of 3.69 Å) in the vicinity is not. These sites may instead be chloride-binding sites with a very high level of disorder or they may simply be noise in the Fo-Fc map due to partial occupancy of the lead at this site. Pb 504 represents the most disordered lead-binding site within the model (although Pb 503 is comparable in anomalous peak height), and our confidence in this ligand-sphere only extends to the two aspartate residues that chelate the lead ion.



**Figure 4.** Lead binding at the periphery of the protein. The anomalous difference electron density map is contoured at 15-sigma. Atoms are shaded as in Figure 2.

Table 6. Binding Specifics of Pb 504		
Amino acid/ligand Distance from Pb 50		
Aspartate 344 (bidentate)	3.03 and 3.07	
Aspartate 348 (bidentate)	2.99 and 3.05	
Water 997	3.54	
Water 998	2.88	
Water 996	3.69	

### Inter-Hexamer Contacts in the Crystal

There is only a single type of crystal contact that holds all the hexamers together to make up the crystal. It involves a rare pi-stacking between two symmetry-related Arg 32 residues (Figure 5). The surrounding negatively charged residues and high salt concentrations of the crystallization buffer are likely factors in allowing this close van der Waals contact.



**Figure 5.** Pi-pi stacking interactions of symmetry-related arginine residues at the single type of crystal contact in the crystal. (a) Closest contacts are shown with interatomic distances (Å). (b) Surrounding acidic residues that help to screen the positive charges on the arginine side chains. Carbon atoms in residues from different protein subunits are shaded light (bottom) or very light grey (top), oxygens are a darker grey, and nitrogen atoms are shaded with the darkest grey.

### Comparison with Other Lead-Binding Proteins

As of the writing of this paper, there are 45 accessible entries in the Protein Data Bank (PDB) for structures that display lead binding. Three (PDB identifiers 2QD5, 2QKL, and 3EC8) were excluded from our analysis because they did not explicitly show coordinated lead binding to residues or describe coordination in associated literature. Four others (1NBS, 1TN1, 1TN2, and 1Y95) were excluded as they show lead binding to nucleotides instead of proteins. Of these 45 structures, 38 clearly define binding between lead ions and protein residues, making them appropriate candidates for a direct comparison to our model of lead-bound MSH. A summary of this data is contained in Table 7 (see references 6, 18-51).

Table 7. Specifics of Lead Binding to Common Residues				
Residue/	Number	Total	Mean	Median
molecule	Of	Bonds	Bonding	Bonding
	Residues	to Lead	Distance (Å)	Distance (Å)
	38 Analy	zed Structu	res in the PDB	
Aspartate	86	145	2.72	2.71
Glutamate	81	140	2.68	2.72
Serine	4	4	3.09	3.13
Water	69	69	2.74	2.77
Haloferax volcanii Malate Synthase Isoform H (MSH)				
Aspartate	6	11	2.60	3.01
Glutamate	4	8	2.54	2.98
Serine	1	1	3.03	3.03
Water	7	7	3.22	3.20
Glyoxylate	3	6	2.45	2.84

Overall, the residues aspartate (-0.12 Å), glutamate (-0.14 Å), and serine (-0.06 Å) have slightly closer mean bonding distances in MSH compared with the other models. Only water ligands (+0.48 Å) have a longer overall mean distance. But these seven water molecules include several with higher temperature factors and one with a bond distance that is longer than the sum of the van der Waals radii and are therefore not reliable in their precise positioning or even in their elemental identities. As discussed above, these may actually be chloride ions with very high B-factors or may perhaps simply represent positive noise peaks in the Fo-Fc electron density map caused by disorder or partial occupancies in these lead-binding sites with higher temperature factors (Pb 503 and Pb 504). However, it appears that the lead–ligand interactions to protein residues within the reported structure are comparable with those observed in these 38 previously described structures.

### Conclusion

The model of the lead-bound structure of malate synthase reported here has been refined to a respectable agreement with the observed xray diffraction data and with satisfactory stereochemistry and bond parameters. This refined model has allowed us to examine the detailed interactions of the lead-binding sites with protein side chains and provides an additional example of lead binding to proteins.

We find that the most important determinant for lead binding in this structure is the distribution of acidic residues at the solventaccessible surface of the protein. Where two or more acidic residues are closely positioned and oriented, lead binding may occur. But binding affinity is greatly affected by the number of oxygens with either full or partial negative charges that can cooperatively coordinate the lead ions with minimal bond distortions.

The displacement of the required catalytic magnesium ion in the active site with lead(II) and its accompanying rearrangement of the active site prevents proper binding of substrates and is a good example of how lead poisoning can disable enzyme activities that require divalent metal ions.

The tightest binding lead site (Pb 501) is found at an inter-subunit interface with chelating residues contributed from two different subunits. The bonding of this lead ion by Ser 221 by both its partially charged backbone carbonyl oxygen and its partially charged side-chain oxygen apparently tightens the interactions among trimers. This, in turn, may help to hold each hexamer in a more ordered state within the crystal lattice and may be responsible for the higher-resolution diffraction compared with the native protein. This is a convenient improvement in our particular case. Although this tightening of the protein lattice upon lead binding may not be a general occurrence with other proteins, it may be useful for other proteins that display a high number of acidic residues at their surface. Such acidic proteins are commonly found in halophilic organisms like H. volcanii [52,53]. Therefore, the use of lead derivatives in x-ray crystallography may not only be useful for phasing new structures but may also serve as a potential strategy to improve crystal quality and to achieve higher-resolution data when determining protein structures from these salt-loving creatures.

### Acknowledgments

We would like to thank Chris Hill at the University of Utah for the generous use of his X-ray diffraction equipment and Southern Utah University for providing computing resources and literature access.

### References

- Kornberg HL, Krebs HA: Synthesis of cell constituents from C2units by a modified tricarboxylic acid cycle. *Nature*, 1957, 179(4568):988-991.
- Bracken, CD, Neighbor, AM, Lamlenn, KK, Thomas, GC, Schubert, HL, Whitby, FG, Howard, BR: Crystal structures of a halophilic archaeal malate synthase from *Haloferax volcanii* and comparisons with isoforms A and G. *BMC Struct Bio*, 2011, 11:23.

- 3. Flora, G, Gupta, D, & Tiwari, A: Toxicity of lead: a review with recent updates. *Interdiscip Toxicol*, 2010, 5(2):47-58.
- 4. Jaishankar M, Tseten T, Anbalagan N, Mathew BB, Beeregowda KN: Toxicity, mechanism and health effects of some heavy metals. *Interdiscip Toxicol*, 2014, 7(2):60-72.
- Flora SJ, Mittal M, Mehta A: Heavy metal induced oxidative stress & its possible reversal by chelation therapy. *Indian J Med Res*, 2008, 128(4):501-23.
- 6. Kursula P, Majava V: A structural insight into lead neurotoxicity and calmodulin activation by heavy metals. *Acta Crystallogr Sect F Struct Biol Cryst Commun*, 2007, 1;63(Pt 8):653-6.
- 7. Ordemann, JM, Austin, RN: Lead neurotoxicity: exploring the potential impact of lead substitution in zinc-finger proteins on mental health. *Metallomics*, 2016, 1;8(6):579-88
- Davidovich RL, Stavila V, Marinin DV, Voit EI, Whitmire KH: Stereochemistry of lead(II) complexes with oxygen donor ligands. *Coord Chem Rev*, 2009, 253(9–10):1316–1352.
- 9. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE: The Protein Data Bank. *Nucleic Acids Res*, 2000, 28: 235-242.
- Castagnetto JM1, Hennessy SW, Roberts VA, Getzoff ED, Tainer JA, Pique ME: MDB: the Metalloprotein Database and Browser at The Scripps Research Institute. *Nucleic Acids Res*, 2002, 1;30(1):379-82.
- 11. Rupp B: *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology.* 2010, Garland Science, Taylor & Francis Group, New York.
- 12. Merritt, EA: Anomalous Scattering Coefficients Webtool, http://skuld.bmsc.washington.edu/scatter/AS\_form.html
- 13. Thomas G, Lamlenn K, Howard BR: Crystallization and preliminary x-ray diffraction of a halophilic archaeal malate synthase. *Amer J Undergrad Res*, 2009, 8(2 & 3):15-23.

- Emsley P, Lohkamp B, Scott WG, Cowtan K: Features and development of Coot. Acta Crystallogr Sect D Biol Crystallogr, 2010, 66(4):486-501.
- 15. Murshudov GN, Vagin AA, Dodson EJ: Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr Sect D Biol Crystallogr*, 1997, 53(Pt 3):240-255.
- 16. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr Sect D Biol Crystallogr*, 1994, 50(Pt 5):760-763.
- 17. DeLano Scientific; http://www.pymol.org
- Okvist, M, Dey, R, Sasso, S, Grahn, E, Kast, P, Krengel, U: 1.6A Crystal Structure of the Secreted Chorismate Mutase from Mycobacterium tuberculosis: Novel Fold Topology Revealed. J Mol Biol, 2006, 357:1483-1499. PDB ID 2FP1.
- Vassylyev, DG, Sekine, S, Laptenko, O, Lee, J, Vassylyeva, MN, Borukhov, S, Yokoyama, S: Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature*, 2002, 417:712-719. PDB ID 11W7.
- Arndt, JW, Hao, B, Ramakrishnan, V, Cheng, T, Chan, SI, Chan, MK: Crystal Structure of a Novel Carboxypeptidase from the Hyperthermophilic Archaeon Pyrococcus furiosus. *Structure*, 2002, 10:215-224. PDB ID 1KA4.
- Momany, C, Kovari, LC, Prongay, AJ, Keller, W, Gitti, RK, Lee, BM, Gorbalenya, AE, Tong, L, McClure, J, Ehrlich, LS, et al: Crystal structure of dimeric HIV-1 capsid protein. *Nat Struct Mol Biol*, 1996, 3:763-770. PDB ID 1AFV.
- Shumilin, IA, Kretsinger, RH, Bauerle, RH: Crystal structure of phenylalanine-regulated 3-deoxy-D-arabino-heptulosonate-7phosphate synthase from Escherichia coli. *Structure Fold Des*, 1999, 7:865-875. PDB ID 1QR7.
- 23. West Jr., AP, Llamas, LL, Snow, PM, Benzer, S, Bjorkman, PJ: Crystal structure of the ectodomain of Methuselah, a Drosophila G protein-coupled receptor associated with extended lifespan. *Proc Natl Acad Sci USA*, 2001, 98:3744-3749. PDB ID 1FJR.

- 24. Burkhard, P, Meier, M, Lustig, A: Design of a minimal protein oligomerization domain by a structural approach. *Protein Sci*, 2000, 9:2294-2301. PDB ID 1HQJ.
- 25. Main, E, Xiong, Y, Cocco, M, D'Andrea, L, Regan, L: Design of Stable alpha-Helical Arrays from an Idealized TPR Motif. *Structure*, 2003, 11:497-508. PDB ID 1NA0.
- Wilson, MA, Brunger, AT: Domain flexibility in the 1.75 A resolution structure of Pb2+-calmodulin. *Acta Crystallogr Sect D*, 2003, 59:1782-1792. PDB ID 1N0Y.
- 27. Georgiadis, MM, Luo, M, Gaur, RK, Delaplane, S, Li, X, Kelley, MR: Evolution of the redox function in mammalian apurinic/apyrimidinic endonuclease. *Mutat Res*, 2008, 643:54-63. PDB ID 2O3C.
- Taskinen, JP, van Aalten, DM, Knudsen, J, Wierenga, RK: High resolution crystal structures of unliganded and liganded human liver ACBP reveal a new mode of binding for the acyl-CoA ligand. *Proteins*, 2006, 66:229-238. PDB ID 2FJ9.
- 29. Myrick, LK, Hashimoto, H, Cheng, X, Warren, ST: Human FMRP contains an integral tandem Agenet (Tudor) and KH motif in the amino terminal domain. *Hum Mol Genet*, 2015, 24:1733-1740. PDB ID 4QW2.
- Gonzalez, B, Banos-Sanz, JI, Villate, M, Brearley, CA, Sanz-Aparicio, J: Inositol 1,3,4,5,6-Pentakisphosphate 2-Kinase is a Distant Ipk Member with a Singular Inositide Binding Site for Axial 2-Oh Recognition. *Proc Natl Acad Sci. USA*, 2010, 107:9608. PDB ID 2XAL.
- 31. Erskine, PT, Duke, EMH, Tickle, IJ, Senior, NM, Warren, MJ, Cooper, JB: MAD Analyses of Yeast 5-Aminolaevulinic Acid Dehydratase. Their Use in Structure Determination and in Defining the Metal Binding Sites. *Acta Crystallogr Sect D*, 2000, 56:421. PDB ID 1QNV.
- Hatada, MH, Lu, X, Laird, ER, Green, J, Morgenstern, JP, Lou, M, Marr, C, Phillips, TB, Ram, MK, Theriault, K: Molecular basis for the interaction of ZAP-70 with the T-cell receptor. *Nature*, 1995, 377:32-38. PDB ID 2OQ1.

- Morales, KA, Lasagna, M, Gribenko, AV, Yoon, Y, Reinhart, GD, Lee, JC, Cho, W, Li, P, Igumenova, TI: Pb2+ as Modulator of Protein-Membrane Interactions. *J Am Chem Soc*, 2011, 133:10599. PDB ID 3TWY.
- Park, SY, Borbat, PP, Gonzalez-Bonet, G, Bhatnagar, J, Pollard, AM, Freed, JH, Bilwes, AM, Crane, BR: Reconstruction of the Chemotaxis Receptor-Kinase Assembly. *Nat Struct Mol Biol*, 2006, 13:400. PDB ID 2CH7.
- 35. Thebault, P, Chirgadze, DY, Dou, Z, Blundell, TL, Elowe, S, Bolanos-Garcia, VM: Structural and functional insights into the role of the N-terminal Mps1 TPR domain in the SAC (spindle assembly checkpoint). *Biochem J*, 2012, 448:321-328. PDB ID 4H7X.
- Mourao, A, Nager, AR, Nachury, MV, Lorentzen, E: Structural Basis for Membrane Targeting of the Bbsome by Arl6. *Nat Struct Mol Biol*, 2014, 21:1035. PDB ID 4V0O.
- Schubert, M, Edge, RE, Lario, P, Cook, MA, Strynadka, NC, Mackie, GA, McIntosh, LP: Structural characterization of the RNase E S1 domain and identification of its oligonucleotidebinding and dimerization interfaces. *J Mol Biol*, 2004, 341:37-54. PDB ID 1SN8.
- Park, SY, Crane, BR: Structural insight into the low affinity between Thermotoga maritima CheA and CheB compared to their Escherichia coli/Salmonella typhimurium counterparts. *Int J Biol Macromol*, 2011, 49:794-800. PDB ID 3T8Y.
- Woo, EJ, Kim, YG, Kim, MS, Han, WD, Shin, S, Robinson, H, Park, SY, Oh, BH: Structural Mechanism for Inactivation and Activation of Cad/Dff40 in the Apoptotic Pathway. *Mol Cell*, 2004, 14:531. PDB ID 1V0D.
- Im, YJ, Raychaudhuri, S, Prinz, WA, Hurley, JH: Structural mechanism for sterol sensing and transport by OSBP-related proteins. *Nature*, 2005, 437:154-158. PDB IDs 1ZHW, 1ZHY, 1ZHZ.

- 41. Thirup, SS, Van, LB, Nielsen, TK, Knudsen, CR: Structural outline of the detailed mechanism for elongation factor Ts-mediated guanine nucleotide exchange on elongation factor Tu. *J Struct Biol*, 2015, 191:10-21. PDB ID 4PC1.
- 42. Kumar, S, Gourinath, S: Structure of Calcium Binding Protein-1 from Entamoeba histolytica in Complex with Lead. DOI 10.2210/pdb3qjk/pdb. PDB ID 3QJK.
- Bitto, E, Bingman, CA, Wesenberg, GE, McCoy, JG, Phillips, GN: Structure of pyrimidine 5'-nucleotidase type 1. Insight into mechanism of action and inhibition during lead poisoning. *J Biol Chem*, 2006, 281:20521-20529. PDB ID 2G0A.
- 44. Brillet, K, Meksem, A, Cobessi, D: Structure of the heme/hemoglobin outer membrane transporter ShuA from Shigella dysenteriae. DOI 10.2210/pdb3fhh/pdb. PDB ID 3FHH.
- 45. Voevodskaya, N, Galander, M, Hogbom, M, Stenmark, P, McClarty, G, Graslund, A, Lendzian, F: Structure of the high-valent FeIIIFeIV state in ribonucleotide reductase (RNR) of Chlamydia trachomatis--combined EPR, 57Fe-, 1H-ENDOR and X-ray studies. *Biochim Biophys Acta*, 2007, 1774:1254-1263. PDB ID 2ANI.
- 46. Gorman, MA, Seers, CA, Michell, BJ, Feil, SC, Huq, NL, Cross, KJ, Reynolds, EC, Parker, MW: Structure of the lysine specific protease Kgp from Porphyromonas gingivalis, a target for improved oral health. *Protein Sci*, 2015, 24:162-166. PDB ID 4TKX.
- Van Duyne, GD, Ghosh, G, Maas, WK, Sigler, PB: Structure of the oligomerization and L-arginine binding domain of the arginine repressor of Escherichia coli. *J Mol Biol*, 1996, 256:377-391. PDB ID 1XXA.
- Alayyoubi, M, Leser, GP, Kors, CA, Lamb, RA: Structure of the paramyxovirus parainfluenza virus 5 nucleoprotein -RNA complex. *Proc Natl Acad Sci USA*, 2015, 112:E1792-E1799. PDB ID 4XJN.

- 49. Kamtekar, S, Berman, AJ, Wang, J, Lazaro, JM, de Vega, M, Blanco, L, Salas, M, Steitz, TA: The phi29 DNA polymerase:protein-primer structure suggests a model for the initiation to elongation transition. *Embo J*, 2006, 25:1335-1343. PDB ID 2EX3.
- Hogbom, M, Stenmark, P, Voevodskaya, N, McClarty, G, Graslund, A, Nordlund, P: The radical site in chlamydial ribonucleotide reductase defines a new R2 subclass. *Science*, 2004, 305:245-248. PDB ID 1SYY.
- 51. Beernink, PT, Segelke, BW, Hadi, MZ, Erzberger, JP, Wilson III, DM, Rupp, B: Two Divalent Metal Ions in the Active Site of a New Crystal Form of Human Apurinic/Apyrimidinic Endonuclease, Ape1: Implications for the Catalytic Mechanism. J Mol Biol, 2001, 307:1023. PDB IDs 1E9N, 1HD7.
- 52. Lanyi JK: Salt-dependent properties of proteins from extremely halophilic bacteria. *Bacteriol Rev*, 1974, 38(3):272-290.
- Fukuchi S, Yoshimune K, Wakayama M, Moriguchi M, Nishikawa K: Unique amino acid composition of proteins in halophilic bacteria. *J Mol Biol*, 2003, 327(2):347-357.

# **Exploring Shifting Global Labor Management Practices and Comparative Job and Life Satisfaction**

# Bergen Eskildsen, Joe Light, Kaitlin Carlisle, and Jonathan H. Westover

Utah Valley University

### Abstract

According to recent studies, individuals in Nordic countries consistently enjoy greater levels of life satisfaction than people in other countries. Additionally, a growing body of research has shown that shifting global labor management practices have a large influence on changing worker attitudes and values. More specifically, various intrinsic and extrinsic motivators in the workplace have been shown to improve both worker job satisfaction and their global life satisfaction. This research utilizes comparative data from multiple waves of the World Value Survey to explore and examine the possible reasons for these country differences and to parse out the many variables that contribute to varying job satisfaction and life satisfaction levels across nations, with a specific comparative focus on Sweden, Belarus, and the United States.

### Introduction

### What Is Life Satisfaction and Why Study It?

It has been argued that the level of satisfaction that workers feel and obtain through work is directly correlated with higher production and efficiency (Böckerman & Ilmakunnas, 2012; Petty, Mcgee & Cavendar, 1984). Can life satisfaction really tip the production scale in any given industry? Why would this even matter? By searching out different types of motivators that influence work and life satisfaction, we can find the key motivators with positive influence in the workplace and in society as a whole.

Life satisfaction is an overall assessment of feelings and attitudes about one's life at a particular point in time, ranging from negative to positive. Happiness or life satisfaction is the degree to which an individual determines the overall quality of his/her life (OECD Better Life Index, n.d.). It is one of three major indicators of well-being: life satisfaction, positive affect, and negative affect (Diener, 1985). Although satisfaction with current life circumstances is often assessed in research studies, Diener et al. (2003) also include the following under life satisfaction: desire to change one's life; satisfaction with past; satisfaction with future; and significant other's views of one's life. Additionally, the positive or negative view that we have on our own lives has substantial impact on each of us (Yamasaki et al., 2011).

### **Literature Review**

### *Overview of Political, Economic, and Labor Management Practices in Belarus, Sweden, and the USA*

### Belarus

Many scholars have pointed to the extensive disrupting economic and social changes that followed the collapse of communism in the USSR in 1991 (e.g., Abbott and Wallace, 2010). Furthermore, despite its central geographic location and cultural and political history and current economic ties with its neighbors, to this point Belarus has largely been an understudied country, particularly within the context of the shifting political and economic landscape following the Soviet collapse.

The collapse of the Soviet Union was not the only cause of the many difficulties for Belarusian citizens. Additionally, Abbott and Wallace (2010) examined the economic collapse of the former Soviet Union on the lives of Belarusian citizens and found that a majority of the population have a poor quality of life and good reason to be dissatisfied with their lives. More specifically, they found that the collapse has not only resulted in a decline in economic security of households but also had an effect on social integration, social cohesion, and the ability of people to take control over their own lives, leaving many unable to develop capabilities to adequately function (see also Danilovich and Croucher, 2011). Moreover, Danilovich and Croucher (2011) found that "Labour management practices at enterprise level in Belarus are more negative for workers than those under the Soviet system. Welfare has largely disappeared, as has Soviet-style informal bargaining; wage payment may be in kind; training is minimal; job insecurity is extreme and trade unions perform a corporatist role" (2011, p. 241).

#### Sweden

In Sweden, it is perceived that all public power proceeds from the people. This is the foundation of parliamentary democracy in Sweden. Everyone has the same rights and is free to scrutinize how politicians and public agencies exercise their power. Sweden operates under the Riksdag, which is the Swedish Parliament. This political structure helps implement policies on behalf of its country's people. The government rules by implementing the policies and decisions that come from the Riksdag (Swedish.se, 2016).

Looking back on the past, Sweden in the mid-1850s was a poor agricultural-based county. Sweden saw substantial growth in the next 120 years, becoming one of the wealthiest countries in the world in the 1970s based on gross domestic product (GDP) per capita. With its strong infrastructure, Sweden has been able to adapt to changes in the industrialization processes and has had what it takes to keep long-term sustainable economic growth (Carlgren, 2007).

Labor management in Sweden has changed over the years as more competition has entered the marketplace. Because of the many changes, countries have had to adapt to the new climates. Sweden was one of those early movers in changing its policies of labor laws, its compensation system, and labor market regulations (Ulka & Muzi, 2015). To this day, Sweden manages to keep its workers both happy and extremely motivated, as will be supported by the research presented herein.

#### **United States**

The Constitution of the United States divides the federal government into three branches to ensure a central government in which no individual or group gains too much control: The Legislative branch makes laws (Congress), the Executive branch carries out laws (President, Vice President, Cabinet), and the Judicial branch evaluates laws (Supreme Court and other courts). All of these branches are critical in operating functionally (Usa.gov, 2016).

The United States is usually considered an economic powerhouse, but this has not always been the case. The United States has had its own struggles, just like any other country. During the Industrial Revolution, which took place between 1790 and 1820, the United States began to evolve from a rural to an urban society (Loc.gov, 2016), which brought much grief on the working class; this is one of the main reasons government intervention occurred. The United States government took it upon itself to establish rules and guidelines to help improve the working conditions of the workers, especially in industrial factories. These rules and many more are in place to protect the labor force.

With the roots that have been set from the industrial revolution and all the other laws and regulations that have been put into place, labor management has changed for the better. Now employees have more protection, better work environments, and more extrinsic/intrinsic motivators that keep them coming back to work day after day (Westover, 2013b). Employers have found that there are many ways to get the most out of their workers, mostly through motivators, as will be supported by the research presented herein.

### Methodology

The main source of data used in conducting this research is the World Values Survey. In brief summary:

The World Values Survey is a global network of social scientists studying changing values and their impact on social and political life, led by an international team of scholars, with the WVS association and secretariat headquartered in Stockholm, Sweden (World Values Survey, 1995–2014)

In addition to using the World Values Survey to compile data, charts, and graphs, articles from field anthropologists and other experts were used to gather information about the background of the differences in political, economic, and labor management practices among the three countries. Data from the World Happiness Report and OECD Better Life Index were used to compare the level of life satisfaction among Denmark, Belarus, and the United States.

### Results

# Comparing Life Satisfaction among Sweden, the United States, and Belarus

For years, economists and social scientists have used GDP as the standard metric for measuring the success and health of a country (e.g. Snyder, 1936; BostonFed, n.d.; FewResources, 2015). However, GDP falls quite short in measuring the social wellness of a country because it fails to take into account subjective well-being in response to economic activity in a country (Easterlin, 1974). Measures of subjective wellbeing can only be gathered through surveying/interviewing individuals about their own unique feelings towards their life and circumstance. Figures 1-4 show these results.



Figure 1. Belarus: Satisfaction with your life, 1995–2014. Source: World Values Survey



Figure 2. Sweden: Satisfaction with your life, 1995–2014. Source: World Values Survey.



Figure 3. United States: Satisfaction with your life, 1995–2014. Source: World Values Survey Source: World Values Survey.

To document life satisfaction, participants in the World Values Survey were asked the following question: "All things considered, how satisfied are you with your life as a whole these days?"

Participants were then asked to select their answer on a scale of 1 to 10 (1 being completely dissatisfied with life, 10 being completely satisfied (WVS Questionnaire, 1995–2014). For the purpose of this paper, and for easier readability and comparison; we changed the answers recorded from a 10-point scale to a dichotomous variable. Values 1 through 5 are categorized as 'unsatisfied', while values 6 through 10 are 'satisfied.' Also, participants who chose not to answer the question were not factored into the total percentage.

Additionally, to provide better understanding about what each country's level of satisfaction signifies, Figure 4 illustrates how the United States, Belarus, and Sweden compared in terms of satisfaction to the other countries surveyed around the world. It shows the average level of satisfaction of those surveyed for each country for the sixth wave of the World Value Survey (2010–2014). Our selected countries have been highlighted.



**Figure 4.** Global comparative life satisfaction, 2014. Belarus, Sweden, and United States are highlighted. (Data from the World Values Survey.)

Upon first glance at this graph, Sweden is among the countries with the higher averages of life satisfaction. Additionally, consistently, Sweden, Denmark, and other Nordic countries have been surveyed by the World Happiness Report (2015) as the countries with the happiest people. This cultural phenomenon has become increasingly noticed today, particularly with United States presidential candidate Bernie Sanders's very outspoken admiration for Denmark and its success as grounds for a socialism reformation in the United States (Tasini, 2015). But, according to the data, socialism is not the simple solution to the happiest people, seeing as Belarus is also a socialist country and has one of the lowest life satisfactions of the countries surveyed.

### Comparing Job Satisfaction Using Extrinsic and Intrinsic Factors between Denmark, United States, and Belarus

Wave 2 of the World Values Survey (WVS; data collected 1990-1993) showed that job satisfaction rates in Denmark and Sweden were the highest across the globe, with scores of 8.24 (sd=1.66) and 8.08 (sd=1.79), respectively. The United States had a score of 7.85 (sd=1.88), and Belarus had the lowest score of 6.10 (2.37) (Diener, Oishi, & Lucas, 2003). In the Fourth European Working Conditions Survey, Danish wage and salary earners were also ranked first in terms of work satisfaction, with Sweden falling close behind (Eurofound, 2007). In yet another study among 21 countries, Denmark has the highest rate of job satisfaction with Sweden ranked as number 9 and Norway as number 10. The United States ranked number 7 (Souza-Poza and Souza-Poza, 2000). Tables 1–3 show these results.

Table 1. Important aspects of job mentioned (Belarus)			
	1995-1998	1999-2004	Average
Extrinsic Factors			
Good Pay	81.6	86.9	84.25
Not Too Much Pressure	18.4	59.9	39.15
Good Job Security	54.5	36.7	45.6
Respected by People	38	42.1	40.05
Good Hours	33.3	58.5	45.9
Generous Holidays	27.5	53.2	40.35
Intrinsic Factors	•	•	
Opportunity to Use Initiative	21.7	39.8	30.75
You Can Achieve Something	31.1	42.3	36.7
A Responsible Job	16.3	27	21.65
A Job that Is Interesting	54.3	72.5	63.4
A Job that Meet's One's	39.8	49.5	44.65
Abilities			
Extrinsic Factors Total	42.22	56.22	49.22
Intrinsic Factors Total	32.64	46.22	39.42

Source: World Values Survey

Table 2. Important aspects of job mentioned (Sweden)				
	1995-1998	1999-2004	Average	
Extrinsic Factors				
Good Pay	59.3	58.2	58.75	
Not Too Much Pressure	22.7	35	28.85	
Good Job Security	53.1	51	52.05	
Respected by People	18.2	28.2	23.2	
Good Hours	37.6	44.4	41	
Generous Holidays	11.3	19.5	15.4	
Intrinsic Factors				
Opportunity to Use Initiative	51.1	51.9	51.5	
You Can Achieve Something	73.6	72.3	72.95	
A Responsible Job	52.8	54.6	53.7	
A Job that Is Interesting	71.1	69.9	70.5	
A Job that Meet's One's Abilities	33.2	41.1	37.15	
Extrinsic Factors Average	33.7	39.38	36.54	
Intrinsic Factors Average	56.36	57.96	57.16	

Source: World Values Survey

Intrinsic motivation can be defined as "the doing of an activity for its inherent satisfactions rather than for some separable consequence," whereas extrinsic motivation can be defined as "a construct that pertains whenever an activity is done in order to attain some separable outcome" (Ryan & Deci, 2000). Within the workplace, intrinsic and extrinsic factors of motivation can be applied to understand what makes a person work harder and more efficiently. For the purpose of this study, intrinsic and extrinsic factors of motivation were used to examine the connection that each type of motivation has with job satisfaction.

Table 3. Important aspects of job mentioned (United States)				
	1995-1998	1999-2004	Average	
Extrinsic Factors				
Good Pay	83.9	88.9	86.4	
Not Too Much Pressure	34	37.8	35.9	
Good Job Security	74.3	71.8	73.05	
Respected by People	39.6	45.5	42.55	
Good Hours	52	66.2	59.1	
Generous Holidays	26.2	36.6	31.4	
Opportunity to Use Initiative	52.4	61.7	57.05	
You Can Achieve Something	72.4	83.7	78.05	
A Responsible Job	50.6	54.8	52.7	
A Job that Is Interesting	72.8	81.5	77.15	
A Job that Meet's One's Abilities	54.9	61.8	58.35	
Extrinsic Factors Average	51.67	57.8	54.73	
Intrinsic Factors Average	60.62	68.7	64.66	

Source: World Values Survey

In waves 3 (1995–1998) and 4 (1999–2004) of the World Values Survey, the question was asked: "Here are some more aspects of a job that people say are important. Please look at them and tell me which ones you personally think are important in a job?" The percentages of respondents who mentioned whether a particular aspect of a job was important to them are calculated in Tables 1–3. Furthermore, it is represented in a dichotomous manner whether a specific aspect of job mentioned was an extrinsic factor or intrinsic factors of motivation within the workplace. The total average of extrinsic factors (e.g., good pay, good job security) and the total average of intrinsic factors (e.g., a job that is interesting, a job where you can achieve something) were also calculated to show a multi-country comparison between the relation of extrinsic factors of motivation and intrinsic motivation to overall job satisfaction for each country.'

From Tables 1–3, several theories and conclusions can be drawn about the relation of intrinsic and extrinsic motivators to overall job satisfaction, which will be discussed further below.

### Job Security (extrinsic factor)

One theory that can be drawn from this descriptive data is that because of the current governmental system in Sweden and the surrounding Nordic countries, employees in these Nordic countries do not need to put as heavy of an emphasis on the importance of extrinsic factors such as good pay and good job security. The flexicurity model that was constructed in Denmark and is being adopted by the surrounding Nordic countries is said to provide workers with labor market flexibility as well as job security for workers. The flexicurity model includes the benefits of unemployment insurance, balanced working hours and vacation, finding a new job, pensions: private and a labor market supplementary pension scheme (Oxford Research, Danish Chamber of Commerce, 2010).

In the United States, job security may have been mentioned at a higher percentage as an important aspect of a job because in the United States job security is more important to people since there is a more competitive job market that leaves many people lacking firm job security.

An increasing percentage of people who perceive job security as declining may provide a reason for employees in the United States to value job security at a higher rate. Job security may also be placed at a higher importance in the United States than in Sweden or Belarus because the United States experiences dramatic fluctuations in the state of its economy. While at times it may have a thriving economy, there have been significant economic recessions. Beginning in the 1970s, the United States experienced the mid 1970s recession, the 1980 recession, the early 1980s recession, and the early 1990s recession (McConnell & Perez-Quiroz, 1998). More recent economic recessions include the early 2000s recession and the great recession of 2008. Most of these economic recessions caused major increases in unemployment rates. For example, during the mid 1970s economic recession, the unemployment rate peaked at 15% in 1975. During the 1980 recession, unemployment was also at 11% (Krugman, 1994). This constant fluctuation of the economy may cause people living in the United States to place more importance on job security.

According to the CIA World Factbook, the unemployment rate in Belarus as of 2014 is 0.7% but it is noted that there is large number of underemployed workers who are not factored into the unemployment rate. The unemployment rate in Sweden as of 2014 is 7.9% and the unemployment rate in the United States as of 2014 is 6.2% (CIA World FactBook, 2014).

### Good Hours/Generous Holidays (extrinsic factors)

In the tables, "good hours" and especially "generous holidays" are mentioned as important job aspects at much lesser percentage in Sweden than in Belarus and the United States. In the Nordic countries, there is a greater emphasis on work/life balance than in other developed countries. This can be demonstrated through the shorter more flexible hours, the option to have home offices, greater trust in employees, inexpensive day-care, the possibility to take first day off when child is sick, guaranteed daycare, and paid maternity/paternity leave (Oxford Research, Danish Chamber of Commerce, 2010).

Furthermore, according to the OECD Better Life Index, Denmark ranks as the number one country for work/life balance, with a score of 9.8, and Sweden comes in close behind with a score of 8.1. The United States comes in towards the bottom of the rankings with a score of 5.3. On average, workers in Denmark spend 67% of their day or 16.1 hours on personal care and leisure. Workers in Sweden spend 63% of their day or 15.1 hours on personal care and leisure. In the United States, workers spend 60% of their day or 14.3 hours on personal care (OECD, 2015). Because of the attention to work/life balance that is in place in the Nordic countries versus those in the other countries, it can be inferred that a much smaller percentage of Swedish citizens feel the need to list good working hours and generous holidays as important aspects of a job when responding to the World Values Survey.

### Good Pay (extrinsic factor)

The United States and Belarus have a similar percentage of respondents mentioning "good pay" as an important aspect of a job, with 86.4% and 84.25%, respectively, whereas in Sweden only 58.75% of respondents mentioned "good pay" as an important aspect of a job. A possible conclusion that can be drawn from this is that the different percentage of tax rates in each country may have an impact on whether or not people regard good pay as an important aspect of their job.

Belarus has a flat rate of 12% for its tax residents and for the Belarusian income of nonresidents. There are no capital gains, stamp duty, and inheritance taxes for individuals (Belarus.by, n.d.). In contrast, in Sweden, local taxes range from 28.89% to 34.32%. For national income tax, there is none on income up to SEK 413,200, 20% on SEK 413,201–591,600 and 25% on income over SEK 591,600 (sweden.se, 2013). In the United States, people are taxed on brackets. The top marginal income tax is 39.6% for people with a taxable income of \$415,000 USD and higher. People at the lowest end of the making \$0– 9275 (single filers) and \$0–18,550 (married joint filers) pay 10% in taxes (TaxFoundation, 2015). Additionally, the United States and Belarus have significantly lower personal tax rates than Sweden. Sweden has one of the highest personal tax rates in the world, which may cause fewer people to place a high importance on good pay in a job (Wagstaff et al., 1999).

### Respected by People (extrinsic factor)

The United States and Belarus have a similar percentage of respondents mentioning "a job that is respected by people" as an important aspect of a job, with 42.55% and 40.05%, respectively, whereas in Sweden only 23.2% of respondents mentioned "a job that is respected by people" as an important aspect of a job. A possible conclusion that can be drawn from this data is that because Sweden and the other Nordic countries have a more collectivist culture and workplace where everybody is treated the same regardless of position, fewer people in Sweden place importance on having a job that is respected by people. The United States has more of an individualistic hierarchical culture where prestige and status in a hierarchical system are more prevalent, which may cause more people to place importance on having a job that is respected by people.

Having a job that is respected by people may also lead to higher job and life satisfaction. According to Maslow's hierarchy of needs, after first having met their physiological needs (e.g., hunger), people must then have safety (e.g., a harness), followed by love (e.g., affection: loving and receiving) before they can reach esteem (e.g., respect of others), which leads to self actualization (e.g., self fulfillment). A person at the top of Maslow's hierarchy of needs, who has reached self actualization, will have a desire for more challenging projects, opportunities for innovation and creativity, and enjoy learning at a high level.

### You Can Achieve Something/Opportunity to Use Initiative/A Responsible Job (intrinsic factors)

The United States and Sweden have relatively similar percentages of respondents mentioning "a job with opportunity to use initiative," "a job where you can achieve something," and "a responsible job" as important job aspects, with the United States being slightly higher on the first two but not "a responsible job." There is a striking difference between the results of these two countries on these three intrinsic factors and the results of Belarus. An explanation of why a significant percentage of Belarusian employees do not place these intrinsic factors as important aspects of a job could be that there is very little opportunity in Belarus compared with the other two countries. During the 1990s, Belarus became more dependent on Russia by strengthening its economic relationship with Russia. A treaty was signed between the two countries in April 1996 to support the reintegration of Belarus to Russia. In August 1998, because of this reintegration, Belarus was hit hard by the disastrous financial and economic crisis that occurred in Russia (Abdelal, 2001). Although between 1993 and 2003 there was a huge recovery of the GDP in Belarus, the economic potential Belarus has built is quickly crumbling down, causing it to fall behind its Western neighbors (BelarusInFocus, 2016). It has been said that the financial performance of the Belarusian economy leaves no hope for economic growth to resume (BelarusInFocus, 2016). This crumbling of the economy along with the lack of hope for future economic growth leaves the people of Belarus with little hope for general opportunity. "Labour management practices at enterprise level in Belarus are more negative for workers than those under the Soviet system. Welfare has largely disappeared, as has Soviet-style informal bargaining; wage payment may be in kind; training is minimal; job insecurity is extreme and trade unions perform a corporatist role." (Danilovich & Croucher, 2011). Because of the lack of general opportunity in Belarus, intrinsic motivational factors of a job such as a job where you can achieve something, use initiative, or have a responsible job are placed on less of an important degree and would not be particularly salient aspects of a job for Belarusians.
# A Job That Is Interesting (intrinsic factor)

The United States had the highest percentage of respondents who mentioned "a job that is interesting" as an important job aspect with 78.05%, followed by Sweden with 70.05%, and Belarus with 63.4%. The reason that Belarus may have a much lower percentage of people who value an interesting job as an important job aspect may be because many of the same reasons mentioned for the other intrinsic factors. Because there is less general opportunity in Belarus than in the other two countries, Belarusian employees may place less of an importance on having an interest in the work they do. With the high level of job insecurity and because labor management practices have become more negative, jobs in Belarus become more of a means to an end to achieve financial stability (Danilovich & Croucher, 2011).

#### Hard Work Brings Success

In waves 5 (2005–2009) and wave 6 (2010–2014) of the World Values Survey, respondents were asked to assess whether hard work or luck and connections bring success. Participants responded on a scale from 1 to 10 (1 being "In the long run, hard work usually brings a better life," 10 being "Hard work doesn't generally bring success- it's more a matter of luck and connections"). For the purpose of this paper, and for easier readability and comparison; we changed the answers recorded from a 10-point scale to a dichotomous variable. Values 1 through 5 are categorized as 'Hard work brings success,' while values 6 through 10 are categorized as 'hard work does not bring success.' Figures 5–7 show these results.



Figure 5. Belarus: Hard work brings success, 1995–1999, 2005–2014. Source: World Values Survey.



Figure 6. Sweden: Hard work brings success, 1995–1999, 2005–2014. Source: World Values Survey



**Figure 7.** United States: Hard work brings success, 1995–1999, 2005–2014. Source: World Values Survey

It is interesting to observe from these tables that in the survey, the United States has the highest percentage of respondents who believe that hard work brings success. A conclusion that can be drawn from this is that because of the capitalist free market system in America, people in the United States may have a greater belief in the "American Dream" of prosperity and success by working hard. Another interesting note is that the data from all three countries show a substantial increase in people who believe that hard work brings success. This could be due to the fact that there is a generational gap between the respondents in the survey in which the later generation has an increase in people who believe hard work brings a better life. Belarus has the most dramatic change in results among the three countries between Wave 3 and Wave 6. Another possible conclusion from this is that after the Soviet collapse of 1991 there was a shift in work attitudes and values among Belarusian workers (Westover, 2013a). This shift may have caused fewer people to believe that hard work brings success in Wave 3 (1995–1998), during the most turbulent government and economic transition year, than in Wave 6 (2010–2014).

# Discussion

The United States has the highest total average of extrinsic factors mentioned as important aspects of a job, with an average score of 54.73. The United States has the highest score on all extrinsic factors of important job aspects mentioned except for "generous holidays" and "not too much pressure" (where it falls close behind Belarus). The United States also has the highest total average of intrinsic factors mentioned as important aspects of a job with an average score of 64.66.

In contrast, Belarus has the lowest total average of intrinsic factors mentioned as important aspects of a job, with an average score of 39.42. Belarus also has the lowest scores on all intrinsic factors of important job aspects mentioned except for "a job that meets one's abilities."

Sweden, a Nordic country with the highest overall life satisfaction and overall job satisfaction among the three countries, has the lowest total average of extrinsic factors that people mentioned as important aspects of a job between the three countries, with an average score of 36.54. Sweden has the lowest scores on all extrinsic factors of important job aspects mentioned except for "good job security" (where Belarus falls shortly behind.)

Based on these research findings, the Nordic countries, including Sweden, do consistently enjoy greater levels of life satisfaction and job satisfaction (consistent with the findings of Rice, Near, & Hunt, 1980; Schmitt & Pulakos, 1985) than other countries around the world. Using information about life satisfaction from the World Values Survey, this study confirmed what several other studies and a growing body of academic literature have mentioned about the high levels of life/job satisfaction in the Nordic countries.

Based on our research findings, it is clear that various intrinsic and extrinsic motivators in the workplace environment have a significant impact on improving the overall job satisfaction of workers, as well as the life satisfaction of citizens in a country. Extrinsic and intrinsic factors of motivation can play a key role in predicting the overall level of job satisfaction in a country because they help determine why and how an employee is driven and finds meaning and value in work. This is consistent with the findings of many studies that show there is a direct correlation between employee motivation and satisfaction, which in turn increases employee performance (e.g., Tietjen & Myers, 1998; Locke & Latham, 1990; Roos & Eeden, 2008). To further the validity of this study, several more countries could be added.

The results presented herein can be used by businesses and corporations to help increase employee motivation and job satisfaction by identifying and making changes to the different extrinsic and intrinsic motivators within their workplace environment. This study, with further exploration, could also be used by government employees and policy makers to determine the best and worst causes and effects of certain public policy implementations on life satisfaction and worker satisfaction.

# Limitations

In the early stages of the research, it was originally intended for Denmark to be used in the comparisons as opposed to Sweden. Denmark was intended to be used as it is the single country with the highest overall life satisfaction and also because of how Denmark has been in terms of overall satisfaction culture. Since Denmark is not included in recent data on the World Values Survey, Sweden was used instead, at it has a similar cultural climate as Denmark and is in close proximity geographically and in overall satisfaction scores. Additionally, although Sweden and Denmark are very similar in regards to social, political, economic, and labor management practices, the two countries have very different immigration approaches, which may play a part in difference in life satisfaction that does exist between the two countries.

# Conclusion

Results presented herein demonstrate that the countries with the highest job satisfaction also have the highest life satisfaction. In regards to job satisfaction, it can be concluded that countries where employees most frequently mention extrinsic factors as important aspects of a job have a lower job satisfaction, whereas countries who most frequently mention intrinsic factors as important aspects of a job have a higher job satisfaction.

Public policy and form of government play a notable role in influencing intrinsic and extrinsic factors of motivation in regards to job satisfaction. However, from the results presented herein, it cannot be inferred that any specific form of government completely determines the level of life satisfaction in the country, based upon the fact that both Belarus and Sweden are socialist countries and Sweden is a country with one of the highest overall life satisfaction, while Belarus is a country with one of the lowest overall life satisfaction levels.

It can also be concluded that the different fluctuations of the economy, including recessions and rises, as well as different tax rates in each country do have an impact on the job and life satisfaction of people in different countries. The level of unemployment rates in each country also affect different factors of job motivation such as having a secure job, which may be attributed to better job satisfaction.

Furthermore, the work/life balance and family life outside of work, as demonstrated by the flexicurity model and shorter working hours in Sweden, has a significant impact on the happiness of employees. It can be inferred that having a better work/life balance and family life outside of work leads to greater employee satisfaction and overall life satisfaction, which is consistent with the findings of other studies on work-life balance (e.g., Byrne, 2005; Valcour, 2007).

Whether a country has more of a collectivist culture or an individualistic culture also plays a role in the difference in extrinsic and intrinsic factors that employees regard as important job aspects. This was demonstrated in our findings by the fact that in places with a collectivist culture, employees feel less of a need to be in a position of power to be respected by people and thus reach self actualization. Furthermore, it has been noted that this difference in collectivist and individualistic culture may not be limited to only being regulated by various forms of government, as corporations in places with a more individualistic culture, such as the United States, can create more of a collectivistic workplace culture at their own discretion.

This study touched upon findings that may affect both the public sector and the private sector. While various studies should be carefully considered and evaluated before being used to implement policy, whether it be company policy in the private sector or legislative policy in the public sector, this study offers a deeper look into how the difference in extrinsic and intrinsic factors of motivation are influenced by laws and governmental policy and how they can be used to improve both the lives of people in various sectors across different country contexts.

# References

Abbott, P., & Wallace, C. (2010). Explaining Economic And Social Transformations In Post-Soviet Russia, Ukraine And Belarus. *European Societies*, *12*(5), 653-674. doi:10.1080/14616691003786350

Abdelal, R. (2001). National Purpose in the World Economy: Post-Soviet States in Comparative Perspective. Ithaca: Cornell University Press

Belarus.by. (n.d.). Taxation in Belarus. Retrieved May 05, 2016, from http://www.belarus.by/en/invest/investment-climate/taxation

BelarusInFocus. (2016, May 26). Financial performance of Belarusian economy leaves no hope for economic growth to resume. Retrieved June 18, 2016, from http://belarusinfocus.info/economy/financial-performance-belarusian-economy-leaves-no-hope-economic-growth-resume

Böckerman, P., & Ilmakunnas, P. (2012). The job satisfactionproductivity nexus: a study using matched survey and register data. ILR Review, 65(2): 244-262. doi:10.1177/001979391206500203

BostonFed. (n.d.). How Do We Measure "Standard of Living" Retrieved May 17, 2016. https://www.bostonfed.org/-/media/ Documents/ledger/ledger2003/measure.pdf

Byrne, U. (2005). Work-life balance: Why are we talking about it at all? *Business Information Review*, 22(1): 53. doi:10.1177/0266382105052268

Carlgren, Fredrik. (2007). *Agricultural toward Industrial*. Retrieved May 15, 2016, from the Ekonomifakta website: http://www.ekonomifakta.se/en/Swedish-economic-history/swedish-economic-history/Agricultural-toward-Industrial/

Central Intelligence Agency. (n.d.). Retrieved May 10, 2016, from https://www.cia.gov/library/publications/the-world-factbook/

Danilovich, H., & Croucher, R. (2011). Labour Management in Belarus: Transcendent Retrogression. *Journal of Communist Studies and Transition Politics*, 27(2): 241. doi:10.1080/13523279.2011.64095

Diener, E. D., Robert A. Emmons, Randy J. Larsen, and Sharon Griffin. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, 49(1): 71-75. Diener, E., Oishi, S., & Lucas, R. E. (2003). Personality, culture, and subjective well-being: Emotional and cognitive evaluations of life. *Annual Review of Psychology*, 54:403-425

Easterlin, R. A. (1974). Does Economic Growth Improve the Human Lot? Some Empirical Evidence. In *Nations and Households in Economic Growth* (ed. P.A. David and M.W. Reder), pp. 89-125. Elsevier, Philadelphia.

Eurofound. (2007, December 3). Fourth European Working Conditions Survey. Retrieved May 4, 2016, from http://www.upf.edu/cisal/\_pdf/ doc8.pdf

Few Resources. (2015). GDP & Well-being. Retrieved May 18, 2016, from http://www.fewresources.org/gdp--well-being.html

Krugman, P. (1994). Past and Prospective Causes of High Unemployment. Federal Reserve Bank of Kansas City, Economic Review, Fourth Quarter, 23-43.

Loc.gov. (2016). The Industrial Revolution In The United States. Retrieved May 23, 2016, from the official Library of Congress website: http://loc.gov/teachers/classroommaterials/primarysourcesets/industrial -revolution/pdf/teacher\_guide.pdf

Locke, E. A., & Latham, G. P. (1990). Work Motivation and Satisfaction: Light at the End of the Tunnel. *Psychological Science*, 1(4): 240-246. doi:10.1111/j.1467-9280.1990.tb00207.x

McConnell, M. M., & Perez-Quiros, G. (1998). Output Fluctuations in the United States: What Has Changed Since the Early 1980s? *SSRN Electronic Journal*. doi:10.2139/ssrn.938810

OECD Better Life Index. (n.d.). Life Satisfaction. Retrieved May 22, 2016, from http://www.oecdbetterlifeindex.org/topics/life-satisfaction/

OECD Better Life Index. (2015). Work-Life Balance. Retrieved May 10, 2016, from http://www.oecdbetterlifeindex.org/topics/work-life-balance/

Oxford Research. (2010). Living and Working in Denmark. An Expat Perspective. Retrieved May 4, 2016, from http://www.oxfordresearch. dk/media/48201/451880-GuidetilExpats.pdf

Petty, M. M., Mcgee, G. W., & Cavender, J. W. (1984). A Meta-Analysis of the Relationships between Individual Job Satisfaction and Individual Performance. *The Academy of Management Review*, 9(4): 712-721. doi:10.2307/258493

Rice, R., Near, J., & Hunt, R. 1980. The Job-Satisfaction/Life-Satisfaction Relationship: A Review of Empirical Research. *Basic & Applied Social Psychology*, 1(1): 37-64.

Roos, W., & Eeden, R. V. (2008). The relationship between employee motivation, job satisfaction and corporate culture. *SA Journal of Industrial Psychology* 34(1): 54-63. doi:10.4102/sajip.v34i1.420

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, 25(1): 54-67. doi:10.1006/ceps.1999.1020

Schmitt, N., & Pulakos, E. 1985. Predicting Job Satisfaction from Life Satisfaction: Is There a General Satisfaction Factor? *International Journal of Psychology*, 20(1): 155-167.

Snyder, C. (1936). The Capital Supply and National Well-Being. *The American Economic Review*, 26(2): 195-224.

Sousa-Poza, A., & Sousa-Poza, A. A. (2000). Well-being at work: A cross-national analysis of the levels and determinants of job satisfaction. *The Journal of Socio-Economics*, 29(6): 517-538. doi: 10.1016/s1053-5357(00)00085-8

Sweden.se. (2013). Swedes and taxes. Retrieved May 05, 2016, from https://sweden.se/society/why-swedes-are-okay-with-paying-taxes/

Swedish.se. (2016). The Swedish System of Government. Retrieved May 15, 2016, from the official website of Sweden: https://sweden.se/society/the-swedish-system-of-government/

Tasini, J. (2015). *The essential Bernie Sanders and his vision for America*. Chelsea Green Publishing, White River Junction, VT.

TaxFoundation. (2015, October 14). 2016 Tax Brackets. Retrieved May 05, 2016, from http://taxfoundation.org/article/2016-tax-brackets

Tietjen, M. A., & Myers, R. M. (1998). Motivation and job satisfaction. *Management Decision*, 36(4): 226-231. doi:10.1108/ 00251749810211027

Ulku, Hulya; Muzi, Silvia. 2015. Labor market regulations and outcomes in Sweden: a comparative analysis of recent trends. Policy Research working paper; no. WPS 7229. World Bank Group, Washington, DC.

Usa.gov. (2016). The United States of America Government. Retrieved May 15, 2016, from the official United States website: https://www.usa.gov/branches-of-government

Valcour, M. (2007). Work-based resources as moderators of the relationship between work hours and satisfaction with work-family balance. *Journal of Applied Psychology*, 92(6): 1512-1523. doi:10.1037/0021-9010.92.6.1512

Westover, Jonathan H. (2013a). Political, Economic, and Social Transformations in Post-Soviet Belarus, 1990-2012: A Research Brief and Overview. *Global Journal of Management and Business*, 13(4): 27-29.

Westover, Jonathan H. (2013b). The Impact of Comparative State-Directed Development on Working Conditions and Employee Satisfaction. *Journal of Management and Organization*, 19(4): 498-518.

WHR. (2015). World Happiness Report 2015. Retrieved May 15, 2016, from http://worldhappiness.report/wp-content/uploads/sites/2/ 2015/ 04/ WHR15.pdf

World Values Survey. (n.d.). WVS Database. Retrieved May 15, 2016, from http://www.worldvaluessurvey.org/WVSContents.jsp

World Values Survey. (1995-2014). WVS Database Wave 6 Questionnaire. Retrieved May 18, 2016, from http://www. worldvaluessurvey.org/WVSDocumentationWV6.jsp

Yamasaki, K., Sasaki, M., Uchida, K., & Katsuma, L. (2011). Effects of positive and negative affect and emotional suppression on short-term life satisfaction. *Psychology, Health & Medicine*, 16(3): 313-322. doi:10.1080/13548506.2011.554564

# **Corporate Payouts and Macroeconomic Influences**

Lauren Lo Re,<sup>1</sup> Robert Patterson,<sup>1</sup> and Mahfuz Raihan<sup>2</sup> <sup>1</sup>Westminster College and <sup>2</sup>University of Utah

#### Abstract

Corporate payouts are economically significant and have fluctuated dramatically for decades. Several researchers have documented these trends and waves in payout activity; however, there has been little explanation for why these patterns occur. The existing literature focuses on examining cross-sectional variation in firm-level characteristics as the basis for explaining payout policy, yet questions remain. In this paper, the role of macroeconomic factors in the payout equation is examined. Results indicate that payout activity fluctuates with changes in macroeconomic factors and macroeconomic factors contribute to an explanation of trends in payout level and form of payout.

# **1. INTRODUCTION**

Corporate payout activity results from firm-level decisions and incorporates a broad range of issues and influences. Payout decisions (paying dividends and repurchasing stock) reflect investment opportunities, future outlook, and financing decisions, as well as other issues, including financial flexibility and even management compensation. Payout decisions impact stock prices, engage investors in the decision-making structure, are the result of many decisions, and are influenced by many factors. It is not surprising that following decades of work, there is not one overall theory of corporate payouts, yet, practical and academic motivations for further work remain.

Practical motivations primarily relate to how corporate payout decisions impact investors and stock prices. Payout decisions are made on an ongoing basis, within increasingly complex environments. It is no surprise that an informal review of business strategy textbooks, websites, and case studies includes the macroeconomic environment as a significant external factor to be considered when designing business strategy and decision-making models. Yet, payout research tends to assume macroeconomic influences are exogenous to the payout equation. This work questions that assumption.

Academic motivations are based on a theoretical framework that includes two primary areas of focus: firm-level determinants of payout activity and firm value relative to payouts. Earnings have consistently emerged as a primary determinant of payouts. The Lintner (1956) model reflects this in the following equation, in which current dividends are explained by earnings and previous dividends:

# $D_{i,e} = \alpha + \beta \operatorname{Earning}_{i,e} + \beta D_{i,e-1} + \varepsilon_{i,e}$ (1.1)

In this model, the dependent variable is current dividends, noted as  $D_{1t}$ . Previous dividends (lag 1) are noted as  $D_{1t-1}$ . Earning<sub>tt</sub> denotes current earnings, followed by the error term  $g_{1t}$ .

Firm value relative to payouts is based on the Gordon Growth Model (Gordon, 1959), which suggests that the value of a stock is based on the future dividend payments. Stock values, as reflected in stock returns, are typically used as leading economic indicators based on the assumption that they reflect expectations about future economic activity; however, the intuition is incomplete. Intuition suggests that stock returns, driven by future dividend payments, predict aggregate economic activity because dividends grow when the economy grows. Empirical evidence does not provide consistent support for this view. Cochrane's (2011) recent findings support Fama (1990) and suggest that there is a significant link between aggregate output growth and stock returns, but not between dividend growth and stock returns. These conflicted findings support the need for further investigation.

In addition to Lintner (1956) and Gordon (1959), explanations for payouts have evolved over time and incorporated the need to balance

the risk and benefits of retaining excess funds and paying them out to shareholders. Research suggests that earnings, profitability, and investment are the primary determinants of corporate payout policy. Following Lintner (1956), three theories are frequently cited that capture the importance of earnings and the tension between payout and retention: the agency cost theory (Jensen, 1986) and the pecking order theory of capital structure (Myers and Majluf, 1984), and the life-cycle theory of corporate payouts (Grullon, Michaely and Swaminathan, 2002; DeAngelo and DeAngelo, 2006). The agency cost theory of corporate payouts (Easterbrook, 1984; Jensen, 1986) suggests that managers may retain excess cash to benefit themselves at the expense of shareholders, and they may adjust corporate decisions to advance their own interests. An expressed concern (Jensen and Meckling, 1976) is that when a firm's excess capital is more than enough to finance current investments, managers could use cash to overinvest in risky projects or increase their own compensation, instead of returning extra cash to shareholders. Myers and Majluf (1984) offer a different perspective that emphasizes retention, as opposed to payout. They focus on how payout decisions impact capital structure and financing decisions and suggest that firms prioritize sources of financing based on the principle of least resistance. The result is a pecking order that encourages the use of internal funds first, followed by debt, and finally by equity. Internal funds are obtained by retaining free cash flow and limiting payouts. The result is that there is no need to issue additional equity. Myers and Majluf (1984) encourage retention to avoid flotation costs, as well as the economic dilution that results from issuing additional equity. The more recently developed life cycle theory of payouts (Grullon, Michaely, and Swaminathan, 2002; DeAngelo and DeAngelo, 2006), considers the time-varying costs and benefits of payouts and retention as the foundation of a more comprehensive explanation of payout behavior. The focus remains on the need to deliver free cash flow to shareholders, but additional issues are taken into account. This explanation considers together the managerial incentives for payout and the capital requirements for investment as the basis of an optimal payout policy. This theory allows that the tradeoffs between retention and payout evolve and change over time as the firm's earnings and investment opportunity set change.

Although limited, some existing work examines aggregate payout activity within the context of the overall economy. This provides further guidance on factor and model selection. Rau and Stouraitis (2009) examine stock repurchases using Vector Autoregressive (VAR) methods. They find distinct patterns in aggregate corporate payout activity. Dittmar and Dittmar (2004) also examine the stock repurchase equation and use a Vector Error Correction Model (VECM). They find a cointegrating relation between earnings and gross domestic product (GDP). Chen, Roll, and Ross (1986) and Lettau and Ludvigson (2001) find macroeconomic variables systematically affect stock market returns, which are also linked to earnings and dividends. Serfling and Miljkovic (2011) examine how macroeconomic factors impact dividend yield and find significant interactions with industrial production index (IPI), consumer price index (CPI), money supply (M1), and the S&P 500 index level and short-term Treasury rates.

This research contributes to the existing literature in several ways. First, the common methods used in payout literature are multivariate regression models and logit models. This work analyzes payout activity using a VAR/VECM, which has been used as previously noted. This approach follows Enders (2010) and allows for the analysis of the dynamics of short-run and long-run relations, as well as the simultaneous interactions among the variables of interest. A second contribution is that this work includes total payouts, total dividends, and total repurchases. Previous work has not included all three payout variables and has not provided a comprehensive view of payout activity.

The purpose of this study is to examine whether macroeconomic variables impact payout policy. This research examines this issue within a very simple context using earnings as the only firm-level control variable. By examining permanent and transitory effects, these results indicate that macroeconomic factors influence corporate payouts in this model. This work begins where many studies do and assumes that earnings are a primary determinant of payouts and that aggregate earnings and macroeconomic variables share a long-term trend. We estimate a cointegrating relation between payout variables and several macroeconomic variables from 1984 to 2011. The following two research questions are addressed: Do total corporate payouts and macroeconomic variables share long-term relations have significant long-term and/or short-term relationships?

The analysis is done by estimating cointegrating vectors using the methods of Johansen and Juselius (1990) and including the following factors. Payout activity is measured by total dividends, total repurchases, and total payout, which is the sum of total dividends and total repurchases. The firm-level control variable is earnings, and the market-level control variable is the price index for the S&P 500. The macroeconomic factors included are IPI, M1, inflation (CPI), and short-term interest rate using the 3-month Treasury rate. Because of the cointegration results, we use a VECM model coupled with Granger causality tests to find long-term and short-term causal relationships.

These results support some generally agreed-upon findings; however, they offer new insight into the complex and interwoven relationship between payout activity, firm-level factors, and macro-economic factors. Granger causality examines short-term causality, and the findings indicate previously unexamined relationships between macroeconomic influences and payouts. These results highlight differences between dividends and repurchases that have not been mentioned in the literature. Changes in IPI and the 3-month Treasury rate Granger cause total repurchases. No Granger causality is found for dividends. These findings support the theory that dividends are "sticky," and once set, firms tend to be reluctant to make changes. Repurchases, on the other hand, are considered to be a more flexible payout method that can be responsive to changes in the payout environment as well as the macroeconomic environment. Two cointegration equations are found, indicating the presence of shared long-term trends. These findings suggest that all three payout variables share a long-term trend with IPI and shortterm interest rates. Total payout and total dividends also share a longterm trend with CPI. It is noted that in the presence of more than one cointegrating vector, the first is most useful (Johansen & Juselius, 1990). These findings suggest that corporate payouts and macroeconomic variables do share a long-term trend and that there are short-term and long-term macroeconomic influences woven throughout the payout equation.

This research is viewed as a first step, and these results shed light on new opportunities for further work. Such opportunities could include additional firm-level factors and control variables such as total assets, investment, cash, and equity. This study could be further enhanced by including a time variable, additional macroeconomic factors, as well as specific examination of the impact of the financial crisis within the context of testing for a structural break.

# 2. DATA AND SUMMARY STATISTICS

The selected macroeconomic variables include the IPI, the CPI, the price level of the S&P 500 index (S&P), the 3-month Treasury rate (short-term interest rate), and the monetary base (M1). The macroeconomic data were retrieved from the website of the Federal Reserve Bank of St. Louis. Data on the S&P 500 price level were obtained from the Home Page of Robert J. Shiller of Yale University (http://www.econ.yale.edu/~shiller). Data are seasonally adjusted. Independent variables were reported monthly and dependent variables were reported quarterly. Data were converted monthly to quarterly by taking an average.

The payout variables are from COMPUSTAT North America quarterly data from 1984 through 2011, available through Wharton Research Data Services (WRDS), and include 205,121 firm-level observations. I follow existing literature and use quarterly data (Dittmar and Dittmar 2004) and eliminate utilities because of the high level of regulation in that industry (Floyd, Li and Skinner, 2015). Dividends are the total amount of cash dividends paid (DVQ). Earnings are earnings before extraordinary items (IBQ) and represent the income of a firm after all expenses, including special items, income taxes, and minority interest, but before provisions for common and/or preferred dividends. This item does not reflect discontinued operations or extraordinary items. Stock repurchases (RP) are purchases of common and preferred stock (PRSTKCQ). This item represents any use of funds that decreases common and/or preferred stock. Descriptive statistics for all firm-level variables are reported in Table 1.

Table 1. Descriptive statisticss								
	<b>A Dividends</b>	Δ Earn	ΔRP	$\Delta$ Total payout				
Mean	0.00724	0.1345	0.0067	0.0124				
Median	0.00321	0.0124	0.0015	0.0203				
Maximum	11.69330	15.000	1.7400	6.2345				
Minimum	-1.449700	-881.00	-1.1500	-1.1500				
Std. Dev.	0.118262	8.6848	0.0418	0.0418				

Notes: Based on 205,121 firm-level observations 1984–2011. The  $\Delta$  denotes the percent change in the variable.

# **3. METHODOLOGY, PRELIMINARY TESTS, AND RESULTS**

VAR models and the VECM have been used to address a wide variety of financial problems (Campbell & Ammer, 1991; Lettau & Ludvigson, 2001; Sarig, 2004) and are applied here. This model allows us to investigate Granger causality among a number of variables that dominate the literature. The model makes no assumptions regarding directional causality and also captures the effects of unobservable variables with the inclusion of lagged values of the dependent variable.

Following Enders (2010), four steps are included to develop this model: (a) pretest all variables for order of integration; (b) estimate the model to determine the number of cointegration equations; (c) identify causal linkages using Granger causality; and (d) interpret the VECM to determine statistical significance.

#### 3.1 Pretest all variables for order of integration

The first step is to test all series for unit roots. Although it is common practice in time series modeling to apply the augmented Dickey-Fuller and Phillips-Perron tests to determine whether a series has a unit root, improved tests with better statistical properties are now available. We apply the GLS Dickey-Fuller test to all variables. The NG-Perron sequential t-test is used to determine the optimum lag length. The results reflect the existence of unit roots in all but one variable. All series are stationary when first differenced with the exception of the corporate yield curve. Because the natural logarithm is applied to the total payout, CPI, IPI, S&P, and M1 before differencing, these variables can be interpreted as percent change. The yield variables, when differenced, are a change in the percent.

#### 3.2 Test for cointegration

The next step is to test for cointegration to determine whether a VECM or VAR model is appropriate. These results are reported in Table 2. Because these models are very sensitive to lag length, Schwarz Information Criteria are used to determine lag length. If there is no cointegration present, a VAR model is appropriate. If cointegration is present, a VECM is the appropriate choice. The Johansen (1991) method to test for cointegration is used to detect whether there is a long-run relation between the variables. After taking the first difference, any linear trend is removed, and, for this reason, the model includes an intercept and not a trend. This test result suggests two cointegration equations when estimating the VECM. Comparing the trace statistic to the critical value for Rank = 1, the trace statistic exceeds the 5% critical value, providing support for Rank > 1. This result indicates that there are two cointegrating relations in the data, and, therefore, a VECM is the appropriate choice for model selection. This model accommodates the fact that the variables have unit roots and also are cointegrated.

For total payout, total dividends, and total repurchases, two cointegration equations exist. The first is for the payout IPI, CPI, S&P level, M1, and the 3-month Treasury rate. The model proposed here suggests that these variables share a long-term trend and long-term equilibrium relationship. This finding provides insight into the longterm relations that will be further clarified with the application of a VECM. The following section addresses short-term relations.

Table 2	Table 2: Results of Johansen cointegration test									
				5%		5%				
Max.		Eigen-	trace	crit.	max	crit.				
rank	LL	value	statistic	value	stats	value				
0	1025.0383		143.2524	82.49	67.4624	36.36				
1	1058.7695	0.47402	75.79	59.46	39.176	30.04				
2	1078.3575	0.31141	36.6140*	39.89	20.1764	23.8				
3	1088.4457	0.17482	16.4376	24.31	12.221	17.89				
4	1094.5562	0.10987	4.2166	12.53	3.3891	11.44				
5	1096.2507	0.03176	0.8274	3.84	0.8274	3.84				
6	1096.6645	0.00785								

The hypothesis being tested is always a certain cointegrating rank with the alternative hypothesis being that cointegrating rank is greater than the hypothesis being tested. In comparing the trace statistic to the 5% critical value for Rank 1, the 5% critical value > the trace statistic. Thus, the hypothesis that Rank = 1 is rejected at the 5% level of significance in favor of the hypothesis that Rank > 1.

# 3.3 Granger causality

Granger causality tests determine whether past values of variable X can help explain current values of variable Y. Because the results of the cointegration test revealed two cointegrating equations, the Granger causality tests involve the coefficients on nonstationary variables, which lead to the standard interpretations of the F-statistic as being inappropriate (Enders, 2010). The Wald test is applied, and Table 3 presents the results of the (VEC) Granger causality/Wald Test. These results indicate that past values of industrial production and short-term interest rates influence current values of total payouts and total repurchases, but do not exert a short-term influence on total dividends. The results indicate significant relationships among macroeconomic factors, repurchases, and total payouts and also highlight differences between dividends and repurchases.

# 4. RESULTS: VECTOR ERROR CORRECTION MODEL (VECM)

The VECM is estimated for total payout, total dividend, and total repurchase. Each VECM includes four macroeconomic variables:

Wald test							
Null Hypothesis	Chi_sq	p-value					
D.lnipi does not granger cause D.lntotalpay	6.02	0.02*					
D.lncpi does not granger cause D.lntotalpay	2.35	0.22					
D.lnsp does not granger cause D.lntotalpay	19.05	0.01*					
D.threemnt does not granger cause							
D.Intotalpay	4.72	0.08					
D.lnm1 not granger cause D.lntotalpay	13.12	0.00*					
D.Intotalpay does not granger cause D.Inipi	2.62	0.25					
D.lncpi does not granger cause D.lnipi	3.34	0.18					
D.lnsp does not granger cause D.lnipi	43.43	0.00*					
D. does not granger cause D.lnipi	5.93	0.05					
D.lnmbdoes not granger cause D.lnipi	12.17	0.02					
D.Intotalpay does not granger cause D.Incpi	6.74	0.03*					
D.lnipi does not granger cause D.lncpi	5.56	0.06					
D.lnsp does not granger cause D.lncpi	4.55	0.10					
D.threemnt does not granger cause D.lncpi	1.18	0.55					
D.lnm1 not granger cause D.lncpi	2.93	0.23					
D.Intotalpay does not granger cause D.Insp	0.67	0.62					
D.lnipi does not granger cause D.lnsp	3.32	0.19					
D.lncpi does not granger cause D.lnsp	1.61	0.45					
D.threemnt does not granger cause D.lnsp	8.21	0.02*					
D.lnm1 not granger cause D.lnsp	2.45	0.29					
D.Intotalpay does not granger cause D.Inm1	2.80	0.21					
D.lnipi does not granger cause D.lnm1	17.70	0.00*					
D.lncpi does not granger cause D.lnm1	1.22	0.54					
D.threemnt does not granger cause D.lnm1	4.10	0.13					
D.lnsp does not granger cause D.lnm1	3.48	0.18					
D.Intotalpay does not granger cause							
D.threemnt	11.65	0.00*					
D.lnipi does not granger cause D.threemnt	22.73	0.00*					
D.lncpi does not granger cause D.threemnt	5.62	0.06					
D.lnsp does not granger cause D.threemnt	2.27	0.32					
D.lnmb does not granger cause D.threemnt	11.63	0.00*					

# Table 3: Results from Granger causality/Block exogeneity

Notes: For each of the above variables, D.In indicates the first difference of the natural log of total payout, CPI, IPI, S&P level, and M1. The natural log transformation is not applied to the 3-month Treasury rate (threemnt). The first difference of the logged variables results in growth rates. \* indicates significance at 5% level.

percentage change in industrial production index, percentage change in consumer price index, percentage change in money supply (M1), and change in 3-month Treasury rate, while controlling for earnings and the price level of the S&P.

The results from the cointegration and VECM analysis results support current theoretical explanations for payouts but also indicate interesting differences among the three payout variables. Cointegration equations are reported in Table 4. Two cointegration equations reflect the long-term relations, whereas the error correction terms and speed of adjustment coefficients represent how each variable responds to deviations from the long-run equilibrium. The first cointegration equation in each model is the primary focus. The first cointegration equation for total payout indicates that changes in total payout share a long-term trend with IPI, CPI, and short-term interest rates. The first cointegration equation for total dividends also indicates that changes in total dividends share a long-term trend with IPI, CPI, and short-term interest rates. The first cointegration equation for total repurchases indicates that changes in total repurchases share a long-term trend only with IPI and short-term interest rates. CPI is dropped from the equation and is not significant.

Tab	le 4: Cointe	gration equa	ations from	i VECM re	esults				
Ce1	L.Intotalpay	-9.52L.lnIPI	-5.05L.lncpi	+5.67L.lnsp	-5.12Lthmnt	+39.16			
		(0.00)	(0.02)	(0.00)	(0.00)				
Ce2	L.Intoearn	-8.81L.lnipi	-4.91L.lncpi	+4.62L.lnsp	+29.39				
		(0.00)	(0.06)	(0.00)					
Cel	L.Intotaldiv	-6.45L.lnIPI	-5.82L.lncpi	+1.64L.lnsp	-1.48Lthmnt	+25.24			
		(0.00)	0	(0.00)	(0.00)				
Ce2	L.Lntoearn	-7.24L.lnipi	-6.44L.lncpi	+2.30L.lnsp	-0.15Lthmnt	30.61			
		(0.00)	(0.06)	(0.00)	(0.01)				
Cel	L.Intotalrp	-19.45L.lnIPI	+4.38L.lnsp	38Lthmnt	+66.26				
		(0.00)	(0.00)	(0.00)					
Ce2	L.Lntoearn	-8.31L.lnipi	+2.55L.lnsp	+28.03					
		(0, 00)	(0, 00)						

Note: Table 4 includes the estimated coefficients for the two cointegration equations for a VECM model that includes total payout, total dividends, total repurchases and IPI, CPI, S&P and the three month Treasury rate. L.ln indicates the lag of the log of the specified variable. P values are in parentheses. Following the literature, the focus is on the first cointegration equation for each payout variable.

VECM results are reported in Tables 5, 6, and 7. Results indicate a significant relationship between payouts and IPI and short-term interest rates. Results for total payout indicate that IPI, CPI, the S&P, and the 3-month Treasury rate share a long-term trend with total payout. IPI, CPI, and the 3-month rate all have significant, negative relations with total payout. The VECM results for total dividends are similar to total payout. Total dividends share a long-term trend with IPI, CPI, the S&P, and the 3-month Treasury. As with total payout, all relations are negative. VECM results for total repurchases are somewhat different. Total repurchases share a long-term trend with IPI, S&P, and the 3-month rate. IPI and the 3-month rate are significant, negative relations, and CPI is not significant relative to repurchases. These results are not surprising. IPI is a coincidental economic indicator. Both theoretical and empirical evidence suggest that corporate cash flows are related to a measure of aggregate output such as GDP or IPI. The findings support a theoretical framework in which payouts decrease when aggregate economic activity is increasing, thus driving an increase in investment activity.

Table 5: VECM I	Cable 5: VECM Results for total payout										
Explanatory Variables	∆i ntotald îv	∆ntotalearn	∆in1p1	Δincpi	Δinsp	∆InM1	∆thrmn				
	-0.03	0.63	0.02	<0.01	-0.06	0.01	0.11				
	(0.27)	(0.00)	(0.00)	(0.99)	(0.03)	(0.00)	(0.45)				
	<0.01	-0.51	<-0.01	<0.01	0.02	<0.31	0.09				
	(0.63)	(0.00)	(0.9Z)	(0.54)	(0.23)	(0.53)	(0.33)				
∆ L.Intotaldiv	-0.10	-0.82	<0.01	<-0.01	-0.05	-0.0Z	-0.29				
	(0.33)	(0.22)	(0.90)	(0.87)	(0.66)	(0.41)	(0.65)				
ΔL. Intotalearn	<0.01	-0.25	<0.01	<0.01	<0.01	<-0.01	0.23				
	(0.79)	(0.00)	(3.46)	(0.70)	(0.58)	(0.68)	(0.01)				
∆Llnipi	0.79	5.38	0.32	0.03	0.13	-0.28	3.12				
	(0.12)	(0.11)	(0.00)	(0.58)	(0.8Z)	(0.00)	(0.32)				
ΔLIncpi	0.51	-8.49	-0.07	-0.35	-2.47	0.18	-12.3				
	(0.54)	(0.13)	(0.61)	(0.61)	(0.00)	(0.31)	(0.02)				
∆Linsp	0.08	2.49	2.07	0.01	0.39	-0.01	-0.70				
	(0.43)	(0.00)	(0.00)	(0.35)	(0.00)	(0.41)	(0.27)				
∆LInM1	0.24	-3.75	-0.23	-0.37	0.72	0.36	1.49				
	(0.60)	(0.22)	(0.00)	(0.20)	(0.17)	(0.00)	(0.61)				
∆Lthrmn	<0.01	0.16	<0.01	<0.01	<0.01	<-0.01	0.59				
	(0.78)	(0.06)	(0.14)	(0.74)	(0.99)	(0.01)	(0.00)				
constant	0.02	<0.01	<0.01	<0.01	0.03	<0.01	<0.01				
	(0.11)	(0.97)	(0.18)	(0.00)	(0.00)	(0.02)	(0.93)				

Notes: Included is the speed of adjustment (SOA) coefficient of the first and second cointegration equations,  $\Delta$  denotes the change of the natural log of the total dividend, total earning, IPI, CPI, the S&P level, M1, and the 3-month Treasury rate. L.ln indicates the lag of the log of the specified variable. P values are in parentheses. Table includes the estimated error correction terms for a VECM model that includes total dividends, IPI, CPI, S&P and the 3-month Treasury rate.

In many aspects, these findings are consistent with current theoretical explanations; however, these results shed light on new significant influences on the corporate payout equation. The results indicate strong similarities between the significant relations found for the two forms of payout: dividends and repurchases. The directional relations are not surprising. What is surprising is the level of influence and sens itivity, as indicated by the coefficients. These findings suggest that repurchases are more sensitive to changes in IPI, and dividends are more sensitive to changes in short-term interest rates.

Fable 6: VECM results for total dividends										
Explanatory Variables	∆l ntotald iv	∆ntotalearn	∆lnipi	∆Incpi	∆lnsp	ΔInM1	∆thrmn			
	-0.03	0.63	2.02	<0.01	-0.06	0.01	0.11			
	(0.27)	(0.00)	(0.00)	(0.99)	(0.03)	(0.00)	(0.45)			
	<0.01	-0.51	<-0.01	<0.01	0.02	<0.31	0.09			
	(0.63)	(0.00)	(0.92)	(0.54)	(0.23)	(0.53)	(0.33)			
Δ L.Intotaldiv	-0.10	-0.8Z	<0.01	<-0.01	-0.05	-0.0Z	-0.2 <del>9</del>			
	(0.33)	(0.22)	(0.90)	(0.87)	(0.66)	(0.41)	(0.65)			
ΔL. Intotalearn	<0.01	-0.25	<0.01	<0.01	<0.01	<-0.01	0.23			
	(0.79)	(0.00)	(2.46)	(0.70)	(0.58)	(0.68)	(0.01)			
∆Llnipi	0.79	5.38	0.32	0.03	0.13	-0.28	3.12			
	(0.12)	(0.11)	(0.00)	(0.58)	(0.82)	(0.00)	(0.32)			
∆ L.Incpi	0.51	-8.4 <del>9</del>	-0.07	-0.35	-2.47	0.18	-12.3			
	(0.54)	(0.13)	(1.61)	(0.61)	(0.00)	(0.31)	(0.02)			
ΔLinsp	0.08	2.4 <del>9</del>	3.07	0.01	0.39	-0.01	-0.70			
	(0.43)	(0.00)	(0.00)	(0.35)	(0.00)	(0.41)	(0.27)			
ΔLInM1	0.24	-3.75	-0.23	-0.07	0.72	0.36	1.49			
	(0.60)	(0.2Z)	(0.00)	(0.20)	(0.17)	(0.00)	(0.61)			
ΔLthrmn	<0.01	0.16	<0.01	<0.01	<0.01	<-0.01	0.59			
	(0.78)	(0.06)	(0.14)	(0.74)	(0.99)	(0.01)	(0.00)			
constant	0.02	<0.01	<0.01	<0.01	0.03	<0.01	<0.01			
	(0.11)	(0.97)	(0.18)	(0.00)	(0.00)	(0.0Z)	(O.93)			

Notes: Included is the speed of adjustment (SOA) coefficient of the first and second cointegration equations,  $\Delta$  denotes the change of the natural log of the total dividend, total earning, IPI, CPI, the S&P level, M1, and the 3-month Treasury rate. L.ln indicates the lag of the log of the specified variable. P values are in parentheses.

Table includes the estimated error correction terms for a VECM model that includes total dividends, IPI, CPI, S&P and the 3-month Treasury rate.

This finding supports the literature, which suggests that repurchases and dividends are very different avenues for payout and not simply substitutions for one another. The repurchase form is considered to be more flexible and designed to disgorge nonoperating or transitory earnings, whereas dividends are considered to be more consistent and less likely to be changed. The different coefficients for IPI clearly stand out in the results. A 1% change in IPI leads to a 9.52% decline in total payouts, a 6.45% decline in total dividends, and a 19.45% decline in total repurchases. This result reveals that repurchases are much more sensitive to changes in 3-month interest rates. A 1% change in the 3-month Treasury rate leads to a 1.48% decline in total dividends and only a .38% decline in total repurchases. Results suggest that IPI and

Table 7. VECM	I Results	for Tota	al Repu	rchase			
Explanatory Variab es	∆intotairp	∆intotaleam	Ainipi	Δincpi	∆Insp	ΔInM1	Δthrmn
	0.01	0.18	<0.01	<-0.01	-C.03	<0.01	0.01
	(0.54)	(0.00)	(0.00)	(0.38)	(0 02)	(0.00)	(0.84)
	-0.05	-0.50	<-0.01	<0.01	0 02	<-0.01	0.05
	(0.09)	(0.00)	(0.59)	(0.23)	(0 25)	(0.99)	(0.56)
∆ Lintotairp	-0.15	-0.08	<0.01	<-0.01	<-0.01	<-0.01	-0.03
	(0.1Z)	(0.76)	(0.29)	(0.83)	(0 90)	(0.94)	(0.9Z)
ΔL ntotalearn	<c.01< td=""><td>-0.25</td><td>&lt;0.01</td><td>&lt;0.01</td><td>0 01</td><td>&lt;0.01</td><td>0.25</td></c.01<>	-0.25	<0.01	<0.01	0 01	<0.01	0.25
	(0.90)	(0.00)	(0.35)	(0.70)	(0 47)	(0.97)	(0.00)
∆Llnipi	3.45	5.59	0.31	0.04	0 21	-0.25	4.61
	(0.00)	(0.10)	(0.00)	(0.51)	(0 72)	(0.02)	(0.16)
∆ Lincpi	-1.90	-8.79	-0.14	-0.07	-Z.40	0.15	-13.31
	(0.34)	(0.12)	(0.35)	(0.49)	(0 01)	(0.42)	(0.01)
∆ Linsp	0.83	2.58	0.60	<0.01	0 4Z	-0.0Z	-0.53
	(0.00)	(0.00)	(0.00)	(0.36)	(0 00)	(0.33)	(0.42)
∆LInM1	-0.61	-3.05	-0.23	-0.03	0 82	0.38	3.39
	(0.57)	(0.32)	(0.00)	(0.56)	(0 11)	(0.00)	(0.24)
∆ Lthrmn	0.08	0.17	<0.01	<0.01	<-0.01	<-0.01	0.61
	(0.00)	(0.05)	(0.02)	(0.95)	(0 69)	(0.11)	(0.00)
constant	0.02	<0.01	<0.01	<0.01	0 02	<0.01	0.01
	(0.49)	(0.99)	(0.00)	(0.00)	(0 07)	(0.00)	(0.84)

short-term interest rates influence payout decisions, both amount and form, in ways not previously understood.

Notes: Included is the speed of adjustment (SOA) coefficient of the first and second cointegration equations,  $\Delta$  denotes the change of the natural log of the total dividend, total earning, IPI, CPI, the S&P level, M1, and the 3-month Treasury rate. L.ln indicates the lag of the log of the specified variable. P values are in parentheses. Table includes the estimated error correction terms for a VECM model that includes repurchases, IPI, CPI, S&P and the 3-month Treasury rate.

# 5. CONCLUSION

Does payout activity share a long-run equilibrium relation with the overall economy, and are there significant short-term and long-term influences in the payout equation that extend beyond firm-level determinants? These findings document significant relationships between macroeconomic factors and aggregate payout variables, suggesting that total payouts, total dividends, and total repurchases interact within a complex and dynamic system. This finding suggests that macroeconomic factors are not exogenous to the corporate payout equation and should be included/controlled for in future work. This evidence provides support for the life-cycle theory of corporate payouts that suggests that the macroeconomic factors influence the payout cycle through earnings and firm investment. There is also evidence in these results for the existence of significant differences between dividends and repurchases that have not been identified in past work. Overall, this research suggests that macroeconomic factors significantly influence corporate payouts and contribute to understanding previously unexplained variations in aggregate payout activity.

The findings suggest that dividends and repurchases need to be examined separately, as well as together in the form of total payouts. Payout activity results from responses to the broad economic environment, as well as firm-level characteristics. IPI significantly influences both repurchase and dividend activity. Specifically, IPI, as a measure of aggregate output, operates within a feedback loop in which economic activity drives investment and investment drives economic activity. In periods of increased investment, aggregate payout declines. These findings show a greater decline in repurchases over dividends. Repurchases generally imply an increased level of financial flexibility in payout policv. With economic expansion also comes varying degrees of uncertainty (Dittmar and Dittmar, 2004), and therefore an increased need for flexibility may be indicated by the use of repurchases. These findings are consistent with this explanation. Short-term interest rates also contribute to explaining aggregate dividend and repurchase activity. The pecking order theory of corporate payouts (Myers and Majluf, 1984) suggests that as costs of external finance increase, firms prefer using internal funds and this may lead to a decrease in payouts resulting in a negative relation. These results provide evidence for a negative relation between short-term rates and payout activity, as well as a stronger relation with dividends, as opposed to repurchases. Given that economic growth relates to each of these factors, the results indicate that payout activity and the form of payout fluctuate with changes in the macroeconomic, environment. This result provides support for including macroeconomic variables in the corporate payout equation.

# References

Chen, N.F, Roll, R., & Ross, S. A. (1986). Economic forces and the stock market. *J Bus*, 59(3), 383-403.

Cochrane, J. (2011). Presidential address: Discount rates. J Finance, 66(4), 1047-1108.

DeAngelo, H., & DeAngelo, L. (2006). The irrelevance of the MM dividend irrelevance theorem. *J Financ Econ*, 79(2), 293-315.

Dittmar, A.K., & Dittmar, R.F. (2004). Stock repurchase waves: An explanation of the trends in aggregate corporate payout policy. Working Paper, University of Michigan.

Easterbrook, F.H. (1984). Two agency-cost explanations of dividends. *Am Econ Rev,* 74, 650-659.

Enders, C.K. (2010). *Applied Econometric Time Series* (3<sup>rd</sup> ed.) New York, NY: Guilford.

Fama, E.F. (1990). Stock returns, expected returns, and real activity. *J Finance*, 45, 1089-1108.

Floyd, E., Li, N., & Skinner, D. J. (2015). Payout policy through the financial crisis: The growth of repurchases and the resilience of dividends. *J Financ Econ*, 118(2), 299-316.

Gordon, M. (1959). Dividends, earnings and stock prices. *Revi Econ Stat*, 52 (4), 126-134.

Grullon, G., Michaely, R., & Swaminathan, B. (2002). Are dividend changes a sign of firm maturity? *J Bus* 75(3), 387-424.

Johansen, S., & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration with applications to the demand for money. *Oxf Bull Econ Stat, 52*, 169-210.

Lettau, M., & Ludvigson, S. (2001). Consumption, aggregate wealth, and expected stock returns. *J Finance*, *56*(3), 815-850.

Lintner, J. (1956). Distribution of incomes of corporations among dividends, retained earnings, and taxes. *Am Econ Rev, 46*(2), 97-113.

Myers, S.C., & Majluf, N.S. (1984). Corporate financing decisions when firms have investment information that investors do not. *J Financ Econ*, *13*, 187-220.

Rau, P.R., & Stouraitis, A. (2009). Patterns in the timing of corporate event waves. Available at SSRN: http://ssrn.com/abstract=946797.

Sarig, O. (2004). A time-series analysis of corporate payout policies. *Rev Financ*, *9*, 1-22.

Serfling, M.A., & Miljkovic, D. (2011). Time series analysis of the relationships among (macro) economic variables, the dividend yield and the price level of the S&P 500 index. *Appl Financ Econ, 21*, 1117-1134.

# A Framework for Teaching the Goal of the Firm in Introductory Business Classes: Shareholder Wealth Maximization Ethicality and Classical Philosophical Paradigms

Rebekah Inez Brau Brigham Young University

# Abstract

In many business textbooks and courses, particularly in finance curricula, students a taught that the single goal of the firm is shareholder wealth maximization (SWM). Although this statement presents a clear and succinct goal, which adheres well to rational expectations and financial economic models, I argue it is oversimplistic and may be detrimental to student learning and professional conduct. Anecdotal evidence suggests strict adherence to the SWM goal might lead business students to act unethically after graduating. I propose a framework that includes discussion of 4 classical ethics camps vis-à-vis SWM in foundational business classes. I assume students have not had a philosophical or business ethics course and begin the discussion there. The crucial point is that managers of public companies do not own the businesses they run. They are employed by the firms' owners to maximise the long-term value of the owners' assets. Putting those assets to any other use is cheating the owners, and that is unethical. If a manager believes that the business he is working for is causing harm to society at large, the right thing to do is not to work for that business in the first place. Nothing obliges someone who believes that the tobacco industry is evil to work in that industry. But if someone accepts a salary to manage a tobacco business in the interests of its owners, he has an obligation to those owners. To flout that obligation is unethical. In addition, of course, managers ought to behave ethically as they pursue the proper business goal of maximising owner value—and that puts real constraints on their actions.

- 1/20/2005, The Economist (*http://www.economist.com/ node/3555286*)

#### Introduction

In a fairly scathing article titled, "Business Ethics and Social Responsibility in Finance Instruction: An Abdication of Responsibility," Hawley (1991) argues that the shareholder wealth maximization (SWM) goal of the firm as described in the opening *Economist* quote is overemphasized in finance textbooks to the detriment of students. After stating that SWM may be "a very effective tool for decision making," the author goes on to claim, "[I]t can also be used to rationalize the commission of unethical or socially irresponsible actions." Further stating, "Overemphasis on the SWM objective by some companies can lead to dangerous or disastrous consequences for consumers, employees, or the general population," Hawley (1991) examines 22 undergraduate and MBA finance textbooks and finds that SWM is virtually universal as the only goal of the firm presented or seriously considered.

The Hawley (1991) article is a call to action. Despite the careful analysis and discussion the author provides, a survey of contemporaneous finance textbooks pertaining to the goal of the firm look remarkably like they did nearly three decades ago between 1988 and 1990 when the Hawley article was written. Many of the leading business and corporate finance books at both the undergraduate and graduate levels through the years state the goal of the firm simply as something akin to, "maximize shareholder wealth" (e.g., Ross et al. (2013), Keown et al. (2005)). Consider just two examples from introductory finance texts. Block and Hirt (2005, p. 12) state, "The broad goal of the firm can be brought into focus if we say the financial manager should attempt to *maximize the wealth of the firm's shareholders* through achieving the highest possible value of the firm. Shareholder wealth maximization is not a simple task, since the financial manager cannot directly control the firm's stock price, but can only act in a way that is consistent with the desires of the shareholders." Note how the authors, in the original text, use italics and bold to emphasize that SWM *is* the goal of the firm.

Levy (1998, p.24) states, "As indicated earlier, management operates the corporation for the stockholders. The only legitimate criterion in managerial decision making is stockholders' welfare, as measured by their wealth. The corporation's goal should be to maximize the stockholders' equity (the stock price times the number of outstanding shares), or simply maximizing the stock price." Note that Levy uses very strong terms, "the only legitimate criterion..." By this text, it is as if any discussion of any conditions under which SWM may not be ethical (i.e., not the right goal) is simply not worth considering.

My primary recommendation is that instructors should spend sufficient time discussing not only the SWM goal as a stated fact, but to do so within an ethical context of what conditions are necessary for SWM to be an ethical goal. The framework I suggest is using as a structure the four philosophical ethics camps of Deontology, Justice/Fairness Theory, Utilitarianism/Consequential Theory, and Virtue Ethics Theory. Inherently if students think about various conditions that make SWM ethical, they will include thinking about other stakeholders, not just shareholders. SWM can be argued as an ethical goal, but it must meet certain necessary (and sufficient) conditions to be such. These conditions vary depending on the philosophical camp that is used as the underpinning for the definition of ethical. As students contemplate the various ethics camps and how they apply to SWM, they can learn that SWM is not as obvious as is proposed (and assumed) in most finance (and other business) textbooks.

The intent of my paper is to add to the business (and particularly finance) ethics literature as it applies to the notion of teaching students that maximizing shareholder wealth is the ultimate goal of the firm. Given the work of Bernardi et al. (2008), finance ethics research has lagged other management fields, specifically accounting and marketing, and perhaps others such as organizational behavior. Whereas accounting and marketing have seen significant increases in the number of articles discussing ethics in their disciplines, the field of finance has not (Bernardi et al. 2008).

The order of the paper is as follows: The next section consists of the literature review, beginning with the underpinnings of the SWM

goal.<sup>1</sup> I next discuss SWM under each of the ethics schools of thought in the context of student instruction. The final section includes discussion for implementing the ethics framework and then summarizes and concludes.

# Literature Review

# The "Stakeholder versus Shareholder" Dilemma

The version taught to introductory business students of the SWM goal seems to have been significantly advanced through the efforts of Berle (1931) and was probably solidified as Stewart (2011) argues, "With the success of Manne's (1964) perspective, the shareholder wealth maximization norm was born, firmly defining the interest of shareholders and planting the seeds for the financialization of the firm." Through a series of articles (Manne 1959, 1961, 1962a, 1962b, 1964, 1965), Manne argued for the efficacy of Berle's early logic for the shareholder primacy theory of the firm. For example, in his 1931 *Harvard Law Review* article, Berle suggested that corporations should maximize benefits for shareholders. Under Berle's perspective, managers only consider the interest of all stakeholders inasmuch as these interests maximize the wealth of the shareholders.

In 1919, Ford Motor Company was brought to court by minority shareholders to decide whether the company could continue to operate in the charitable manner Henry Ford had implemented. Ford, the majority shareholder, proclaimed in 1916:

My ambition is to employ still more men, to spread the benefits of this industrial system to the greatest possible number, to help them build up their lives and their homes. To do this we are putting the greatest share of our profits back in the business (Ford and Miller, 1922).

The court concluded:

A business corporation is organized and carried on primarily for the profit of the stockholders. The powers of the directors are to be employed for that end. The discretion of directors is

<sup>&</sup>lt;sup>1</sup> The topic of ethics has been addressed to some degree outside of the finance journals measured in the figure above as in Horrigan (1987) who published "The Ethics of the New Finance" in the *Journal of Business Ethics*. Horrigan discusses five major theories of finance but he does not focus on SWM as one of them.

to be exercised in the choice of means to attain that end, and does not extend to a change in the end itself, to the reduction of profits, or to the non-distribution of profits among stock-holders in order to devote them to other purposes (Ford and Miller, 1922).

The Dodge v. Ford Motor Company is an important legal case in the context of SWM because it is the initial case to rule in favor of shareholder primacy. However, several years after the court ruled on the Dodge v. Ford Motor Company case, researchers began developing stakeholder theory. Under the paradigm of the stakeholder theory, the goal of maximizing shareholder wealth is challenged. Stakeholder theory argues that other stakeholders have a vested interest in what firms do (and what they should be doing) as well and should be considered directly in the goal of the firm. For example, a year after Berle (1931), Dodd (1932), also in the Harvard Law Review, challenged the goal of shareholder maximization and stated that, in addition to shareholder interest, corporations should consider the interest of the organization's workers, consumers, suppliers, and society as a whole in the decisionmaking process. Dodd held the belief that organizations have a responsibility to all stakeholders of an organization, not just shareholders. This perspective was supported by a number of prominent business leaders. For example, Henry Ford also stated, "For a long time people believed that the only purpose of industry was to make a profit. They are wrong. Its purpose is to serve the general welfare." (Donaldson, 1982).

The case Shlensky v. Wrigley, 1968, acknowledges the changing perspectives from SWM to stakeholder theory. However, the court concluded that the president had a right to make decisions that were not in accordance with what the stakeholders desired as long as the decision was not, "...fraud[ulant], illeg[al], or conflict of interest." (Shlensky v. Wrigley, 1968).

In 1970, one of the most influential economists of the 20<sup>th</sup> century, Nobel laureate Milton Friedman, reiterated Berle's perspective and the Dodge v. Ford decision that organizations have the sole responsibility to provide financial gain for the organization's shareholders and suggested that, "The social responsibility of business is to increase its (the organization's) profits." (For the purpose of this article, ethics and social responsibility are considered under the same umbrella, although it is recognized these can be two different concepts.) Under Friedman's assumption, the executives of an organization are the employees of the shareholders. As such, employees have the fiduciary responsibility to make as much money as possible for the owners of the firm. (An underlying assumption of Friedman's argument is that while businesses should maximize their profits, they must still obey the legal rules and ethical customs of society.) Friedman states:

> There is one and only one social responsibility of business—to use its resources and engage in activities designed to increase its profits so long as it stays within the rules of the game, which is to say, engages in open and free competition without deception or fraud (Friedman, 1970).

It seems the very large majority of finance textbooks agree wholeheartedly with Friedman's assertion above. For example, Keown et al. (2005) argue that:

Not only will this goal [shareholder wealth maximization] be in the best interest of the shareholders, but it will also provide the most benefits to society. This will come about as scarce resources are directed to the most productive use by businesses competing to create wealth (p. 4).

My concern is that teaching introductory business students should be more careful and more nuanced than simply adhering to the logic of Friedman (1970) or Jensen (2001, 2002). For example, other articles such as Mitchell et al. (2015) argue for pluralism and multiobjective corporations. If students are not properly taught to think about a more complex goal of the firm, then the ethical breach concerns of Hawley (1991) may be a result of strict acceptance of and adherence to SWM. Such ethical breaches can be seen in the Enron and Worldcom frauds, for example.

# The Pricing of Ethics and the Firm as a Conduit of Ethics

Students should recognize that the discussion itself of whether or when SWM is an ethical goal presupposes that a firm goal *can* reflect an ethical stance in the first place. Prior to discussing the various ethics traditions, it is instructive to consider this notion of whether a firm can be ethical or not. In an article that does just this, Chambers and Lacey (1996) first state what they perceive is a problem with the interpretation of SWM:

An unfortunate theme of the stakeholder approach is that the traditional view of shareholder wealth maximization (SWM) is inconsistent with ethical behavior. It is often suggested or

concluded that people's desire for money in general and shareholders' wealth maximization in particular are in direct opposition to ethical concepts such as honesty, fairness, and the good of society (Chambers and Lacey, 1996).

In addressing this line of reason, Chambers and Lacey's argument is that firms themselves do not serve as a determinant of ethics, but rather as a conduit of ethics. Relying on financial economics, they argue that if investors care about ethics, then they *price* ethics into the stock price, just as they would with any attribute of a stock they desire. The logic goes that if society (investors) feel that a firm is acting unethically, they will not invest in that firm, and as such, the stock price will fall (i.e., ethics are priced). In order for management to maximize shareholder value then, they need to act ethically (as defined by potential investors) so the perceived firm's "lack of ethics" as judged by the investors does not discount the stock price. Thus, the goal of SWM serves as a mirror that reflects the ethics of investors instead of a source of ethics in and of itself.

A counterpoint to their own argument is that perhaps shareholders choose not to act ethically in their investment decisions. If this is the case, then the conduit of the firm to price ethics may be turned on its head. Considering such a concern, Chambers and Lacey (1996) continue:

> Another important implication or aspect of economic laws is their ability to provide incentives for shareholders to behave according to the ethical standards of the product market participants. In other words, even if the shareholders do not care about behaving ethically, the behavior of consumers and other product market participants can provide incentives for the shareholders to act as if they cared about the issue....

> Thus, many product markets reflect value that depends on reputation. The objective of shareholder wealth maximization can force firms operating in these markets to behave as if they cared about ethical standards.

Using the logic of Chambers and Lacey (1996), markets may be able to provide two critical functions—one on investors and one on firms. The first mechanism is that product market participants may impose ethics on shareholders so shareholders act in accordance with those ethics. The second mechanism is that the concern for the pricing of firm reputation capital gives firms incentives to act ethically so their share prices do not suffer.

Providing some degree of evidence for Chambers and Lacey (1996), Rao and Hamilton (1996) offer an example of how ethics (in this case, unethical behavior) is priced by the market (with stock price decreases). Rao and Hamilton use event study methods to test five categories of unethical behavior:

Bribery, scandals, white collar crime, illegal payment,
Employee discrimination,
Air pollution, water pollution, environmental cleanup, pollution,
Insider trading, and
Business ethics.

Combining all 58 events, Roa and Hamilton find an 11-day cumulative abnormal return around *Wall Street Journal* announcements of -14.84% and an event single-day return of -5.67%. Roa and Hamilton show persistent negative abnormal monthly returns for five months after the announcement. Roa and Hamilton do not perform regressions on the abnormal returns to determine whether any of the five unethical behavior categories have specific impact on stock price. The conclusion of the article is that shareholders do price ethics—at least what is perceived as breaches of ethical behavior.

In a more general analysis of ethical firm behavior and SWM, Roman et al. (1999) provide an analysis of the literature between the correlation (if any) of corporate social performance (CSP) and corporate financial performance (CFP). In this case, CSP can be seen as a proxy for perceived ethics and CFP can be viewed as a measure of SWM. Roman et al. (1999) reconstruct the work of Griffin and Mahon (1997) who determined that of 62 research results in 51 articles, 33 show positive correlations between CSP and CFP, 20 show negative correlation, and 9 show insignificant correlation. Roman et al. (1999) reclassify these 62 results and update the data set of articles to conclude that most studies show either a positive relationship or no relationship between CSP and CFP. Although loose in interpretation for the pricing of ethics, this meta-analysis indicates that many studies show a positive pricing of ethics.

Along with the pricing of ethics in stock prices, Chambers and Lacey (1996) also argue in support of SWM using a markets argument:

The reason that shareholders are not unanimous with regard to ethical issues is the existence of corner point solutions. For example, politically conservative shareholders will not view a contribution to a liberal organization as being attractive even if their personal contribution level to the organization is reduced to zero. It is generally not possible for the shareholders to hedge the corporate donation by short selling goodwill generated by the donation. In other words, conservative shareholders generally can not sell this "commodity" (goodwill) to liberals who would enjoy it.

In the absence of well-functioning markets in which shareholders can buy and sell claims related to ethical issues, a strong efficiency argument can be made that everyone is at least as well off or better off if the financial manager of the corporation maximizes shareholder wealth and therefore allows shareholders to best maximize their private and social utility functions. In other words, the money should be passed through to the shareholders and then they can donate to the causes of their choice.

The above quotation is an example of the defense of SWM as an ethical matter. The statement that "everyone is at least as well off or better" rings of an ethical argument for social justice and the statement "to best maximize their private and social utility functions" also rings of obtaining a degree of optimality. In the subsequent section, I discuss the goal of SWM within the framework of student understanding of four major ethics traditions in an effort to think carefully about not if but when SWM may be an ethical firm goal.

# **Philosophical Ethics Traditions and SWM**

To begin the presentation of the SWM objective in introductory business classes, introducing students to the four ethics camps of Deontology, Justice/Fairness Theory, Utilitarianism/Consequential Theory, and Virtue Ethics Theory can provide a framework for carefully thinking about SWM as an ethical goal. The differences between the various theoretical camps are not trivial, in that moral philosophers have debated ethics for literally thousands of years. As it may be too strong of an assumption to assume that introductory business students have already taken an ethics or philosophy class, it may be best to discuss and motivate each ethical camp along with covering the SWM in the context of each camp. In general, trying to measure and define ethics in business and finance is a fundamentally tricky exercise. For example, to create an operational definition of (un)ethical behavior, in their study of earnings management in seasoned equity offerings, Jo and Kim (2008) state:

> In this article, we define unethical firms as firms that manipulate their earnings aggressively. (Schipper (1989) defines earnings management as "purposeful intervention in the external reporting process, with the intent of obtaining some private gain to managers or shareholders." Healy and Wahlen (1999) define earnings management as follows: Earnings management occurs when managers use judgment in financial reporting and in structuring transactions to alter financial reports to either mislead some stakeholders about the underlying economic performance of the company, or to influence contractual outcomes that depend on the reported accounting numbers.) Similarly, we define ethical firms as firms choosing ethical reporting, i.e., conservative earnings management.

Jo and Kim (2008) proceed to use the method of Kothari et al. (2005) to construct discretionary accruals and then use this variable to (subjectively) determine ethical or unethical levels of earnings smoothing. It is important to note that IFRS (International Financial Reporting Standards) and GAAP (Generally Accepted Accounting Principles) permit income smoothing through revenue and expense recognition rules. That is, typically firms that manage earnings are acting completely within the law. Whereas Jo and Kim (2008) use the Kothari et al. (2005) method to define ethical and unethical firms, many other studies in accounting and finance use the exact method, or very similar methods, to compute strategically determined "aggressive" or "conservative" levels of earnings management. Within the accounting and finance mainline journals, managing earnings is rarely, if ever, considered unethical until it has crossed the law and has become illegal (see e.g., Brau and Johnson, 2009). Thus, thinking about ethics in the finance paradigm requires a framework, or else the definition and operationalization of ethics becomes purely a subjective exercise. SWM can be presented to students using the four following classic camps of philosophy to form that framework.

#### The Four Ethical Camps

The discussion below is designed as a primer for students who have not taken a philosophical ethics or business ethics class. It intro-
duces four classical ethics camps and when the goal of SWM can be ethical under each camp.

#### Normative Deontology

Under the school of thought of Deontology, right and wrong is defined by a sense of duty. To be ethical, one must act in a way that is intrinsically right or wrong, based on duty, without thought of consequence. The definition of right and wrong is based upon conforming to a moral norm. Perhaps the greatest deontologist, Immanuel Kant, stressed that humanity is the ends and is not to be treated as the means.

Carreira et al. (2008) explain deontology as:

Reflecting upon rules is what is called "deontology", which comes from the conjugation of the Greek words "déon, déontos", which mean duty, and "lógos", that means speech or treaty. Under this perspective, deontology would be the treaty of the duty or the set of duties, principles and norms adopted by a specific professional group. It is a normative subject that deals with the duties that must be followed in specific social circumstances within a specific profession. Deontology is therefore the science that sets the guiding norms of professional activities under the sign of morality and honesty. Pro-fessional deontology systematically elaborates the ideals and the norms that should guide professional activity.

One can imagine under SWM, a professional norm may be that management holds a fiduciary responsibility (duty) to the shareholder. Indeed, given the fact that SWM is overwhelmingly the goal taught to both undergraduate and MBA students, it seems very plausible that many finance professionals (and others) will view SWM as the social norm. In such cases, deontology would argue that the manager has the absolute duty to do everything within his or her means to benefit the shareholder. Logic such as that used in the opening *Economist* quote of this article appeal to deontology. Shareholders are seen as the people who are to be represented at all costs from Kant's point of view.

Of course, the crux of using deontology to support SWM is the underlying assumption that management's duty is to the shareholder.<sup>2</sup> If one makes an alternative assumption, that management's duty should

<sup>&</sup>lt;sup>2</sup> See Morck (2008) who offers rational and behavioral finance explanations for excessive subservience (obedience to superiors, or hyper-duty diligence) in what he terms "type II agency problems." Within deontology applied to finance even the notion of duty can have multiple definitions and levels.

really be directed towards all stakeholders, society, itself, or some other fiduciary, then deontology does not support SWM anymore. Extending the *Economist* example from the introduction, let's say, a manager of a tobacco firm feels a fiduciary responsibility to humanity to decrease the number of smokers. In an attempt to decrease the number of smokers, he or she may purposely break or delay a cigarette manufacturing facility. As a result, the share price of the firm may go down, hurting shareholders (clearly against SWM), but the manager's actions may make it more difficult for some people to smoke, thus relieving them of the health hazards of smoking. In this case, deontology, in a strict sense, would not support SWM. Under deontology, the outcome of shareholders losing money or people smoking less is actually irrelevant—the consequences are not the determining factor. The determining factor is that management honors his or her duty.

To decide between strict SWM and some other goal as being justified by deontology, one must agree on the appropriate "moral norm." By definition though, a moral norm is not necessarily a fixed law with which every market participant agrees. Students should understand that the necessary and sufficient condition for SWM to be deemed ethical under the deontology framework is that the moral norm defines management as holding a fiduciary responsibility (duty) strictly to the shareholder. If the duty of management is placed anywhere else, wholly or partly, then SWM would be deemed unethical if management duty was not exercised in those areas.

Adhering to deontology, the study by Carreira et al. (2008) provides evidence that an individual's ethical behavior can be increased by ethics and deontological education. (See Cagle and Baucus (2006) for another example of education promoting personal finance ethics and Cagle et al. (2008) for an example of ethics education changing student perceptions of business ethics.)

> Concerning the cognitive dimension of attitude towards ethics, the subject of ethics and professional deontology strengthened the answers to the questions with lower scores [of ethical behavior]. Concerning the affective/assessing attitude of ethics, the subject of ethics and professional deontology strengthened the students' convictions about the importance of the existence of a deontological code, of ethical principles and of accounting information, as well as the question with the lowest score (the entity's interest is more important). (Carreira et al. (2008)

Under deontology, these respondents adhered more towards social norms that the authors considered ethical. Herein lies another caveat to

the evaluation of SWM as ethical. Not only must the fiduciary be established, as discussed above, but the social norm must be agreed upon to determine ethicality. If alumni of finance programs who follow the school of Berle (1931), Manne (1959), and Friedman (1970) strongly feel that SWM is the appropriate social norm for the goal of the firm, but other students trained under the school of thought of Dodd (1932) and more modern-day CSR advocates support a stakeholder social norm, the ethicality of SWM again becomes contingent. This contingency, based on the selection or belief in a specific social norm, determines deontological ethicality of the SWM paradigm.

Survey data of DeLoughy et al. (2011) from 680 respondents offers some evidence of finance professionals' views of norms as they pertain to CSR and shareholder value. The authors point out that nearly 80% of finance professionals agree that firms "have a social responsibility beyond their interests of the shareholders, with only 4.8% not concurring." This result suggests that a norm other than strict SWM seems to exist in practice. However, the next question, "The socially responsible manager must occasionally place the interests of society over the interests of the company," only receives 40% concurring, with 21.8% not concurring and the rest remaining neutral. This latter question seems to suggest that the SWM goal continues to be persuasive when finance professionals must choose directly between shareholder wealth and social responsibility. These survey data seem to indicate that agreement upon the social norm of the intent of the firm is not very universal and, as such, the deontological framework of ethics becomes vague as firm managers have ambiguity as to who their fiduciaries are and to whom they should be duty-bound.<sup>3</sup> Such an idea may promote healthy debate among students.

# Justice/Fairness Theory

Based on the work of Rawls (1971), in societies with sufficient surplus (such as one with profitable corporations), it is argued that surplus should be allocated to individuals within society. The justice the-

<sup>&</sup>lt;sup>3</sup> Another complication of defining a single social norm is that social norms may vary by culture and geography. For example, Phatshwane et al. (2014) show differences in ethical beliefs between students in Botswana, South Africa, and Turkey who were in accounting and finance classes. In addition, Chang (2004) concludes geopolitical culture affects ethics when he states, "Low politics and civility, collectivism and closed ethical boundaries, formality-based pretentiousness, and public mistrust and discredit are seen as major cultural factors contributing to Korea's unhealthy state of financial ethics." In a survey of eight countries, Statman (2007) demonstrates that "ethics" varies around the world because of culture, globalization, income, education, and law enforcement.

ory argues for the optimal way to distribute this surplus. It is based on two primary principles: 1) "each person is to have an equal right to the most extensive scheme of basic liberties compatible with a similar scheme of liberties for others," and 2) "social and economic inequalities are to be arranged so that they are both reasonably expected to be to everyone's advantage, and attached to offices and positions open to all." (Rawls, 1971). Similar to utilitarianism, discussed below, justice theory believes in using corporate surplus to benefit society directly, but unlike utilitarianism, justice theory prescribes a specific prescriptive path on how to allocate the surplus.

Under the justice theory paradigm, students can discuss whether SWM is inherently unethical in that the surplus of corporations go to the shareholders, as they are the residual claimants. The shareholders themselves then decide what they prefer to do with the surplus. Surplus is not allocated through the Rawls's system of justice. Perhaps the only way SWM is ethical under justice theory is if shareholders take all surplus and allocate it through society according to Rawls's method. Of course, this is highly unlikely in practice, as shareholders all have their own unique utility functions and for 100% of shareholders to follow the justice paradigm is not a practical assumption. Students may debate about fundamental assumptions of human nature as related to justice theory and conclude for themselves whether this philosophical camp adheres to their own vision of the business world.

#### Utilitarianism Theory

Utilitarianism falls under a class of ethics known as consequentialist theories. For the utilitarian, ethicality is based on the end result, or consequence, of each action. The key to utilitarianism is the basis of a Pareto optimum, that is, maximizing society's benefit (good) while minimizing costs (bad). There are various types of utilitarianism such as hedonistic, preference, act, and rule-based utilitarianism. Within the umbrella of utilitarianism though, is the idea that: 1) it is the consequence of a decision that matters and 2) decisions should be made to maximize the benefit (and minimize the cost) to society.

The framework of utilitarianism becomes both a philosophical and empirical matter to determine whether SWM is the mechanism that provides the greatest good for society at the lowest cost (bad). For example, Jensen (2002) and Friedman (1970) both argue that SWM is the optimal mechanism to meet utilitarianistic goals in society. However, Jones and Felps (2013a, 2013b) argue the opposite.

Jones and Felps (2013a) actually title their paper, "Shareholder Wealth Maximization and Social Welfare: A Utilitarian Critique." The

intent of Jones and Felps is to argue that SWM "is not a good way to promote social welfare." Indeed they conclude, "By casting serious doubt on one of market capitalism's most entrenched institutions shareholder wealth maximization—we hope to alter that way of thinking." The approach of Jones and Felps (2013a) is the closest to this current paper in that they consider SWM under an ethical camp utilitarianism. It differs, however, in that their purpose is to attack SWM and my purpose is to illustrate what conditions are necessary for SWM to be considered ethical in a pedagogical framework. Having said this, Jones and Felps provide a very thorough and thoughtful critique of SWM. In their analysis, they are careful to make a distinction between act and rule utilitarianism (the two in their study).

> Act utilitarianism instructs the agent to make the decision that results in the greatest net social benefit *with respect to the decision at hand. Rule* utilitarianism, on the other hand, directs the agent to follow rules that are intended to produce the greatest net social benefit *over time*. (Jones and Felps, 2013a, p. 212, italics theirs)

They provide a hypothetical that if a defense attorney knows that her client is guilty of murder, under act utilitarianism, she may provide a weak defense or no defense at all to get the murderer off the street. In this case, it benefits society the most with the decision at hand. Under rule utilitarianism, however, she should vigorously defend the murderer for acquittal to uphold the criminal justice system, which is intended to produce optimal legal outcomes over time. Jones and Felps (2013a) cite prior literature on SWM and conclude that SWM is a rule-utilitarian part of the capitalist system designed to provide long-term social benefits.

Having set the stage under a rules-based utilitarian system, Jones and Felps examine the logic of going from SWM to social welfare improvement and argue that the logic flow is flawed:

> In summary, four conclusions are relevant: (1) many markets are not sufficiently competitive, (2) maximizing shareholder wealth is not always the best way to achieve firm efficiency, (3) increases in efficiency may not increase aggregate economic welfare, and (4) greater economic welfare is only weakly linked to greater human happiness. The theoretical chain connecting SWM to improved social welfare has substantial weaknesses at every link. (Jones and Felps, 2013a, p 224)

In their own way, Jones and Felps (2013a) outline when (the necessary and sufficient conditions) SWM is ethical under a utilitarian framework. Given efficient (or mostly efficient) markets, situations where SWM achieves firm efficiency, contexts where increased firm efficiency does increase economic welfare, and cases where greater economic welfare helps promote human happiness, then SWM is an ethical goal. Jones and Felps argue that all four of these situations must be in place and if more than one does not hold, the problem becomes multiplicatively worse. Students may participate in a robust discussion between the Jones and Felps camp and the Friedman and Jensen camp as it applies to the utilitarian framework as they consider the efficacy and ethicality of SWM.

#### Virtue Ethics

Derived from the work of Aristotle, virtue ethic theory argues that ethics depends on the character traits (virtues) of the people making decisions (agents). The decision for action then is not measured by an allegiance to duty of a moral norm (deontological), an appeal to fairness of surplus allocation (justice theory), or the societal outcome (utilitarianism), but instead from an internal virtue within the person deciding on the action. The underlying assumption of virtue ethics is that agents either intrinsically have or can develop virtuous character traits and avoid negative character traits. Any positive character trait can be claimed with the virtue ethics tradition, and moral education is seen as a path to help people become more ethical. As virtues increase, ethics increase.

A thorough presentation of the rationale underlying virtue ethics is provided in a series of scholarly articles that John Dobson published in the *Financial Analyst's Journal*. His first piece printed in 1993 is titled, "The Role of Ethics in Finance." The second article printed in 1997 is titled, "Ethics in Finance II." His third article in this thread, printed in 1999, is closest to our study and is titled, "Is Shareholder Wealth Maximization Immoral?" (Dobson 1993, 1997a, 1999).

Dobson (1993) speaks more generally about ethics vis-à-vis the field of finance than limiting to the goal of the firm. The sentiment of the paper, however, can be directly applied to the goal of the firm. Dobson (1993) states, "Contemporary financial economists view ethics in the context of objective wealth maximization. In this context, ethics functions primarily as a constraint on behavior." He goes on to argue that the logic driving the wealth maximization objective, and the ethics constraint, has serious problems. Defining ethics as the motivation for what people do, and not necessarily the actions that people take, Dob-

son argues that one can judge an individual only if his or her motivations are known and not by observing actions. Dobson goes on to argue that the goal of the firm should be the opposite of Friedman's (1970) view (that ethics is a constraint and profits are the endgame) in that, "If financial theory recognizes ethical motivations, these motivations will be implicitly condoned. The accumulation of external goods, namely wealth, will then be viewed as merely a means to the ultimate end of achieving internal goods, namely respect and integrity." In essence, Dobson (1993) is a plea for finance practitioners to set the real goal of finance to be gaining "respect and integrity." Dobson argues that "throughout the 2000-year history of moral philosophy, ethics has generally been viewed as a behavioral motivation, not as a constraint." Extending this logic to the goal of the firm, SWM is thus seen as a means to achieve the true goal of the firm—acquisition and maintenance of respect and integrity.

In Dobson (1997a), he reiterates that, "Ethics becomes an ideal or an excellence to be pursued as an ultimate objective." This follow-on article to the 1993 piece seeks to clarify the earlier arguments. Dobson (1997a) appeals to virtue ethics theory as the primary driver for a finance professional. The virtue ethics theory stresses personal virtue and the quest for "moral excellence as a goal in and of itself." From Dobson (1997a):

> Virtue ethics is concerned with pursuing a certain type of morally inclusive "excellence." In The Nicomachean Ethics, Aristotle called it *eudaimonia*, which can be roughly translated as happiness or human flourishing through moral excellence. For present purposes, this approach to ethics can be thought of as exhibiting four basic attributes. Its primary attribute is a strong emphasis on the importance of certain generally accepted virtues of character; indeed, through honing and perfecting these virtues, an individual becomes truly ethical. Second, strong emphasis is placed on the existence of an active community that nurtures these virtues. Third, virtue-ethics theory makes clear that in the moral life, one cannot rely merely on rules or guidelines; in addition, an ability to exercise sound moral judgment is requisite. Finally, the successful identification and emulation of moral exemplars or role models is essential for the dissemination of morality within the nurturing community.

Dobson (1997a) expands and applies this framework of virtue ethics to finance professionals. The paper does not fully develop the corre-

lation of virtue ethics theory and the goal of the firm. That is, SWM could be either ethical or nonethical based upon why individuals do what they do in the firm. If one manager maximizes shareholder wealth because: 1) she has the positive virtues of loyalty, diligence, and hard work; 2) she feels inside like she is representing an active community of shareholders who are depending on her to fund their retirements, kid's college funds, medical costs, etc.; 3) she exercises sound moral judgment in her fiduciary responsibility to her ultimate bosses (i.e., these same shareholders); and 4) she not only models herself after highly ethical role models, but serves as an excellent role model herself in her community, then by Dobson (1997a), this manager is acting ethically by maximizing shareholder wealth. On the other hand, if the same manager instead does the identical actions, but her goal is simply to enrich herself and not the four steps above, then she would be acting unethically. In the latter case, the goal of shareholder maximization would be seen as unethical under the virtue ethics paradigm.

Dobson (1997b) is an entire book dedicated to the topic of finance ethics and is titled, *Finance Ethics: The Rationality of Virtue*. As the title suggests, Dobson relies heavily on virtue ethics theory to motivate his arguments in this book, very similar to his 1997 *Financial Analyst Journal* discussed above. He expounds on Dobson (1997a) considerably through the rational expectations paradigm of economics and arguing for the rationality of the virtue ethics paradigm. From Dobson (1997b):

> This approach [to modeling rational behavior in finance] does not attempt to add on ethics as some form of appendage or constraint to the existing theory; rather it goes right to the heart of the theory: to the finance paradigm's very notion of rationality. I draw on the moral philosophy of virtue ethics to provide an alternative rationality premise for financial economics. This alternative premise brings notions of ethical behavior within the rationality rubric. Thus, such action as 'honoring trust' become rational in and of themselves, and do not have to be justified in material terms.

In his book, Dobson focuses on the definition of a firm and how the firm is viewed from an economical vis-à-vis an ethical perspective. He motivates the book with the SWM theory in the opening paragraph of the introduction and then immediately challenges SWM in the second paragraph with potential alternative goals of the firm. Dobson argues that the fields of business ethics and financial economics ask similar questions as to the purpose of the firm, but "Even the most cursory review of their respective literatures, however, will reveal that they tend to do so from distinctly different perspectives." The bulk of the book is therefore Dobson's attempt to bridge this gap, by proposing the rationality (i.e., financial economics) of virtue ethics (i.e., business ethics).

Because of sufficient confusion left from Dobson (1993) and Dobson (1997a), Dobson tackles the issue of the SWM head on in Dobson (1999). He argues that shareholder maximization is not amoral (neutral) or immoral (bad). After dismissing SWM as amoral or immoral, he sets out an argument for under what conditions the SWM goal may be "moral," which is the answer to the article title, "Is Shareholder Wealth Maximization Immoral." His argument is based on his earlier articles, particularly on virtue ethics. He states:

> In accepting shareholder wealth maximization as the objective, however, business professionals should not abrogate all moral common sense. As a version of might makes right, shareholder wealth maximization clearly needs tempering, which is where the character and judgment of the manager comes in.... In exercising the judgment necessary to answer such questions, ethical character traits, virtues, are essential virtues such as prudence, courage, wisdom, and compassion. Only through sound moral judgment on the part of individual managers can the organizational premise of shareholder wealth maximization be morally justified.

This quote is the crux of the article. In fact, Dobson footnotes the second to the last sentence in this paper with a note to "See my 1993 and 1997 articles." Dobson's (1999) extends his earlier paper's assertion that it is not revealed preferences that indicate ethics, but the motivation of the manager that indicates ethics, and as such, it is individual-specific to the manager as to whether the goal of SWM is moral or not.

The articles by Dobson can vitally contribute to class discussion on the ethicality of SWM as the goal of the firm. For example, students could debate whether Dobson's logic for his 1999 article suffers from the same critique as the 1997 article that we discuss above in that the goal of SWM is a behavior that we question, as a mantra taught in many finance classes, and not the motivation behind the behavior. In addition, discussion of Zingales (2015) could lead to student interaction about what conditions must exist for SWM to benefit society better than some other alternative goal.

# Implementation in the Business Course and Conclusion

# Implementation of the Four Ethical Camps and SWM in the Business Class

Having discussed a brief history of SWM and how it has traditionally been taught in business (especially finance) classes, four classical ethics camps, and the conditions under which they motivate SWM as an ethical goal of the firm, I now discuss application of the framework in business courses. As the goal of the firm is a central element of any business class, a statement in the syllabus encouraging students to think throughout the semester how they can apply these ethical paradigms to the goal of the firm may be warranted. For example, in a finance class, there may be a statement in the syllabus such as:

> In this class, we will learn how financial managers make decisions as to what assets to buy for the firm and the best ways to finance these investments. As you learn these finance skills, think about how they contribute to the goal of the firm and how ethics informs these management decisions. If the goal of the firm is to maximize shareholder wealth, think about under what conditions the particular financial skill you are learning can fit within, or outside, various ethical paradigms such as utilitarianism, deontology, justice, and virtue ethics.

Such a statement in the syllabus could help students to not blindly accept SWM as "[t]he only legitimate criterion in managerial decision making is stockholders' welfare, as measured by their wealth" (Levy, 1998) as quoted in the introduction. If the goal of the firm is discussed in the first or second class period to include the four ethics paradigms, students have a framework for thinking about SWM (or alternative goals) for the remainder of the semester.

Each lecture during the course, professors can apply the concepts of this article into student discussion. For example, a supply chain professor who is teaching about the pros and cons of offshoring labor can encourage a discussion of under what conditions is this an ethical supply chain to maximize shareholder wealth? Students may bring up topics such as how direct foreign investment helps workers out of poverty in impoverished countries. Others may discuss how child labor laws are not enforced in some economies and how children work in horrible conditions for virtually no pay. Professors may lead discussions about a business decision being legal versus being ethical. By considering more than one ethical camp, such as the four discussed in this article, students have a framework of how to define ethical behavior. Professors can use the four ethics camps of deontology, justice/fairness theory, utilitarianism/consequential theory, and virtue ethics theory as the framework for teaching students how to think carefully about the goal of the firm.

Instructors can test the knowledge and thought process of students through short-answer or essay questions. Instead of multiple-choice questions simply asking: "What is the goal of the firm?" and one of the choices being, "To maximize shareholder wealth;" students could be asked to, "Write a concise essay on under which conditions SWM is an ethical goal using the four philosophical camps of deontology, justice, utilitarianism, and virtue ethics." Students who can carefully think through these types of thought questions pertaining to the goal of the firm in a traditional business class (and not just in their business ethics course) may be better prepared to face the real world of business after graduating.

#### Summary and Conclusion

For decades, introductory business students have been taught in business (and particularly finance) textbooks that the goal of the firm is simply "to maximize shareholder wealth." This goal, abbreviated SWM, may actually lead to unethical behavior if taken to the extreme. In this article, I provide a framework on how to think about SWM using four classical philosophy camps and how to potentially discuss it with students when teaching the goal of the firm.

Assuming these students have little or no philosophical ethics training, I appeal to the four main camps of ethics: deontology, justice/fairness theory, utilitarianism/consequential theory, and virtue ethics theory, at foundational levels. Table 1 summarizes the ethics camps and illustrates how they relate to SWM. This table could be used in class as a concise slide to explain the relationship between SWM and the four ethics camps. If a more thorough coverage is desired, students could be assigned to read this article as part of class readings.

Discussing the caveats of each camp may provide a framework for students to not blindly accept SWM as the only and best corporate goal. In fact, the answer to the question of whether SWM is an ethical goal of the firm depends on the underlying school of ethics the person in judgment accepts. Except under very specific conditions, students may discover that SWM can be seen as an unethical goal of the firm. However, in some settings, SWM can be the ethical approach for the goal of the firm. It is these underlying assumptions and predispositions that have driven the literature for and against SWM. The work of Berle, Mann, and Friedman represents the work of authors who lived in an ethical paradigm, informed mostly by economics, which support their arguments in favor of SWM. Most authors of finance textbooks rely on this same logic to teach modern-day undergraduates and MBA students in their finance classes that the goal of the firm is "to maximize shareholder wealth" complete with italicized and bolded fonts.

Table 1. Major Ethical Traditions and Their Relationship with SWM			
Ethical Tradition	Summary	Conditions Under Which SWM Is Ethical	
Deontology	Focuses on actions, does not waver according to situations. Certain acts are intrinsically right or wrong. The goal is honor one's duty.	The moral norm is that management holds a fiduciary responsibility (duty) strictly to the shareholder.	
Justice/Fairness	Focuses on allocation of surplus. The goal is to allocate surplus to society in the optimal manner.	Any surplus paid to shareholders is then in turn allocated according to Rawls's (1971) allocation method of justice.	
Utilitarianism/ Consequentialism	Focuses on the ends rather than the means. The goal is to maximize society good (or minimize society bad)	SWM provides the optimal mechanism for maximizing good (or minimizing bad) for society as a whole.	
Virtue Ethics	Focuses on the intrinsic virtues of the individual. The goal is to develop and act on positive character traits.	SWM is ethical when management acts in accordance with internal character traits (virtue).	

Those who disagree with the SWM goal and argue for its unethicality approach the question from different paradigms. Their underlying assumptions and views of the fundamental framework from ethics allow them to argue against SWM with strong and clever cases. As analyzed in this article, both sides of the argument may be right, and both may be wrong, depending on what core assumptions are made pertaining to duty, fiduciaries, norms, and which school of ethics to accept. Inclass discussions of these assumptions have the promise to address the concerns of Hawley (1991) and others that the nearly universal goal in business/finance texts of SWM may incorrectly motivate business school graduates to act unethically to maximize shareholder wealth, as seen in cases such as Enron and Worldcom.

# References

Berle, Adolf A., 1931, Corporate Powers as Powers in Trust. *Harvard Law Review* 44, 7, 1049-1074.

Bernardi, Richard A., Michael R. Melton, Scott D. Roberts, and David F. Bean, 2008, Fostering Ethics Research: An Analysis of the Accounting, Finance and Marketing Disciplines, *Journal of Business Ethics* 82, 157-170.

Block, Stanley B. and Geoffrey A. Hirt, 2005, *Foundations of Financial Management*, 11<sup>th</sup> ed., McGraw-Hill, Boston.

Brau, James C., and Peter M. Johnson, 2009, Earnings Management in IPOs: Post-engagement Third Party Mitigation or Issuer Signaling? *Advances in Accounting* 25, 125-135.

Cagle, Julie A.B., and Melissa S. Baucus, 2006, Case Studies of Ethics Scandals: Effects on Ethical Perceptions of Finance Students, *Journal of Business Ethics* 64, 213-229.

Cagle, Julie A.B., Philip W. Glasgo, and Vanessa M. Holmes, 2008, Using Ethics Vignettes in Introductory Finance Classes: Impact on Ethical Perceptions of Undergraduate Business Students, *Journal of Education for Business* 84, 2, 76-83.

Carreira, Francisco A., Maria do Amparo Guedes, and Maria da Conceição Aleixo, 2008, Can we Teach Ethics and Professional Deontology? An Empirical Study Regarding the Accounting and Finance Degree, *Social Responsibility Journal* 4, 89-103.

Chambers, Donald R., and Nelson J. Lacey, 1996, Corporate Ethics and Shareholder Wealth Maximization, *Financial Practice and Education*, 6, 1, 93-96.

Chang, S. J., 2004, A Cultural Inspection of Financial Ethics in South Korea, *The Asia Pacific Journal of Economics and Business*, 8, 2, 65-77.

DeLoughy, Sara T., K. Gregory Jin, and Ronald Drozdenko, 2011, Corporate Social Responsibility and Organizational Ethics Research Issues: A Preliminary Look at Finance Professionals' National Survey Data, *Journal of Leadership, Accountability and Ethics* 8, 4, 9-25.

Dobson, John, 1993, The Role of Ethics in Finance, *Financial Analysts Journal* 49, 57-61.

Dobson, John, 1997a, Ethics in Finance II, *Financial Analysts Journal*, 53, 1, 15-25.

Dobson, John, 1997b, *Finance Ethics: The Rationality of Virtue*, Rowman and Littlefield Publishers, Oxford.

Dobson, John, 1999, Is Shareholder Wealth Maximization Immoral? *Financial Analysts Journal* 55, 5, 69-75.

Dodd, E. Merrick, Jr., 1932, For Whom Are Corporate Managers Trustees? *Harvard Law Review* 45, 7, 1145-1163.

Donaldson, Thomas, 1982, *Corporations and Morality*, Prentice Hall, Upper Saddle River, NJ, p 57.

The Ethics of Business: Good Corporate Citizens, Wise Governments, Should be Wary of CSR, *The Economist* 374, 17-S22, 1/20/2005.

Ford, H., and J.M. Miller. 1922. The Amazing Story of Henry Ford: The Ideal American and the World's Most Famous Private Citizen; a Complete and Authentic Account of His Life and Surpassing Achievements. M. A. Donohue & Company, Chicago.

Friedman, Milton, The Social Responsibility of Business is to Increase its Profits, *New York Times Magazine*, 9/13/70.

Griffin, J.J., and J.F. Mahon, 1997, The corporate social performance and corporate financial performance debate: Twenty-five years of incomparable research, *Business and Society*, 36, 1, 5-31.

Hawley, Delvin D., 1991, Business Ethics and Social Responsibility in Finance Instruction: An Abdication of Responsibility, *Journal of Business Ethics*, 10, 711-721.

Healy, P., and J. Wahlen, 1999, A Review of the Earnings Management Literature and Its Implications for Standard Setting, *Accounting Horizons* 13, 365–383.

Horrigan, James O., 1987, The Ethics of the New Finance, *Journal of Business Ethics*, 6, 2, 97-110.

Jensen, M. C. 2001. *A Theory of the Firm: Governance, Residual Claims, and Organizational Forms.* Harvard University Press, Cambridge, MA.

Jensen, M. C. 2002. Value maximization, stakeholder theory, and the corporate objective function. Business Ethics Quarterly, 12, 235-256.

Jo, Hoje, and Yongtae Kim, 2008, Ethics and Disclosure: A Study of the Financial Performance of Firms in the Seasoned Equity Offerings Market, *Journal of Business Ethics* 80, 855-858.

Jones, T. M., and W. Felps, 2013a. Shareholder wealth maximization and social welfare: a utilitarian critique. *Business Ethics Quarterly*, 23: 207-238.

Jones, T. M., and W. Felps, 2013b. Stakeholder happiness enhancement: a neo-utilitarian objective for the modern corporation. *Business Ethics Quarterly*, 23: 349-379.

Keown, Arthur J., John D. Martin, J. William Petty, and David F. Scott, Jr., 2005, *Financial Management: Principles and Applications*, 10<sup>th</sup> Edition, Pearson/Prentice Hall, New Jersey.

Kothari, S. P., A. J. Leone and C. Wasley, 2005, Performance Matched Discretionary Accrual Measures, *Journal of Accounting and Economics* 39, 163–197.

Levy, Haim, 1998, Principles of Corporate Finance, South-Western, Cincinnati.

Manne, Henry G., 1959, Accounting for Share Issues Under Modern Corporation Laws, *Northwestern University Law Review* 54, 285.

Manne, Henry G., 1961, Current Views on the "Modern Corporation," *University of Detroit Law Journal* 38, 559-588.

Manne, Henry G., 1962a, The "Higher Criticism" of the Modern Corporation, *Columbia Law Review*, 62, 399-432.

Manne, Henry G., 1962b, Corporate Responsibility, Business Motivation, and Reality, *The Annals of the American Academy of Political and Social Science*, 343, 55-64.

Manne, Henry G., 1964, Some Theoretical Aspects of Share Voting— An Essay in Honor of Adolf A. Berle, *Columbia Law Review*, 64, 1427

Manne, Henry G., 1965, Mergers and the Market for Corporate Control, *Journal of Political Economy*, 73, 110-120.

Mitchell, Ronald K., Gary R. Weaver, Bradley R. Agle, Adam D. Bailey, and James Carlson, 2016, Stakeholder agency and social welfare: Pluralism and decision making in the multi-objective corporation, *Academy of Management Review* 41, 2 252-275.

Morck, Randall, 2008, Behavioral Finance in Corporate Governance: Economics and Ethics of the Devil's Advocate, *Journal of Management and Governance*, 12, 2, 179-200.

Phatshwane, Percy, Mogotsinyana Mapharing, and Elang Basuhi, 2014, Attitudes towards Business Ethics Held by Accountancy and Finance Students in the University of Botswana, *International Journal of Business and Management* 9, 2, 17-28.

Rao, S.M., and J. B. Hamilton, III, 1996, The effect of published reports of unethical conduct on stock prices, *The Journal of Business Ethics* 15, 1321-1330.

Rawls, John, 1971, A Theory of Justice, Harvard University Press, Cambridge.

Roman, Ronald M., Sefa Hayibor, and Bradley R. Agle, 1999, The relationship between social and financial performance, *Business and Society*, 38, 1, 109-125.

Ross, Stephen A., Randolph W. Westerfield, and Bradford D. Jordan, 2013, *Fundamentals of Corporate Finance*, 10<sup>th</sup> Edition, McGraw-Hill/Irwin New York.

Schipper, K., 1989, Commentary on Earnings Management, *Accounting Horizons* 3, 91–102.

Shlensky v. Wrigley, 237 N.E.2d 776, 778, (Ill App 1968).

Statman, Meir, 2007, Local Ethics in a Global World, *Financial Analysts Journal* 63, 3, 32-41.

Stewart Jr., Fenner L., 2011, Berle's Conception of Shareholder Primacy: A Forgotten Perspective For Reconsideration During the Rise of Finance, *Seattle University Law Review*, 34, 1457.

Zingales, Luigi, 2015, Presidential Address: Does Finance Benefit Society? *Journal of Finance* 70, 4, 1327-1363.

# **Evaluation of the Effect of Study Skills and Lifestyle Factors on Performance in Organic Chemistry**

Don R. Davies, Heather Root, Valerie Herzog

Weber State University

#### Abstract

Student performance in Sophomore Organic Chemistry courses was measured against lifestyle and study skill factors at a large, openenrollment, public university in the Mountain West region. The survey items were first evaluated individually for correlation to student actual performance on a midterm examination. Items requiring significant student output, such as verbally explaining concepts, participating in study groups, and working practice exams, showed highest correlation to actual performance. Student chemical foundation, as measured by their 2<sup>nd</sup> semester General Chemistry grade, also was a significant contributor to student success in Organic Chemistry. A factor analysis grouped the items into 4 general categories including sleep patterns, student output (study groups, verbally explaining principles), student input (traditional learning methods such as attending lecture, reading the textbook, and being tutored), and foundation and attitude, which included prior performance, level of anxiety experienced on exams, and how well they liked Organic Chemistry. The factor that correlated most strongly to actual performance was foundation and attitude, but that was closely followed by student output activities. Based on the outcomes of this study, emphasis has been placed on providing in-class opportunities for students to work with a classmate or in groups of 5 or 6 students in solving and discussing problems. Furthermore, student preparation has moved beyond just completion of homework problems to teaching principles to another individual and creating assessment questions. So far, the outcome of these modifications is that the top students in the class perform about the same as before, but student retention has significantly improved.

#### Introduction

This study was conducted at Weber State University, a large, open-enrollment, public university located in Ogden, Utah. Of the 25,000+ students at Weber State University, 56% are nontraditional students, meaning they are over 25 years of age, married/widowed/ divorced, or are a parent. Because of additional responsibilities held by these students to provide for their families, most work 20+ hours/week.

Besides being a requirement for Chemistry majors, Organic Chemistry is a prerequisite course for many graduate programs. As one of the most difficult classes required for acceptance into these programs, Organic Chemistry often becomes an insurmountable barrier, sending many students to alternative career paths. In the Organic I and II Chemistry classes taught by the first author during the past 5 years. the average percentage of students who passed with a C grade or better was 57.0% and 65.8% for the two courses, respectively. Factors that may contribute to the difficulty of Organic Chemistry are a combination of its uniqueness from previous completed physical science courses, the need for spatial perception,<sup>1</sup> attention to detail, the creative use of basic principles to predict reactivity and to do it multiple steps at a time, compared with simply identifying data and plugging the data into equations as is often done in General Chemistry. Therefore, advancing from General Chemistry to Organic Chemistry is often a big step for students at this level of their education. As students have struggled to perform up to their expectations, they have sought direction from their instructor on how to improve their performance in the course. Based on conversations with and observations of students who have failed and succeeded in the course during the past 15 years of teaching this course, the first author has come to counsel current students to 1) dedicate additional time to studying by reducing other time

constraints, 2) complete all forms of assessment including homework that was assigned but not graded and practice examinations and guizzes from previous semesters made available on the instructor's website with a focus on comprehension over just completion,<sup>2</sup> 3) engage in activities that require students to incorporate what they read by highlighting text and notes and summarizing content in their own words, and 4) regularly attend supplemental instruction sessions. The authors wanted to determine which factors contributed most strongly to the student's actual performance on a summative assessment item, such as a midterm exam. To carry out this study, we devised a survey that students had the opportunity to voluntarily complete at the end of one of their midterm exams. Items chosen for the survey included the four factors described above along with assessment of the students' past success in chemistry as indicated by their grades from the second semester of General Chemistry, the effect of whether they enjoyed the subject or not, the level of anxiety they experienced, and the effect of sleep patterns, which were shown to strongly influence academic performance in other studies as discussed below. This research investigated which study practices and lifestyle factors most strongly influenced actual student performance on midterm exams in Organic I and Organic II Chemistry courses at a large, public university.

# **Related Work**

Student ability to assess their level of comprehension depends largely on the complexity of the knowledge being assessed and how it is assessed. Student learning can be divided into lower-order cognitive skills (LOCS) and higher-order cognitive skills (HOCS).<sup>3,4</sup> LOCS include activities such as information retrieval and recognition, generally assessed by multiple-choice-type questions. HOCS activities generally require free-response-type answers and frequently assess application of principles, production of reaction mechanisms, differentiation of reactions, and deductive reasoning. The difference between fluency in LOCS activities compared with fluency in HOCS activities has been compared with being able to read a foreign language but not speak it.<sup>5</sup> Tsaparlis et al. demonstrated that students accurately self-assess their performance in LOCS activities as compared with an assigned grade by their professor.<sup>3</sup> However, in HOCS activities, they significantly overestimated their actual performance as measured against an assigned grade given by a professor. Organic Chemistry represents a new way of thinking for most students in that it requires students to use basic principles in solving new and sometimes complex problems. Students are required to differentiate one reaction type from another type.<sup>4</sup> Success

in the subject of Organic Chemistry requires student ability in using both LOCS- and HOCS-type activities.

In addition to levels of comprehension, lifestyle factors have also been associated with student performance and ability to assess level of preparedness. A study involving 200 students living in on-campus residence halls examined variables affecting grade point average (GPA), such as exercise, eating, sleep habits, mood states, perceived stress, time management, social support, spiritual or religious habits, number of hours worked per week, sex, and age. Of these factors, sleep habits accounted for the greatest variance in GPA.<sup>6</sup> In addition to total amount of sleep, a consistent sleep-wake pattern was strongly related to academic success.<sup>7,8</sup> Sleep deprivation was found to mainly affect higherorder cognitive functions such as attention, memory, and problem solving.<sup>9,10</sup> Adequate sleep after rigorous studying or training aids in retention of information and ability to a perform task.<sup>11,12</sup> A study involving 44 college students subjected to 1 night of total sleep deprivation demonstrated the expected performance decrease on cognitive deduction and recognition of assumptions. Interestingly, these same students reported higher levels of estimated performance and greater effort expended than non-sleep-deprived subjects, thus showing that sleep deprivation alters a student's awareness of actual ability.<sup>13</sup>

Frequently, students report that they knew the material but performed poorly on a test because of anxiety. Anxiety disrupts focus on task-related activities and is correlated with lower exam performance.<sup>14</sup> Factors contributing to anxiety include focusing on outcomes more than processes, especially outcomes that are not entirely in the student's control such as questions on an exam and how those questions are scored. Anxiety can also come when outcomes from studying are less than a student's expectations, creating a negative academic cycle. The more anxiety a student has, the less focus they have on the comprehension, leading to lower productivity and increased anxiety. Twenty-nine percent of students with high test anxiety have GPAs below the 2.0 mark required for graduation. Students with high test anxiety have less effective study skills, which can be partially compensated for by an increased number of hours studying.<sup>15</sup>

Student awareness of the depth and breadth and development of their comprehension comes best when they have the opportunity to practice "speaking the language" in settings such as collaborative and cooperative learning formats.<sup>16–23</sup> Collaborative and cooperative learning methods allow large classes to be broken down into small group sizes, where individuals have greater participation and accountability.<sup>21</sup> These groups are assigned problems to solve. In the process, students learn to work and communicate together and support each other. Some-

times these groups are initiated with a written contract and are rewarded for the success of fellow group members. <sup>19,20</sup> These groups are often led by a student who has previously completed the course successfully, similar to our Supplementary Instruction sessions at Weber State University.<sup>24–27</sup> Verbally explaining principles causes students to think about, associate, and gain ownership of what they have learned and become keenly aware of what they have not yet learned.

The facilitating of group activities is greatly enhanced with the availability of powerful and inexpensive applications (apps) for laptop computers, tablets, and smart phones.<sup>28–38</sup> Using this available technology, teachers can create groups and provide available sites on Canvas or other learning management systems where students can virtually assemble and converse. The idea of a flipped classroom is gaining momentum with the availability of screencasting apps (apps that enable the projecting of information produced on a tablet onto a remote screen) such as Explain Everything, Air Sketch, and Doceri.<sup>32,33,39,40</sup> The app Notability, installed on an iPad, has been implemented in making an Organic Chemistry course nearly paperless, as the app functions not only as a note-taking device, but also as a laboratory notebook.<sup>28</sup> Several other studies have focused on the use of technology that allows for the success of small-group activity inside or outside of a physical classroom.<sup>18,32,38,41</sup> In addition to screen-casting lectures outside of class and the facilitation of group work, technology also allows for rapid feedback, which helps clear up misconceptions early.<sup>38,42</sup> Audience survey items are also available to help an instructor assess the understanding of a class during a lecture using cell phones or smart phones.<sup>25,31,37,43</sup> All of these forms of technology can help to cultivate a rich active learning environment for students, where the majority of the time is devoted to student participation and expression.

#### **Experimental Methods**

The survey instrument was administered by Dr. Don Davies to his students in three Organic Chemistry II classes, constituting 85 students, and two Organic Chemistry I classes, constituting 107 students (192 students total), beginning with the Spring 2013 semester and ending at the conclusion of the Spring 2014 semester.

Students in Organic Chemistry I were assessed by three midterm exams and a comprehensive final, while Organic Chemistry II students were subject to four midterm exams and the ACS national exam for Organic Chemistry as the course final. Students in both classes were given the survey instrument on either their 2<sup>nd</sup> or 3<sup>rd</sup> midterm exam, so that enough time had elapsed for study skill and lifestyle patterns to be

established. The mean actual score on the exams evaluated by the survey was  $71.44\pm22.63$ . Questions targeted the students' ability to assess their own performance, their patterns of preparation, sleep patterns, anxiety, external demands on their time, prior performance in chemistry, and attitude towards the class (Figure 1). Of the 20 questions on the assessment (Table I), 14 questions asked students to evaluate items on a scale from 1 (never) to 7 (always). The other 6 questions required insertion of a value according to their regular routine.



**Figure 1.** Summary of items and their contribution to actual performance in an Organic Chemistry course.

Because of the large number of variables examined in the study, the data were further examined by distilling the 19 questions (excluding predicted score) to factors by performing a factor analysis in the software R (v. 3.1.1, R development core team) using the call "factanal," which standardizes variables to have a mean of zero and a variance of 1. This statistical approach capitalizes on relationships among questions, which were highly correlated, to develop a reduced number of factors that represent combinations of questions with correlated responses. Factor analysis does not consider performance outcomes in the development of factors. Unfortunately, the analysis required the removal of student surveys with missing responses, thus the sample size for the factor analysis was reduced to 146. After factors were developed, the first four factors were statistically significant and following factors were not. We used these first four factors as predictors of student exam scores using multiple linear regression in the software R (v. 3.1.1, R development core team) and the call "lm."

Tab	Table I. List of question presented to students by the survey		
inst	rument		
1. I	predict I received points out of 100 points possible on		
	this exam. Actual score		
	In the last 2 weeks how often did you:		
2	Read the material planned for the following lecture?		
3	Complete the assigned homework problems?		
4	Complete practice quizzes /exams?		
5	Highlight and summarize the book and/or notes?		
6	Make and review flash cards?		
7	Participate in a study group?		
8	Attend supplemental instruction sessions?		
9	Verbally explain principles to another individual?		
10	Write assessment questions of your own?		
11	Receive 6 or more hours of sleep?		
12	Go to bed & arise at the same time (M-Th)?		
13	Maintain the same sleep pattern on the weekend?		
14	Your level of anxiety on this exam.		
15	Do you like Organic Chemistry?		
16	How many hours of sleep did you receive last night?		
17	In the last 2 weeks about how many hours/day did you study for		
	this class?		
18	In the last 2 weeks how many hours/week committed elsewhere		
	(work, church, etc)?		
19	In the last 2 weeks how many days/week did you meet with a		
	tutor?		
20	What was your letter grade in Chem 1220 (Gen. Chem.)?		

Copy of the survey accounting for the data shown above:

# **Results and Discussion**

#### Individual Item Analysis

Study habits, sleep, tutoring, and the other variables studied statistically significantly predict the actual score on the exam, F(19,142)=11.317, p<0.000. A multiple regression analysis was run to determine which variables could be used to predict the students' exam score. The only variables that contributed significantly to the prediction of a student's actual exam score (p<.05) were participation in a study group, exam anxiety, liking Organic Chemistry, and the student's previous grade in Chem 1220 (Principles of Chemistry II) (see Tables II and III) F(19,142)=11.317, p<0.000, adj.  $R^2=0.549$ . The statistical assumptions of linearity (through a collinearity diagnostics analysis in

SPSS), independence of residuals (assessed by a Durbin-Watson statistic of 2.177), homoscedasticity (assessed by visual inspection of a plot of studentized residuals versus unstandardized predicted values), unusual points (no outliers identified via casewide diagnostics in SPSS), and normality of residuals (based on a visual analysis of histogram of the actual score data, confirmed by a normal P-P plot of regression standardized residual analysis in SPSS) were met. Overall, there was a strong positive correlation between student's actual score and their predicted score ( $R^2=0.847$ , p<0.000). The student's predicted exam score was not included in the multiple regression analysis as the goal was to determine which study habits and other factors such as sleep could best predict a student's grade on the exam. Furthermore, comparison between students who scored 80 or above and those who scored below 80 on their exams showed a statistically significant difference in their ability to predict their exam score, t(122)=10.084, p<0.000, thus verifying that students who performed well were better able to accurately assess their preparedness and performance than less prepared students.

Table II. Arrangement of items (correlation coefficient,significance) as compared with actual performance listed inorder of increasing student contribution			
Item	Pearson's r	Sig.	
Read text	0.053	0.465	
Tutor	0.026	0.026	
Completed homework	0.130	0.074	
Hours studying	0.149	0.042	
Highlight	0.262	0.000	
Flashcards	0.202	0.005	
Create own questions	0.216	0.003	
Supplement instruction	0.331	0.000	
Study group	0.460	0.000	
Practice quizzes/exams	0.425	0.000	
Verbally explain	0.511	0.000	

Correlation coefficient interpretation: small (0.1-0.3), moderate (0.3-0.5), strong (>0.5) correlation.

A Pearson's r correlation was also run to examine the correlation between individual items and the student's actual score on the exam. Listing items in order of increasing correlation (Pearson's r) to actual performance reveals a close parallel to the amount of student activity required in the item (Table II). Items listed at the top of the table are mainly student reception or passive learning items, where the student is receiving information from an external source. Interestingly, when a student is performing poorly in a class and seeks advice on how to improve, the items listed at the top of the table, such as reading the text, completing homework assignments, increase time spent studying, and meeting with a tutor are those generally recommended by most teachers.

Although student learning must begin with information input or reception, not all input methods are equal. Students can receive information by reading the textbook and listening to lectures, but more effective methods are based on student discovery, such as processoriented guided inquiry learning (POGIL).<sup>44,45</sup> Each student must have an appetite and thirst for knowledge for learning to occur. As shown in Table II, input methods alone are inadequate preparation for performance on exams, but they form a starting point on which to build. The items are arranged in increasing requirement of student output as items descend the column. The production (output) of information requires students to connect with their chemical foundation<sup>46</sup> and helps them better assess their knowledge of chemical principles. Verbally explaining principles, writing explanation of principles as explained by Wilson,<sup>47</sup> or creating new assessment questions all require HOCs.

Compared with other high-contribution methods, creating original assessment questions seems to be underrepresented with a lower than expected value. One possible explanation for this low ranking may have been that few students actually participated in this study method since it was not a practice generally recommended in preparation for exams at the time.

Another discrepancy appears to be the difference in placement of completed homework versus completed practice quizzes and exams. The data from this study showed that students who worked through former quizzes and exams, made available online, performed significantly higher on their actual exams and were better able to predict their actual score (correlation coefficient, significance were 0.440, 0.000, respectively) compared with those who reported to have completed the homework in the text (correlation coefficient, significance were 0.216, 0.005, respectively). Interestingly, completion of the assigned homework from the textbook had only a mild positive correlation with student actual performance. The level of difficulty and style of questions may account for this observed difference. Whereas textbook homework problems are directly related to one section or one chapter, questions on exams stem from two or three chapters, requiring greater ability to discern one reaction from another reaction and are generally more challenging than questions found in the text, thus requiring a higher level of assessment and greater connectivity to a chemical foundation.

As shown in Table III, another significant predictor of student actual performance was their grade from the 2<sup>nd</sup> semester of General Chemistry, and this also correlated well with student ability to predict their actual score, coinciding with the study of Lopez et al., which indicated the greatest predictor of student performance was being able to link prior concepts to new concepts through concept maps.<sup>48,49</sup> This correlation would be valuable in advising students who performed poorly (below a C grade) in General Chemistry and are deciding whether to move ahead into Organic Chemistry or to re-take General Chemistry. At Weber State University students must average a C grade in the courses required by their major; however, students are able to advance from Principles Chemistry II to Organic Chemistry I as long as they score a D- grade or better. If a student has aspirations of attending a health-related professional school, they generally need a B grade or better. For students who have made the decision to move on to Organic Chemistry, a foundation in General Chemistry is not an issue that can be immediately improved upon. Although a correlation was found between the students' Principles Chemistry II grade and their actual performance in Organic Chemistry, it is unclear if there is causation. There is some slight overlap in the material taught between the two courses, but for the most part, they are very different subjects. Organic Chemistry is more qualitative in comprehending structure and predicting structure reactivity, while General Chemistry is more quantitative in solving mathematical equations. Occasionally, students who have done poorly in General Chemistry have performed well in Organic Chemistry and vice versa. Also, a few students have managed to take both courses simultaneously (first time for both classes) and scored well in both classes

Table III. Non-study skill items correlation to student actual			
performance			
Item	Corr.	Sig.	
	Co.	)	
General Chemistry II Grade	0.504	0.000	
Like Organic Chemistry	0.38	0.000	
Received >6 h of sleep the night before	0.214	0.003	
exam			
Hours committed elsewhere	-0.228	0.003	
Anxiety	-0.301	0.000	

Student attitude (like or dislike) toward the course also had a positive significant correlation with actual performance and ability to predict actual score. It is difficult to comprehend a subject when emotional resistance toward that subject exists. Also, it is likely that students who have performed well in the class will have a better attitude toward the course, creating a reinforcing cycle leading toward success.

Related to student attitude toward the course is the level of anxiety experienced during a quiz or an exam. Similar to previous studies,<sup>14,15</sup> the level of anxiety experienced during the exam showed a significant negative correlation with actual performance and ability to predict performance.

In contrast to many studies showing strong positive correlations between academic performance and sleep patterns,<sup>6–13</sup> this study showed only a modest positive correlation between students who received at least 6 hours of sleep the night before the exam and their actual performance and ability to predict their actual score.

In spite of frequent reports from past students who failed to complete the course because of overloading their schedules with too many credits hours or work hours, this study showed only a small negative correlation between hours committed elsewhere and their actual performance.

# Group Item Analysis

The questions asked in the survey could be interpreted as four factors cumulatively representing 39% of the variance in the multivariate data set. A test of the null hypothesis that four factors was sufficient  $(\chi^2 = 118.9 \text{ on } 101 \text{ df}, p = 0.108)$  led us to focus on these rather than considering more factors that explained a decreasing amount of the variance. These factors can be interpreted as sleep habits (sleep), assessment and student contribution (active learning), reading, review, and doing homework-low student contribution (passive learning), and foundation in general chemistry, attitude, and anxiety (foundation). These factors suggest that the responses to these questions are related to each other; for example, students with strong foundations in General Chemistry tended to have positive attitudes and low test anxiety. Similarly, students who engaged in one active learning approach also engaged in other active learning approaches and supplemental instruction. Hours committed elsewhere (question 18) was negatively related to all factors, suggesting that more hours committed elsewhere negatively influences the other positive behaviors. By considering correlations of question responses, factor analysis gave a different interpretation than examination of individual items by showing which student attributes and behaviors were related to each other.

Linear regression suggested that all factors contributed to actual performance (Table IV). The strongest contributor was foundation,

suggesting that foundation in General Chemistry, positive attitude, and low test anxiety were most important in determining success. However, factors associated with active learning and supplemental instruction, had nearly as strong an effect on actual scores as did foundation. Sleep and passive learning were weaker predictors of actual scores, but both had significant positive effects. Question 17, about how many hours students were committed to outside of the class, was quite unique from the other questions. In adding this question to the regression in Table 3, it was found to have a small statistically insignificant positive effect on test scores (*est.*=0.69, *t*=1.78, *p*=0.77).

Table IV. Regression results predicting final score from factors				
Factor	Est.	St error	Т	p-value
(intercept)	72.78	1.35	53.90	< 0.001
Sleep	3.988	1.48	2.688	0.008
Active Learning	10.29	1.50	6.907	< 0.001
Passive Learning	3.261	1.56	2.096	0.038
Foundation	13.05	1.77	7.389	< 0.001

slope estimate (est.), standard error of the slope estimate (St. error), t-statistics (T stat.), and p-value

# Application

After receiving the results of this study, instruction in the course was modified in several ways to support increased opportunities for active learning: a portion of class time has been devoted to students explaining principles to one another, review sessions have involved the use of the Air Sketch app to allow students the opportunity to produce answers on a practice exam rather than being passively presented with answers by the instructor, and as students have sought assistance during office hours, they have been given the opportunity to teach principles while being guided with occasional prompts. The result of these course modifications has been mixed. The Spring 2016 Organic Chemistry II class performed significantly better than the Spring 2015 class in regard to scoring a C grade or better (Table V). Furthermore, traditionally the attrition rate has been especially high in the accelerated Summer Organic Chemistry II course, where 14 weeks of instruction is condensed to 7 weeks of instruction. However, this past summer semester (Summer 2016), 66.7% of students passed with a C grade or better compared with only 34.8% of students the prior year. However, the academic

performance of students in the Spring 2016 Organic Chemistry I course was significantly lower than the 5-year average percent of students receiving a C grade or better. One possible explanation for this lower value may be that this particular class was not as responsive to group participation compared with the other classes with improved scores. The authors are encouraged by results obtained thus far, but realize six classes represents a small sample set and assessment over the next couple of years will be needed to provide statistically meaningful results.

Table V. Percent of students receiving a C grade or better over			
the past year for Organic Chemistry I and II courses that were			
Semester O. Chem. $I \ge C$ (# O. Chem. $II \ge C$ (#			
	stud)	stud)	
Fall 2015	60.7% (61)	Not taught	
Spring 2015	Not taught	69.8% (43)	
Spring 2016	50.0% (58)	75.8% (33)	
Summer	Not taught	34.8% (23)	
2015			
Summer	Not taught	66.7% (18)	
2016			

#### Conclusions

Past experience, as commonly reported anecdotally by students who failed to pass the class, has been that they had overloaded their schedules with too many classes or too many hours at work each week. However, this study found that neither the number of hours spent studying nor the duration of commitments outside of class were strong contributors to whether a student performed well on an exam. Traditional study methods (attending class, taking notes, and doing assigned homework) alone were insufficient to adequately prepare students for examinations.

Students with past success in Chemistry tended to continue to have success in Organic Chemistry. Since the value for active learning methods was nearly as large as the value for foundation in the group item analysis, deficiencies in chemical foundation can largely be compensated for by participation in active learning study skill items requiring significant student contribution, such as verbally explaining principles, writing assessment questions, working through practice exams and quizzes, and participating in meaningful study groups. We have provided evidence supporting the need for and effectiveness of learning methods that are student centered and that actively engage them in expressing their knowledge.

# Acknowledgments

We express our appreciation for the students who chose to give of their time in participation of this study.

# References

- Pribyl JR, Bodner GM. Spatial ability and its role in organic chemistry: A study of four organic courses. *J Res Sci Teach*. 1987;24(3):229-240. doi:10.1002/tea.3660240304.
- Pickering M. Further studies on concept learning versus problem solving. *J Chem Educ.* 1990;67(3):254. doi:10.1021/ed067p254.
- Tsaparlis G, Zoller U, Fastow M, Lubezky A. Students' Self-Assessment in Chemistry Examinations Requiring Higher- and Lower-Order Cognitive Skills. *J Chem Educ*. 1999;76(1):112. doi:10.1021/ed076p112.
- 4. Grove NP, Bretz SL. Perry's Scheme of Intellectual and Epistemological Development as a framework for describing student difficulties in learning organic chemistry. *Chem Educ Res Pr.* 2010;11(3):207-211. doi:10.1039/C005469K.
- Wiediger SD, Hutchinson JS. The Significance of Accurate Student Self-Assessment in Understanding of Chemical Concepts. *J Chem Educ*. 2002;79(1):120. doi:10.1021/ed079p120.
- Trockel MT, Barnes MD, Egget DL. Health-Related Variables and Academic Performance Among First-Year College Students: Implications for Sleep and Other Behaviors. *J Am Coll Heal*. 2000;49:125-131.

- Link SC, Ancoli-Israel S. Sleep and the teenager. *Sleep Res.* 1995;24a:184.
- Lowry M, Dean K, Manders K. The link between sleep quantity and academic performance for the college student. *Sentience*. 2010;3:16-19.
- Curcio G, Ferrara M, Gennaro L. Sleep loss, learning capacity and academic performance. *Sleep Med Rev.* 2006;10(5):323-337. doi:10.1016/j.smrv.2005.11.001.
- Thomas M, Sing H, Belenky G, et al. Neural basis of alertness and cognitive performance impairments during sleepiness. I. Effects of 24 h of sleep deprivation on waking human regional brain activity. *J Sleep Res.* 2000;9(4):335-352. doi:10.1046/j.1365-2869.2000.00225.x.
- Stickgold R, James L, Hobson JA. Visual Discrimination learning requires sleep after training. *Nat Neurosci*. 2000;3:1237-1238.
- Karni A, Tanne D, Rubenstein B, Askenasy J, Sagi D. Dependence on REM sleep of overnight improvement of a perceptual skill. *Science (80- )*. 1994;265(5172):679-682. doi:10.1126/science.8036518.
- Pilcher JJ, Walters AS. How Sleep Deprivation Affects Psychological Variables Related to College Students' Cognitive Performance. *J Am Coll Heal*. 1997;46(3):121-126. doi:10.1080/07448489709595597.
- 14. Wine J. Test anxiety and direction of attention. *Psychol Bull*. 1971;76(2):92-104. doi:10.1037/h0031332.
- Culler RE, Holahan CJ. Test anxiety and academic performance: The effects of study-related behaviors. *J Educ Psychol.* 1980;72(1):16-20. doi:10.1037/0022-0663.72.1.16.

- Paulson DR. Active Learning and Cooperative Learning in the Organic Chemistry Lecture Class. J Chem Educ. 1999;76(8):1136. doi:10.1021/ed076p1136.
- Bowen CW. A Quantitative Literature Review of Cooperative Learning Effects on High School and College Chemistry Achievement. J Chem Educ. 2000;77(1):116. doi:10.1021/ed077p116.
- Hagen JP. Cooperative Learning in Organic II. Increased Retention on a Commuter Campus. J Chem Educ. 2000;77(11):1441. doi:10.1021/ed077p1441.
- Dougherty RC. Grade/Performance Contracts, Enhanced Communication, Cooperative Learning and Student Performance in Undergraduate Organic Chemistry. *J Chem Educ.* 1997;74(6):722. doi:10.1021/ed074p722.
- Carpenter S, McMillan T. Incorporation of a Cooperative Learning Technique in Organic Chemistry. *J Chem Educ*. 2003;80(3):330. doi:10.1021/ed080p330.
- Poole MJ, Glaser RE. Organic Chemistry Online: Building Collaborative Learning Communities through Electronic Communication Tools. *J Chem Educ*. 1999;76(5):699. doi:10.1021/ed076p699.
- Herreid CF. Why Isn't Cooperative Learning Used to Teach Science? 1998;48(7):553-559. http://www.jstor.org/stable/1313317.
- Dinan FJ, Frydrychowski VA. A Team Learning Method for Organic Chemistry. J Chem Educ. 1995;72(5):429. doi:10.1021/ed072p429.
- Lyle KS, Robinson WR. A Statistical Evaluation: Peer-led Team Learning in an Organic Chemistry Course. *J Chem Educ*. 2003;80(2):132. doi:10.1021/ed080p132.

- Ryan BJ. Line up, line up: using technology to align and enhance peer learning and assessment in a student centred foundation organic chemistry module. *Chem Educ Res Pr.* 2013;14(3):229-238. doi:10.1039/C3RP20178C.
- Wamser CC. Peer-Led Team Learning in Organic Chemistry: Effects on Student Performance, Success, and Persistence in the Course. *J Chem Educ*. 2006;83(10):1562. doi:10.1021/ed083p1562.
- Bradley AZ, Ulrich SM, Jones M, Jones SM. Teaching the Sophomore Organic Course without a Lecture. Are You Crazy? *J Chem Educ*. 2002;79(4):514. doi:10.1021/ed079p514.
- Amick AW, Cross N. An Almost Paperless Organic Chemistry Course with the Use of iPads. J Chem Educ. 2014;91(5):753-756. doi:10.1021/ed400245h.
- Lewis MS, Zhao J, Montclare JK. Development and Implementation of High School Chemistry Modules Using Touch-Screen Technologies. *J Chem Educ*. 2012;89(8):1012-1018. doi:10.1021/ed200484n.
- Libman D, Huang L. Chemistry on the Go: Review of Chemistry Apps on Smartphones. J Chem Educ. 2013;90(3):320-325. doi:10.1021/ed300329e.
- Williams AJ, Pence HE. Smart Phones, a Powerful Tool in the Chemistry Classroom. J Chem Educ. 2011;88(6):683-686. doi:10.1021/ed200029p.
- Silverberg LJ, Tierney J, Bodek MJ. Use of Doceri software for iPad in online delivery of chemistry content. *J Chem Educ*. 2014. doi:10.1021/ed4009057.
- Silverberg LJ. Use of Doceri Software for iPad in Polycom and Resident Instruction Chemistry Classes. *J Chem Educ*. 2013;90(8):1087-1089. doi:10.1021/ed400093r.

- Li Q, Chen Z, Yan Z, Wang C, Chen Z. Touch NMR: An NMR Data Processing Application for the iPad. *J Chem Educ*. 2014;91(11):2002-2004. doi:10.1021/ed5002784.
- 35. Karatjas AG. Use of iSpartan in teaching organic spectroscopy. *J Chem Educ*. 2014. doi:10.1021/ed4006858.
- 36. Lee AWM, Ng JKY, Wong EYW, Tan A, Lau AKY, Lai SFY. Lecture Rule No. 1: Cell Phones ON, Please! A Low-Cost Personal Response System for Learning and Teaching. *J Chem Educ*. 2013;90(3):388-389. doi:10.1021/ed200562f.
- Wijtmans M, van Rens L, van Muijlwijk-Koezen JE. Activating Students' Interest and Participation in Lectures and Practical Courses Using Their Electronic Devices. *J Chem Educ.* 2014;91(11):1830-1837. doi:10.1021/ed500148r.
- Bryfczynski SP, Brown R, Hester J, et al. uRespond: iPad as Interactive, Personal Response System. 91(3):357-363. doi:10.1021/ed4006453.
- Berrett D. How "Flipping" the Classroom Can Improve the Traditional Lecture. *Chron High Educ*. 2012:1-15. doi:http://dx.doi.org/10.1108/17506200710779521.
- Browne LM, Blackburn E V. Teaching Introductory Organic Chemistry: A Problem-Solving and Collaborative-Learning Approach. J Chem Educ. 1999;76(8):1104. doi:10.1021/ed076p1104.
- Meyer KA. Face-To-Face Versus Threaded Discussions: The Role of Time and Higher-Order Thinking. *JALN*. 2003;7(3):55-65.
- 42. Penn JH, Nedeff VM, Gozdzik G. Organic Chemistry and the Internet: A Web-Based Approach to Homework and Testing Using the WE\_LEARN System. *J Chem Educ*. 2000;77(2):227. doi:10.1021/ed077p227.
- Bunce D, Van den Plas J, Havanki L. Comparing the effectiveness of student achievement of student response system vs. online WebCT quizzes. *J Chem Educ*. 2006;83(3):488-493.
- 44. Hein SM. Positive impacts using POGIL in organic chemistry. *J Chem Educ*. 2012. doi:10.1021/ed100217v.
- Straumanis A, Simons EA. A Multi-Institutional Assessment of the Use of POGIL in Organic Chemistry. In: Moog RS, Spencer JN, eds. *Process Oriented Guided Inquiry Learning* (*POGIL*). Vol 994. Washington, DC: American Chemical Society; 2008:226-239.
- Taagepera M, Noori S. Mapping students' thinking patterns in learning organic chemistry by the use of knowledge space theory. *J Chem Educ*. 2000;77(9):1224-1229. doi:10.1080/0950069970190303.
- 47. Wilson JW. Writing to Learn in an Organic Chemistry Course. *J Chem Educ.* 1994;71(12):1019. doi:10.1021/ed071p1019.
- Szu E, Nandagopal K, Shavelson RJ, et al. Understanding Academic Performance in Organic Chemistry. *J Chem Educ*. 2011;88(9):1238-1242. doi:10.1021/ed900067m.
- Lopez EJ, Shavelson RJ, Nandagopal K, Szu E, Penn J. Factors Contributing to Problem-Solving Performance in First-Semester Organic Chemistry. *J Chem Educ*. 2014;91(7):976-981. doi:10.1021/ed400696c.

# Fabrication of Dye-Sensitized Solar Cells Using Native and Non-Native Nanocrystals in Ferritin as the Dye

Alessandro Perego, Cameron Olsen, J. Ryan Peterson, Kameron Hansen, John S. Colton, Richard K. Watt

Brigham Young University

## Abstract

Recent studies have demonstrated the potential of different nanocrystals synthesized in ferritin to act as light harvesters for high-efficiency solar cells. This study investigates the possibility of using these nanocrystals as alternative dyes for dye-sensitized solar cells. Four different dye-sensitized solar cells were successfully fabricated using iron, cobalt, manganese, and lead-sulfide nanocrystals within ferritin as the dye. The efficiencies obtained were 0.06%, 0.14%, 0.25%, and 0.26%, respectively. These results show that ferritin minerals can indeed act as solar-harvesting substrates and helped us understand how this technology would need to be investigated to make them more practical and efficient.

# **1** Introduction

## 1.1 Dye-sensitized Solar Cells

We live in a world that is moving towards renewable and clean energies. Solar energy is frequently lauded as a potential game changer in the energy landscape, but unfortunately, commercially available photovoltaic technologies are based on inorganic materials (mainly silicon), which require high costs and high-energy-consuming preparation methods.<sup>1</sup>

Dye-sensitized solar cells (DSSCs) present a valuable and sustainable alternative to silicon solar cells. These cells present numerous advantages compared with inorganic photovoltaic systems, such as ability to absorb more sunlight per surface area than standard siliconbased solar panels. DSSCs are also able to work even in low-light conditions such as indirect sunlight and cloudy skies. Finally, they are economical, easy to manufacture, and constructed from abundant and stable resource materials. This makes DSSCs an attractive replacement for current photovoltaic technology.<sup>2</sup>

The mechanism of a dye-sensitized solar cell (DSSC) is shown in Fig. 1. In this cell, the sunlight is absorbed by a dye, which is known as the sensitizer. This dye is adsorbed onto a semiconductor material such



**FIG. 1.** Principle of operation for a dye-sensitized solar cell. The photoexcited dye (small black circles) transfers an electron to the semiconducting  $TiO_2$  layer (larger white circles) via electron injection. The electron is then transported through the porous  $TiO_2$  layer and collected by the conductive layer on the glass. Within the electrolyte, the mediator  $(\Gamma/\Gamma_3)$  undergoes oxidation (and regeneration). The electrons lost by the dye to the  $TiO_2$  are replaced by the electrolyte, which in turn obtains an electron at the catalyst-coated counter electrode (the cathode) as current flows through the electrical load.

as TiO<sub>2</sub>, which not only acts as support for the dye sensitizer but also functions as an electron acceptor and an electronic conductor. The film thickness of the TiO<sub>2</sub> substrate is usually around 10  $\mu$ m and the nanoparticles are about 25 nm in diameter.<sup>2</sup> This film is deposited on a glass plate covered with a transparent conducting oxide (e.g., Indium Tin Oxide or Fluorine-Doped Tin Oxide), which allows light to enter in the cell. At the dye interface, the excited electron created in the dye upon light absorption leaves behind a hole (absence of an electron) in the state from which it was excited, and the electron and hole separate. The excited electron navigates an electric circuit, doing useful work. The circuit is then completed thanks to the presence of an electrolyte (typically a solution of  $\Gamma/I_3$ ), which provides a transport route for the positively charged hole to be filled by a returning electron.

DSSCs have been studied for the past 20 years because of their simplicity in fabrication, but little progress has been made concerning their energy efficiency conversion (best reported efficiency is  $11.9 \pm 0.4\%$ ).<sup>3</sup> Commercially available dyes usually are made using different ruthenium complexes; however, those types of sensitizers are very expensive to manufacture. Organic dyes have been studied in recent years as an alternative to ruthenium-based dyes, but their efficiency is too low to replace ruthenium-based sensitizers. Finding ways to improve the light harvesting and electron transfer reactions of the dye is the key to improving the energy conversion of DSSCs.<sup>4</sup>

## 1.2 Ferritin

Ferritin (FTN) is a 12-nm-diameter spherical protein with an 8-nm hollow interior, which naturally contains iron oxide nanocrystals in the form of ferrihvdrite (Fe(O)OH). The natural core of ferritin can be removed, and other metal oxide nanoparticles can be synthesized inside the empty ferritin. As shown in Fig. 2, the choice of metal used in the growth of the nanoparticles determines the wavelengths of light that can be absorbed.<sup>5</sup> Recent studies have been published to quantify and characterize the light-harvesting and oxidative charge separation reactions of ferritin in order to tune the wavelength of light that is harvested by each nanocrystal. Calculations have also shown that the theoretical maximum efficiency of a solar device made using a combination of (Fe(O)OH), (Co(O)OH), (Mn(O)OH), and (Ti(O)OH) nanocrystals synthesized inside ferritin is up to 38% solar energy conversion.<sup>6</sup> Additionally, ferritin possesses the ability to prevent photocorrosion in metal oxide semiconductors and it is also thermostable up to 80°C. These unique properties make the ferritin nanoarchitecture an in-triguing photocatalyst for DSSCs.5



**FIG. 2.** Left: Cross-section of the ferritin protein 24 subunits assembled to form the hollow spherical protein with a 12-nm-diameter exterior, an 8-nm-diameter interior cavity, and an encapsulated mineral. Right: Exemplification of the ability of ferritin to absorb different wavelengths of light according to the choice of metal used in the growth of the nanoparticles.

# 2 Experimental

## 2.1 Materials

Titanium dioxide electrodes  $(20 \times 20 \times 2 \text{ mm})$  were purchased from Solaronix, made from TCO22-7 FTO-coated glass using an active layer from Ti-Nanoxide T/SP covered by a reflective layer of Ti-Nanoxide R/SP (active area  $6 \times 6$  mm). Platinum counter electrodes were also purchased from Solaronix, made from TCO22-7 FTO-coated glass and prepared by screen-printing using Platisol T/SP precursor. A commercial iodide-based electrolyte (Solaronix AN-50) was selected as a redox solution.

# 2.2 Preparation of dye solutions

For this experiment, we used four different ferritin samples, using Fe, Mn, Co, and PbS nanocrystals as ferritin cores. These were created as solutions, which then coated the  $TiO_2$  (see below), and an anthocyanin dye solution was used as a control.

## 2.2.1 Anthocyanin

Anthocyanins (ANTs) are a class of flavonoids responsible for the red and purple shades of many fruits. Because of their abundance and stability, they have been studied in recent years for DSSC applications.<sup>7</sup> A mixture of cyanin 3-glycoside and cyanin 3-rutinoside was isolated

from California blackberries (*Rubus ursinus*). These are the only ANTs present in blackberries. The pigments were extracted using methanol/ acetic acid/water (25:4:21) and then filtered.

#### 2.2.2 Ferritin Cores

Horse spleen ferritin (HoSF) (SigmaAldrich CAS# F4503) was used as supplied for the Fe(O)OH (native nanocrystal) cores. To synthesize Mn(O)OH cores, we followed established methods,<sup>8,9</sup> which involved adding a solution of MnCl<sub>2</sub> to apoferritin using 3-([1,1,-Dimethyl-2-hydroxyethyl]ammino-2-hydroxypropanesulfonic acid as buffer (AMPSO). The color change from clear to brown is a definite indication of the oxidation of the manganese and of the formation of the new core. Co(O)OH cores were also synthesized by following established methods10 by periodically adding Co(NO3)2 and H2O2 to an apoferritin solution. PbS cores were synthesized in 500 mL of a 0.1 M sodium acetate buffer at pH 5.5 and room temperature. The buffer was prepared and 6.25 nmol of apoferritin and 3.125 µmol of Pb(AcO)2 were added and allowed to mix gently for 2 minutes. This was followed by an injection of 3.125 umol of Na2S, which was then allowed to mix for 5 minutes. The supernatant, containing the protein with the PbS cores, was then extracted after centrifuging.11

## 2.3 Preparation of photovoltaic devices

For the ANT-based DSSC, we soaked the TiO<sub>2</sub>-coated glass plates for 30 minutes in a 0.1 mg/mL cyanin extract solution. Adsorption of cyanin to the surface of TiO<sub>2</sub> and complexation to Ti<sup>IV</sup> sites is rapid, forming a complex capable of electron injection.<sup>7</sup> For the ferritin-based DSSCs, the different solutions were diluted down to 0.1 mg/mL using a TRIS HCl buffer at pH 8.5. The electrode was then soaked in the solution for 72 hours. Once ready, we assembled the cell and injected a drop of the electrolyte solution between the two electrodes and proceeded to the characterization.

## 2.4 Measurements

## 2.4.1 Band Gap Measurements

Absorption was measured with a 500 W tungsten halogen lamp. The light passed through a monochromator, with which we selected wavelengths spaced every 4 nm. The light was mechanically chopped before passing through the cuvette with the sample in solution and was finally measured using a silicon photodiode and lockin amplifier. The absorption was calculated by comparing the intensity of light transmitted through the ferritin solution to the intensity through a control cuvette containing only buffer solution. All samples had a concentration of 0.4 mg/mL. The band gaps for all but the PbS sample were measured by extrapolating linear fits of the square root of the absorption coefficient for indirect transitions and the square of the absorption coefficient for direct transitions. The band gap of the PbS sample was measured by photoluminescence. The sample was deposited on a glass slide, and a 398-nm laser was used to stimulate photoluminescence. The wavelength of the emitted light was measured using a spectrometer and germanium photodiode.

## 2.4.2 Photovoltaic Measurements

Photovoltaic measurements were recorded using a variable resistance load (5 to 5000  $\Omega$ ) while the cell was illuminated with a 150 W Xe arc lamp at standard AM 1.5 condition (100 W/cm<sup>2</sup>) at 298 K. (See Fig. 3.<sup>12</sup>)

The efficiency was then calculated by taking the ratio of  $P_{max}$ , to  $P_{in}$ :

$$\eta = \frac{P_{\text{max}}}{P_{\text{fn}}} 100 \tag{1}$$

In this equation,  $P_{max}$  is the maximum electrical power output and  $P_{in}$  is the product of the irradiance of the light [W cm<sup>-2</sup>] with the active surface area of the solar cells [cm<sup>2</sup>]. The curves thus obtained were then fit to the following I-V equation<sup>13</sup>:

$$I = I_{L} - I_{0} \left\{ ex p \left[ \frac{V + IR_{s}}{nV_{T}} \right] - 1 \right\} - \frac{V + IR_{s}}{R_{SH}}$$
<sup>(2)</sup>

where  $I_0$  is the reverse saturation current, *n* is the diode ideality factor (1 for an ideal diode), *q* is the elementary charge of the electron,  $V_t$  is the thermal voltage ( $V_t = kT$ , *k* being Boltzmann's constant and *T* the temperature, which equals 0.0259 V at 298K), and  $I_L$  is the photogenerated current.  $R_S$  and  $R_{SH}$  represent the series and shunt resistances, which are parasitic resistances that tell us how much efficiency we are losing because of the dissipation of power across internal resistances of the cell. Specifically,  $R_s$  is the result of contact resistance and charge transfer resistance in the semiconductor material while  $R_{SH}$  results from photo-generated current taking an alternate current path instead of going through the electrodes and the load as expected. For an ideal cell,  $R_{SH}$  would be infinite and would not provide an alternate path for current to flow, while  $R_S$  would be zero, resulting in no further voltage



**FIG. 3.** Schematic representation of the circuit used to measure the solar cells' parameters. The variable resistance *R* allowed for the *I*-*V* curve to be traced out and ranged from 5 to 5000  $\Omega$ 

drop before the load. Decreasing  $R_{SH}$  and increasing  $R_s$  will decrease the fill factor (*FF*) and the maximum observable power ( $P_{max}$ ). *FF* is essentially a measure of quality of the solar cell. It is calculated by comparing the maximum power to the theoretical power ( $P_T$ ) that would be output using the product of the open circuit voltage and the short circuit current:

$$F_{F} = \frac{P_{\text{max}}}{P_{\text{T}}} = \frac{I_{\text{MP}} \cdot V_{\text{MP}}}{I_{\text{SC}} \cdot V_{\text{OC}}}$$
(3)

In this equation,  $I_{MP}$  represents the current at  $P_{max}$ ,  $V_{MP}$  the voltage at  $P_{max}$ ,  $I_{sc}$  the short circuit current (I at V=0), and  $V_{oc}$  is the open thewhile increasing  $R_s$  excessively can cause  $I_{SC}$  to drop.

## **3 Results**

The measured direct and indirect bad gaps for Fe, Co, Mn, and PbS nanocrystals within ferritin are listed in Table 1. Fig. 4 shows the ferritin nanoparticle absorptions at different wavelengths.

Table 1. Band gaps measurements of the four ferritin samples						
	Direct Band Gap (eV)	Indirect Band Gap (eV)				
FTN-Fe	$2.95 \pm 0.05$	$1.90 \pm 0.1$				
FTN-Co	$2.76\pm0.05$	$1.38 \pm 0.06$				
FTN-Mn	$2.70 \pm 0.05$	$1.00 \pm 0.1$				
FTN-PbS	1.10	-				

Because of its molecular structure, PbS does not possess an indirect band gap.



**FIG. 4.** Ferritin nanoparticles absorptions on AM 1.5G spectrum. The concentration used was 0.4 mg/mL for all the samples.

The photovoltaic properties of the solar cells constructed from these dyes were measured under exposure to the standard AM 1.5 radiation (100 mW cm<sup>-2</sup>). Figure 5 shows the *I-V* curve of the different cells. The measured open-circuit photovoltage ( $V_{oc}$ ), short-circuit photocurrent density ( $J_{sc}$ ), fill factor (*FF*), and solar-to-electric conversion efficiencies ( $\eta$ ) are listed in Table 2.



**FIG. 5.** Current-voltage curve obtained with our dye-sensitized solar cells under AM 1.5 radiation (100 mW cm<sup>-2</sup>) using a variable load (5 to 5000  $\Omega$ ).

Table 2. Photovoltaic performance on DSSCs based on thedifferent dyes								
Device	Dye	V <sub>oc</sub> (V)	$J_{sc}$ (mA cm <sup>-2</sup> )	FF (%)	η (%)			
1	ANT	0.41	0.95	73.2	0.31			
2	FTN	0.21	0.25	59.0	0.06			
3	FTN-Co	0.38	0.44	55.1	0.14			
4	FTN-Mn	0.40	0.73	67.5	0.25			
5	FTN-PbS	0.37	0.90	55.8	0.26			

## 4 Discussion

The low efficiency obtained can be explained by analyzing the structure of ferritin. Studies have shown that electron transfer does occur in ferritin,<sup>14</sup> but contrary to most of the available dyes used for DSSCs, the protein lacks a highly conjugated system. This factor greatly limits the electron transfer between the protein and the TiO<sub>2</sub> substrate, and we see from the results that it decreases the  $J_{sc}$  and subsequently the efficiency of the cells.

Another important consideration to make comes when we analyze the binding between the protein and the titanium dioxide substrate. In DSSCs, the dye molecules usually possess several =O or –OH groups capable of chelating the  $Ti^{iv}$  sites on the titanium dioxide. The kinetics of this process are usually very quick; in fact, it takes only a few minutes for the ANT to bind on the substrate.<sup>7</sup> In contrast, ferritin does not possess several =O or –OH groups on the surface, and therefore the binding is not as strong.

Our experiments also show that if the soaking process is too long and too much ferritin gets deposited, the series resistance ( $R_S$ ) notably increases and therefore the efficiency of the cell decreases. This is likely happening because of a buildup of high contact resistance between the protein and the semiconductor material. Our study suggests that anywhere between 48 and 72 hours is an optimal deposition time. One possible way to overcome these two challenges could be through using a genetically modified ferritin in which a highly conjugated peptide chain is added to the outer ferritin shell. This peptide chain also has the potential to be engineered to increase its affinity for binding with titanium dioxide.<sup>13</sup>

Our results show that the best efficiencies were obtained by using lead sulfide and manganese cores. This can be explained by analyzing the band gap measurements and the absorptions of the different nanocrystals synthesized inside ferritin (see Fig. 4). The measured direct band gap of FTN-PbS is right at 1.1 eV, whereas the band gap of

FTN-Mn is 2.7 eV, with an indirect band gap of 1.0 eV. These values according to the Shockley–Queisser  $limit^{16}$  will provide the highest efficiency for a single-layer photovoltaic solar cell when compared with the band gap measurements of the other ferritin nanocrystals structures.

# **5** Conclusion

We successfully investigated the possibility of using iron, cobalt, manganese, and lead-sulfide nanocrystals within ferritin as alternative dyes for DSSCs. Our results show that lead-sulfide and manganese nanocrystals best match the necessary properties for an alternative dye.

To increase the power output of the cells, it is necessary to optimize and implement the deposition of ferritin on the titanium dioxide layer. Experimental analysis and computational simulation will be performed to better understand the physical and chemical interactions between the protein and the semiconductor material. After the optimization of a single solar cell, the next step will be to fabricate a photovoltaic module of multiple cells connected to each other. This module, called a multi-junction solar cell, will use the principle mentioned above in which different nanocrystals (namely iron, manganese, cobalt, and lead sulfide) synthesized inside ferritin can absorb different wavelengths of the light This will allow us to harvest more frequencies of light simultaneously and therefore raise the efficiency of the cell.

# References

- Mohr, N., Meijer, A., Huijbregts, M. A., & Reijnders, L. (2009). Environmental impact of thin-film GaInP/GaAs and multicrystalline silicon solar modules produced with solar electricity. Int J Life Cycle Assess, 14(3), 225-235. doi: 10.1007/s11367-009-0062-z
- Jasim, K. E. (2011). Dye sensitized solar cells—working principles, challenges and opportunities. In *Solar Cells—Dye-Sensitized Devices* (ed. L. A. Kosyachenko). InTech Open, Rijeka, Croatia. doi:10.5772/19749
- Green, M., Emery, K., & Hishikawa, Y. (2015). Solar cell efficiency tables (Version45). Prog. Photovolt. Res. Appl., 23(1), 1-9.

- Kim, Y., & Lee, D. (2011). Development of dye-sensitized solar cell for high conversion efficiency. In *Solar Cells—Dye-Sensitized Devices* (ed. L. A. Kosyachenko). InTech Open, Rijeka, Croatia. doi:10.5772/19577
- Watt, R. K., Petrucci, O. D., & Smith, T. (2013). Ferritin as a model for developing 3rd generation nano architecture organic/inorganic hybrid photo catalysts for energy conversion. Catal. Sci. Technol. 3(12), 3103. doi:10.1039/ c3cy0053
- Erickson, S. D., Smith, T. J., Moses, L. M., Watt, R. K., & Colton, J. S. (2014). Non-native Co-, Mn-, and Tioxyhydroxide nanocrystals in ferritin for high efficiency solar energy conversion. Nanotechnology, 26(1), 015703. doi:10.1088/0957-4484/26/1/015703
- Cherepy, N. J., Smestad, G. P., Grätzel, M., & Zhang, J. Z. (1997). Ultrafast electron injection: implications for a photoelectrochemical cell utilizing an anthocyanin dyesensitized TiO<sub>2</sub> nanocrystalline electrode. J. Phys. Chem. B, 101(45), 9342-9351. doi:10.1021/jp972197w
- Douglas, T., & Stark, V. (2000). Nanophase cobalt oxyhydroxide mineral synthesized within the protein cage of ferritin. Inorg. Chem. 39, 1828–30
- Tominaga, M., Ohira, A., Yamaguchi, Y., & Kunitake, M. (2004). Electrochemical, AFM and QCM studies on ferritin immobilized onto a self-assembled monolayer-modified gold electrode J. Electroanal. Chem. 566, 323–9
- Yamashita, I., Iwahori, K., and Kumagai, S. (2010). Ferritin in the field of nanodevices Biochim. Biophys. Acta. 1800, 846– 57
- Hennequin, B., Turyanska, L., Ben, T., Beltrán, A. M., Molina, S. I., Li, M., et al. (2008). Aqueous near-infrared fluorescent composites based on apoferritin-encapsulated PbS quantum dots. Adv. Mater. 20(19), 3592-3596. doi:10.1002/ adma.200800530

- Shrotriya, V., Li, G., Yao, Y., Moriarty, T., Emery, K., & Yang, Y. (2006). Accurate measurement and characterization of organic solar cells. Adv. Funct. Mater. 16(15), 2016-2023. doi:10.1002/adfm.200600489
- Hovinen, A. (1994). Fitting of the solar cell IV-curve to the two diode model. Phys. Scr. T54, 175-176. doi:10.1088/0031-8949/1994/t54/043
- 14. Xu, D., Watt, G. D., Harb, J. N., & Davis, R. C. (2005). Electrical conductivity of ferritin proteins by conductive AFM. Nano Lett., 5(4), 571-577. doi:10.1021/nl048218x
- Fukuta, M., Zettsu, N., Yamashita, I., Uraoka, Y., & Watanabe, H. (2013). The adsorption mechanism of titaniumbinding ferritin to amphoteric oxide. Colloids Surf B Biointerfaces, 102, 435-440. doi: 10.1016/j.colsurfb.2012. 07.042
- Shockley, W., & Queisser, H. J. (1961). Detailed balance limit of efficiency of p-n junction solar cells. Semiconductor Devices: J. Appl. Phys. 32 510–9

# A New Perspective to Antenna Design for Reconfigurable Wireless Networks

Mehedi Hasan, Israfil Bahceci, and Bedri A. Cetiner Utah State University

## Abstract

In recent years, wireless communication technologies have undergone a period of unprecedented growth, leading to dramatic increases in data traffic. Network concepts, like internet of things, build on cloud computing and networks of data-gathering sensors; such developments indicate an exponential increase of information exchange. This calls for reconfigurable wireless networks where each layer of the network shall be reconfigurable to account for a diverse set of operational requirements. Subject to these diverse operational requirements, an efficient reconfigurable physical layer shall consist of reconfigurable antennas. A new class of antennas called multifunctional reconfiguable antennas (MRAs) have been developed at Utah State University. Parasitic tuning–based reactive surfaces on and around their driven antenna surface give MRAs the capability to dynamically modify resonant frequencies, polarizations, and radiations patterns. Until now, when an MRA was designed, the considered parameters did not account for the spatial and temporal statistics of multipath propagation environment (e.g., angular spread, path correlations, coherence time, and bandwidth of the underlying channels). Link level analysis shows that, while traditional design parameters are important, the statistics of radio channeling cannot be overlooked. For example, channel correlations increase spectral efficiency and improve the error performance. This observation gives a new perspective to antenna design. Simulation results indicate considering statistics of propagation environment along with traditional parameters while designing antennas improves overall system performance.

## Introduction

Today's wireless communication technologies face significant challenges with handling the ever-increasing volume of data traffic. To meet the demands of increasing data traffic, the network capacity must be increased. Moreover, because of the wide range of diverse application, the configurability of wireless networks must diversify. A multiple-input and multiple-output (MIMO) wireless system that uses multiple transmission and reception antennas offers improved capacity over single-antenna topologies. In an effort to further improve the performance of MIMO systems and to meet the wide range of diverse operational requirements, MIMO systems equipped with multi-functional reconfigurable antennas (MRA) have been proposed [1], [2]. MRA refers to a new class of antenna that is capable of dynamically changing its properties, i.e., radiation patterns, frequencies, and polarizations a.k.a. modes. Combination of multiple individual MRAs in an array forms a multifunctional reconfigurable antenna array (MRAA). MRAA in conjunction with MIMO gives rise to a multifunctional reconfigurable MIMO (MR-MIMO) system, which has shown great potential and sparked considerable interest in the research community.

The radiation pattern of a standard linear array can only be controlled by the array factor, which is determined by the geometrical position of the individual antenna element (with a fixed element factor) forming the array and their excitations. In contrast, MRAA has a variable element factor, which presents a significant additional reconfigurable ability. Traditionally, the design criteria of the basic building block of an MRAA, i.e., MRA, has been similar to a legacy antenna. For example, antenna properties like reflection coefficient, realized gain, axial ratio, efficiency, and bandwidth have played a major role in designing MRAs. While the traditional design parameters play a vital role, statistics of multipath environment can also be a key in determining the overall system performance [3]. Statistics such as angular spread, path correlations, coherence time, and bandwidth of the underlying channels varies in different propagation scenarios. Moreover, various transmission schemes give rise to various desirable properties of the underlying physical layer. Antenna designers must, therefore, consider the channel statistics when designing MRAs. To this end, in an effort to support our claim, we present a modulation scheme based on MRA and how pattern correlation of MRA can be used to improve the performance of this scheme. Simulation results demonstrate that, for a  $4\times2$  MIMO system comprising MRA, decreasing the average pattern correlation from 0.6 to 0.4 can result in ~6-dB improvement in overall system performance.

## Space-Shift Keying and MRA

Space-shift keying (SSK) is a recently developed modulation technique that uses antenna index rather than modulated signal properties to convey information [4]. The idea of SSK is presented in Figure 1.



Figure 1. Space-shift keying for a 4×4 MIMO system.

Traditional modulation schemes use modulated signal properties, i.e., frequency, amplitude, and phase, to carry information. In SSK, it is the antenna index used during transmission that relays information, rather than the transmitted symbol itself. For example, in Figure 1, if antenna-1 is activated during transmission, the receiver will decode the transmitted symbol as 0 or transmitted bits as 00 regardless of the waveform received, given that the receiver can successfully determine the antenna used for transmission. SSK's strength lies its simplicity. It eliminates the necessity of multiple radiofrequency (RF) chains (note that SSK uses only one RF chain), thereby reducing the hardware complexity. While SSK eliminates the need of multiple RF chains, to

achieve a high data rate it requires a large number of transmit antennas. For comparison, to achieve a data rate of 64-QAM, i.e., 6 bits per symbol, a comparable SSK scheme would have to use 64 antennas to correspond to 64 different symbols. Recently, reconfigurable antenna-based SSK have been proposed in literature [5]. Instead of an antenna array, a reconfigurable antenna that can change its properties can be used to carry information over the change of modes. A receiver can determine the active antenna mode at the transmitter using maximum likelihood decoder. An MRA-based SSK eliminates the need of large number of antennas in the transmitter.

# Results

For MRAs, the channel correlations resulting from different antenna modes are related to the radiation pattern correlation between the corresponding antenna modes. As expected, pattern correlation and resulting channel correlation tends to follow similar trends, which can be seen in Figure 2.



Figure 2. Pattern correlation and resulting channel correlation in a richscattering environment

For reconfigurable antenna-based SSK, the performance of a maximum likelihood decoder depends on the separation of constellation point, which in turns depends on the correlation of channels. Channels with higher correlations are likely to result in erroneous detection as they are closely separated in signal space. So, channels with lower correlation/higher decorrelation are expected to have better system performance. Relating this idea and the similarity of pattern and channel correlation, it is obvious that MRA with modes that generate more pattern decorrelation results in improved performance. A comparison of symbol error rate (SER) with different average pairwise pattern correlation is provided is Figure 3. Here we consider an SSK scheme with the transmitter equipped with a 4-mode MRA. Simulation results show that when the average pattern correlation is 0.6, the transmit SNR has to be ~22 dB to achieve an SER of  $10^{-3}$ ; to achieve the same SER, a transmit SNR of 15 dB is enough when the average pattern correlation is 0.4.



**Figure 3.** Symbol error rate vs. SNR for different pairwise pattern correlation (number of receiver = 2, number of transmitter mode = 4, 2 bits/symbol)

# Conclusion

In conclusion, we have shown that decreasing pattern correlation/increasing pattern decorrelation improves MRA-based SSK schemes performance. MRA geometry can be optimized to generate more diverse radiation patterns, which gives an additional dimension to MRA design.

# References

[1] Z. Li, D. Rodrigo, L. Jofre, and B. A. Cetiner, "A new class of antenna array with a reconfigurable element factor," IEEE Transactions on Antennas and Propagation, vol. 61, no. 4, pp. 1947-1955, April 2013.

[2] B.A. Cetiner, E. Akay, E. Sengul, and E. Ayanoglu, "A MIMO system with multifunctional reconfigurable antennas," IEEE Antennas and Wireless Propagation Letters, vol. 5, no. 1, pp. 463-466, Dec. 2006.

[3] M.A. Jensen and J.W. Wallace, "A review of antennas and propagation for MIMO wireless communications," IEEE Transactions on Antennas and Propagation, vol. 52, no. 11, pp. 2810-2824, Nov. 2004.

[4] J. Jeganathan, A. Ghrayeb, L. Szczecinski, and A. Ceron, "Space shift keying modulation for MIMO channels," IEEE Transactions on Wireless Communications, vol. 8, no. 7, pp. 3692–3703, July 2009.

[5] Z. Bouida, H. El-Sallabi, A. Ghrayeb, and K. Qaraqe, "Reconfigurable antenna-based space-shift keying (SSK) for MIMO rician channels," IEEE Transactions on Wireless Communications, vol. 15, no. 1, pp. 446–457, Jan. 2016.

# Throttled Launch-Assist Hybrid Rocket Motor for an Airborne NanoSat Launch Platform

Zachary S. Spurrier, Sean D. Walker, Stephen L. Merkley, Stephen A. Whitmore

Utah State University

## Abstract

We describe design, integration, and testing of a throttled launch-assist hybrid rocket motor for an airborne nano-launch platform. Gaseous oxygen and additively manufactured acrylonitrile-butadiene-styrene are used as the propellants. We established the requirements for this launch-assist propulsion system, developed the system design features, and developed a closed-loop proportional throttle control law. The detailed end-to-end system design is presented. Initial static tests were performed with a cylindrical fuel port to verify system functionality and establish a baseline for the propellant regression rate and optimal oxidizer-to-fuel massflow (O/F) ratio. Subsequent tests are performed using a helical fuel port to increase the volumetric efficiency of the system and allow operation near the optimal O/F condition. Multiple restarts of each system configuration are demonstrated. Results of open- and closed-loop throttle tests are presented.

## I. Introduction

Since the early days of spaceflight, an unachieved goal has been to create an orbital launch system capable of operating from runways with convenience and flexibility similar to aircraft. Mainly because of propulsion technology limitations with chemical rocket engines, nearly all launch systems developed to date perform takeoff vertically from specialized launch pads and have very limited operational flexibility. Fixed-base launches are restricted to certain azimuths and orbit inclinations (depending on launch site), and launch windows are typically short in duration and infrequent in occurrence.

A recent NASA-DARPA<sup>1</sup> study has concluded that there exists a significant potential for horizontal air-launch to provide critical strategic advantages and "assured" access to space when compared with fixed-base launch operations. Because the launch altitude and airspeed are achieved using a high-efficiency air-breathing propulsion system, there is a significant reduction in the required  $\Delta V$  that must be delivered by the launch vehicle, and a significantly smaller launch vehicle is allowed. The study concludes that a performance boost to orbit of 50% may be obtainable. An air-launched vehicle can also achieve a wide range of launch inclinations and right ascensions from a single deployment site. Performance boost for launches at or near the equator can be accomplished with a 12% to 25% reduction in propellant mass. More importantly, air-launch provides a wide range of operational options including on-demand launch azimuth, flexible launch windows, and nearly all-weather launch opportunities. This capability enhancement can lead to increased launch rates and an associated overall launch-cost reduction.

## A. Towed-Glider Air Launch System (TGALS).

The NASA-DARPA study concluded that a towed, remotely piloted, unpowered glider bottom-launching the space-launch vehicle has the potential to be significantly smaller and operationally cheaper than a dedicated human-crewed carrier aircraft. Because the towed platform is separated from the launch vehicle by a significant distance, the risk to human crew is significantly reduced. Consequently, the launch platform does not require certification for human occupancy.

The high lift-to-drag (L/D) towed platform offers the potential for a significantly increased operational range when compared with a coupled launch vehicle and lift platform. Finally, the glider platform can be towed to the launch altitude using a variety of options; this concept offers a significant increase in operational flexibility. Also, these features offer the potential to dramatically lower launch operating



Figure 1. Concept of operations of towed-glider air launch system.

costs. Such cost savings could represent a market-disruptive potential for the merging commercial spaceflight industry. Figure 1 shows the concept of operations (CONOPS) for a TGALS operational platform.

# B. AFRC Demonstration Prototype of TGALS

Previous air-launch studies<sup>2,3,4</sup> have demonstrated that a key parameter for optimal air-launch trajectories is the launch flight path angle. Conceptually, an optimal air-launch flight path angle at the launch altitude and airspeed would place the launch vehicle onto the trajectory that follows the optimal ground launch trajectory. The glider platform itself is unable to achieve this flight condition, and launch-assist propulsion is required. Currently, Armstrong Flight Research Center (AFRC) is developing a prototype platform to verify the operational feasibility of the towed-launch platform concept. A primary objective of this demonstration project is to tow to altitude, release, and safely return to base with an instrumented, subscale, remotely piloted, twinfuselage glider with a representative scaled small-rocket system. Figure 2 shows a photograph of the demonstration vehicle-scaled prototype. The launch vehicle is attached to the center pylon of the launch platform. This demonstration project will allow AFRC to gain operational experience with the towed-glider platform, understand aerodynamic and structural interactions of the rocket and pylon, and demonstrate that the launch platform can achieve the proper launch attitude.



Figure 2. Demonstration prototype of towed-air launch platform.

# C. Top-Level Requirements for Launch-Assist Rocket System

Based on a preliminary analysis performed by NASA AFRC, the top-level system requirements for the launch-assist propulsion system are:

- 1. Maximum thrust of 200 lbf.
- Capability to throttle from <20% to 100%. Simulation studies verified that a high level of system throttleability was necessary to achieve the required flight profiles.
- 3. Provide sufficient throttle fidelity to allow a 2–2.5g pull-up to 70° flight path angle at 85 knots true airspeed (KTAS) at 4500 ft above mean sea level (MSL).
- 4. Provide sufficient impulse to allow the launch platform to hold the 70° flight path angle for a minimum of 5 seconds.
- 5. Use nontoxic, nonexplosive propellants, and a nonpyrotechnic ignition system.
- 6. A properly engineered, restartable launch vehicle "stage 0" trajectory would retain sufficient impulse to allow contingency energy management for the glider launch platform to return to base. Thus, system restartability is highly desirable.

Multiple options are available to achieve the required launchassist total impulse, including a small solid rocket booster, a bipropellant liquid system, a cold-gas system, a mono-propellant hydrazine system, and a hybrid rocket system. The bi-propellant liquid rocket was discarded because of the associated complexity and expense of engineering the required subsystems. The hydrazine system was discarded because of the potential vapor hazard and the associated operational complexities of working with a toxic propellant. The solid rocket booster, although offering a simple solution, does not deliver the impulse precision and variable thrust required to place the launch platform onto the properly launch attitude. Finally, because of the associated low specific impulse  $(I_{sp})$ , the cold-gas system required more propellant than can be carried by the launch platform with the launch vehicle payload. Thus, by process of elimination a hybrid system was selected for the launch-assist propulsion unit (LAPU).

## **II. TGALS LAPU System Design Overview**

Figure 3 presents a top-level solid-model schematic of the LAPU systems. The prototype system is based on a previous design tested at Utah State University (USU).<sup>5</sup> Pictured are the gaseous oxygen (GOX) oxidizer tanks, the high-pressure fill and relief valves, a tank manifold, a manually set pressure-reducing regulator, a low-pressure burst safety disk, an electronic run-valve, a ball-type throttle valve, the electrical valve actuator, and the motor thrust chamber pressure case. The associated pneumatic assembly piping and connectors are also shown. Major features are described in detail in the following subsections.



Figure 3. Top-level schematic of LAPU hybrid motor system components.

## A. Hybrid Motor Combustion Chamber and Ignition System

The hybrid motor system employs GOX as the oxidizing agent and additively manufactured acrylonitrile-butadiene-styrene (ABS) as the fuel component. These propellants are nonexplosive and nontoxic and remain inert until combined within the motor combustion chamber. The fuel grain is manufactured using the conventional fused-deposition modeling (FDM) technique of additive manufacturing for thermoplastics and features "snap-together" interlocks that allow the grain segments to be manufactured separately and then assembled for combustion. The FDM-processed grain segments also provide for an embedded helical fuel port that enhances the fuel burn rate and combustion efficiency.

#### 206 Engineering



Figure 4. Schematic of LAPU hybrid motor with snap-together helical segments.

Figure 4 shows a cut-away schematic for the hybrid rocket motor case. Pictured are the helical fuel grain interlocks, injector cap with ignition electrodes, and post-combustion chamber with graphite nozzle insert and adapter. The motor case is constructed from a modified Cesaroni solid rocket motor case and is 98 mm in diameter and ~60 cm long. The pictured fuel grain is additively manufactured from commercially available Stratasys ABSplus-340 feed-stock (www.stratasys.com/materials/fdm/absplus). Table 1 lists dimensions and weights of the major thrust chamber system components.

Table 1. Thrust Chamber Component Dimensions andWeights							
Motor case	Length: 70.2 cm	Diameter: 98 mm	Empty weight:	Loaded weight:			
			3.61 kg	6.54 kg			
Injector	Diameter: .716 cm	Single port, aluminum. Area: .402 cm <sup>2</sup>	Cd ~ 0.85	Total oxidizer load 11.2 lbm (3.8 kg)			
Machined graphite nozzle	Diameter: 1.85 cm	Expansion ratio: 4.65	Conical exit angle: 15°	Throat erosion rate: 0.011 cm/sec			
ABS fuel grain	Length: 58.61 cm; Diameter: 8.4 cm	Initial port diameter: 2.286 cm	Fuel weight: 2.932 kg	0.5:1 helix ratio, 19.5 cm, pitch length (3 turns)			

The system is ignited using a patent-pending, arc-ignition technology developed at USU.<sup>6</sup> This technology exploits the unique electrical breakdown properties of additively manufactured ABS to allow ondemand start and restart. The non-pyrotechnic system requires two independent signals to initiate combustion and is thus duel redundant to the Hazards of Electromagnetic Radiation to Ordnance (HERO) as defined by MIL-STD-464.<sup>7</sup> The oxidizer injector consists of a single port injector with a .402-cm<sup>2</sup> area to allow the required massflow of at least 250g/sec (0.55 lbm/sec) into the combustion chamber without choking. The ignition power-processing unit (PPU) and oxidizer delivery system are not shown in Figure 4. The ground test motor systems are designed to reproduce the flight systems as closely as possible.

Figure 5 shows the flight system components in the approximate the flight orientation, as mounted to the pylon between the twin vehicle fuselages. The fully loaded system weight is approximately 23.9 kg (52.6 lbm), and it is approximately 165 cm (65 in) in end-to-end length. Each GOX tank is rated for a 4500 psig maximum fill capacity and holds approximately 1.93 kg (4.24 lbm) of oxidizer when filled at room temperature. The motor dry system weight is ~18 kg (40.3 lbm).



Figure 5. Installed LAPU System schematic.

# B. Flight Test Oxidizer Delivery System.

Figure 6 presents the oxidizer delivery system piping and instrumentation diagram (P&ID) for the flight test system. The system is designed to operate between 4500 and 1500 psig upstream of the pressure regulator and between 750 and 800 psig downstream of the regulator. Required safety-of-flight system instrumentation consists of pressure transducers upstream of the regulator and a chamber pressure transducer.





The oxidizer delivery system components consists of

the electrically actuated throttle ball-valve,

- the manual-set pressure-reducing regulator,
- the DC-solenoid actuated run valve,
- the Thrust chamber injector, and two aviation-rated 4500 psig carbon-composite gaseous oxygen storage tanks, manifolded together.

The throttle ball-valve allows the system to regulate the massflow by adjusting the outlet Cv coefficient. A full-open valve Cv range of approximately 2.5 is required to achieve the desired 250g/sec maximum massflow level at a valve inlet pressure of approximately 750 psig. The valve is actuated using an Invensciencei01300 rotary actuator (http://www.invenscience.com/index\_files/torxis\_rotary\_\_\_\_\_servo.htm/). The 12-V powered ball-valve rotary actuator features 0 to 5 VDC analog input proportional control signal.

The pressure regulator has a lockable, manual set-point. Assuming a full-filled capacity for the  $O_2$  tanks (4500 psig) and the assumed ball-valve Cv (2.5), a regulator set-point range of approximately 750 psia will be required to achieve the prescribed maximum thrust level or ~200 lbf. It is assumed that the Cv for the electronic run valve is >2.5 to ensure that the flow will not choke upstream of the throttling ball-valve. The regulator set-point will be manually tuned to adjust for any potential losses in the system run valve. The regulator valve set point of 750 psig was selected to ensure a choking massflow of greater than 250 g/sec at that pressure set-point.

## C. Ignition System PPU and Control System

The ignition system PPU is based on the UltraVolt AA-series line of high-voltage power supplies (HVPS).<sup>8</sup> These HVPS units take a 24– 28 VDC input and provide a current-limited (30 mA) high-voltage output—up to 1 kV. The output signal is initiated by a commanded TTLlevel signal. Units with output capacities from 4 to 30 watts are available. Previous experience with this ignition system has demonstrated that ignition can be achieved using as little as 6 watts;<sup>9</sup> however, to ensure guaranteed reliable motor ignition a 30-watt model will be employed for this design.

Figure 7 shows the interface to the AA-series HVPS. The unit features current and high-voltage output signals used to monitor the system performance on the flight vehicle. The remote adjust input is set to the maximum value, and the unit output is enabled by driving the system enable pin to ground. Figure 8 shows the complete electronics interface diagram for the launch-assist motor subsystems. At this point in the design process, the complete vehicle electronics interface to the motor subsystems has not been entirely defined.



Figure 7. Schematic of the UltraVolt HVPS System pinouts and interface.



Figure 8. Launch-assist motor systems electronics interface diagram.

# **III. Ground Test System Overview**

The ground test system used to perform the preliminary integration and gualification tests on the LAPU subsystems has a more extensive instrumentation suite including, an in-line Venturi flow meter on the oxidizer feed line downstream of the pressure regulator and multiple thermocouples to monitor the system temperatures at various points. The system is integrated onto a portable ground test cart (Figure 9) with all hot-fire testing performed in the Propulsion Research Laboratory's on-campus test cell. Figure 10 shows the piping and instrumentation test schematic for the ground test system. The ground test system is operated using a National Instruments Universal Serial Bus (USB)cDAO-9174 Acquisition based NI Data and Control Unit (http://sine.ni.com/nips/cds/view/p/lang/en/nid/207535/) with data logging and system control via a LabVIEW interface program.



Figure 9. Ground test cart for TGALS LAPU verification testing.



Figure 10. Ground test motor systems P&ID.

## **IV. Ground Test Results**

A series of ground development tests using three different configurations were performed. Two sets of static fire tests were performed to verify the throttle capability of the system. The final series of tests evaluated the preliminary performance of the closed-loop throttle system. Closed-loop throttle systems control law tuning is current being completed.

## A. Cylindrical Port Hot-fire Tests

The first series of ground tests were performed using an existing ABS fuel grain left over from the earlier nitrous oxide (N<sub>2</sub>O)/ABS testing campaign of Whitmore and Peterson.<sup>10</sup> Stratasys, Inc. printed this fuel grain as a single monolithic piece with a density of  $0.975 \text{ g/cm}^3$ Fortus 900mc production using FDM machine а (www.stratasys.com/3d-printers/production-series/fortus-900mc/). This series of 5 tests were performed at the full throttle position with the ball-valve set in the full opening position. After each test, the motor fuel grain was removed from the motor case, and the consumed fuel mass was measured.

Figure 11 plots the measured resulting regression rates as a function of the oxidizer massflux ( $G_{ox}$ ). For these calculations, the oxidizer massflow was measured using an in-line calibrated Venturi massflow meter. The fuel massflow was calculated as the difference between the measured oxidizer massflow and the nozzle exit massflow. The nozzle exit massflow was calculated based on the measured chamber pressure  $P_{0}$ , nozzle exit area  $A^*$ , and exhaust gas properties using the onedimensional De Laval choking massflow equation.<sup>11</sup>



Figure 11. Regression rate for GOX/ABS cylindrical port tests.

The combustion products for the combustion flame temperature  $T_0$ , gas-specific constant  $R_g$ , and ratio of specific heats  $\gamma$  were calculated using tables developed using the NASA chemical equilibrium program "Chemical Equilibrium with Applications," (CEA).<sup>12</sup> For the CEA calculation, the measured chamber pressure was used as an input, and the oxidizer-to-fuel massflow (O/F) ratio entered into CEA was adjusted to produce a fuel massflow whose integral value exactly equaled the consumed fuel mass measured after each test.

Figure 11 also plots an exponential curve fit of the form

$$\dot{r}_L = a \cdot G_{ox}^n, \tag{1}$$

where  $\dot{r}_L$  is the mean longitudinal regression rate,  $\{a\}$  is the scale factor,  $G_{ox}$  is the oxidizer mass flux,  $\{n\}$  is the burn exponent. For a cylindrical fuel port, it can be shown that the O/F at any burn time is

$$O/F = \frac{1}{4^n \cdot \pi^{1-n}} \frac{\dot{m}_{ox}^{1-n} \cdot D_{pon}^{2n-1}}{a \cdot \rho_{fuel} \cdot L}$$
(2)

Analysis of Eq. (2) shows that when the burn exponent is  $\{n > 1/2\}$ , the *O/F* ratio is progressive and increases as the fuel grain burns and the port opens up. Conversely, when  $\{n < 1/2\}$  the *O/F* burn is regressive and becomes increasingly rich with time, and when  $\{n = 1/2\}$  the burn rate is neutral and implies no *O/F* shift during the burn. The majority of commonly used oxidizer/fuel combinations (including  $N_2O/ABS$ ) have burn exponents greater than 1/2, and thus burn increasingly leaner with time.<sup>13</sup>

For the GOX/ABS grain cylindrical fuel port tests, the resulting best-fit burn parameters are

$$\begin{bmatrix} a \\ n \end{bmatrix} = \begin{bmatrix} 0.048 \frac{cm^{1+2n}}{g-s^{1-n}} \\ 0.45 \end{bmatrix}.$$
 (3)

The value for the burn exponent  $\{n=0.450\}$  is considerably smaller than the value previously measured,  $\{n\sim0.762\}$ .<sup>10</sup> The derived burn exponent suggests that the LAPU motor should exhibit very little O/F shift during the burn. Figure 12 verifies this assertion where O/F is plotted as a function of oxidizer mass flux. The O/F shift is slightly regressive with the motor burning only slightly richer as the fuel port opens up. This quantitative behavior matches the qualitative physical observations of the various motor burns. The larger 98-mm motor plume was observed to show very little change in the plume characteristics during the fuel grain burn lifetime—  $\sim 20$  seconds.



Figure 12. Oxidizer-to-fuel shift of cylindrical ABS fuel port.

## B. Helical Port Static Tests

#### 1. Measuring the Required Regulator Set Point

As shown by Figure 12, the O/F ratios for the cylindrical port fuel grain lie just above the optimal operating value of ~1.5 for GOX/ABS.<sup>5</sup> Thus, for the second series of tests, a moderate helix was printed into the fuel grain to lower the mean O/F ratio. As listed in Table 1, the port helix radius was 1/2 of the initial fuel port diameter. The helix pitch length was 19.5 cm (7.68 in.) resulting in 3 complete turns along the fuel port length. This change was incorporated to slightly lower the O/F ratio so that the motor would burn nearer the optimal operating condition. The grain was printed as three interlocking segments on the Mechanical and Aerospace Engineering department's Dimension 1200es (http://www.stratasys.com/3d-printers/design-series/dimension-

1200es/) using ABSplus-340 feedstock. In addition to the modification of the fuel port, the nozzle retainer exhibited unwanted erosion, and a small redesign was made to reduce the erosion potential.

The testing campaign on the helical fuel port motor was broken into two sets. A primary function of the first test set was to measure the necessary regulator set point to achieve the full required 200-lbf thrust level. A series of four 2-second burns at various regulator set points were performed to determine the required set pressure level. The regulator output "droop" was found to be strongly a function of oxidizer massflow. Figure 13a plots regulator output from the set point (droop) as a function of the oxidizer massflow, and Figure 13b plots the achieved output pressure as a function of the regulator set point. The result is that the achieved full-open thrust and chamber pressure are also strongly a function of the regulator set point. Thrust coefficient is also strongly influenced by the regulator set point. Figure 14 plots this result. With the throttle ball-valve fully open, the full required thrust level of 200 lbf mandates a regulator set point of at least 750 psig.



Figure 13. Regulator output droop as a function of oxidizer massflow set point.



Figure 14. Full throttle thrust chamber pressure and thrust coefficient as functions of regulator set point.

## 2. Oxidizer-to-Fuel Ratio for the Helical Fuel Port

The initial tests of the helical fuel port were also used to verify that the changes to the grain configuration moved the O/F ratio to the optimal operating range. Here Figure 15a plots the mean O/F ratio for each burn as a function of the accumulated burn time on the motor. The resulting O/F range—between 1.25 and 1.67—is plotted overlaid onto the characteristic velocity plot  $C^*$  in Figure 14b. Here the achieved O/F range brackets the optimal performance range, thereby verifying the helix grain design.



**Figure 15.** Achieved O/F range for helical fuel port burn at near full throttle and the resulting effect on  $C^*$ .

Figure 16 compares the achieved C\* calculated by



Figure 16. C\*, combustion efficiency, and  $I_{sp}$  as a function of burn time.

against the 100% combustion efficiency theoretical values for the O/F levels of Figure 15. Figure 16 also plots the measured combustion efficiency and specific impulse,  $I_{sp}$ . The achieved combustion efficiencies, as calculated by Eq. 5, are slightly less than 80%.

$$\eta^* = \frac{C^*_{measured}}{C^*_{theoretical}}$$
(5)

The low observed combustion efficiency is also reflected by the measured specific impulse of the system, which at a value of  $I_{sp}=205$  seconds is approximately 10% lower than predicted. The small drop in C\* is a result of a slight observed nozzle erosion during the 8 seconds of burn time. The reasons for the low observed performance levels have

yet to be determined. Two causes currently being investigated include an incompletely cured printed fuel grain and potential blow-by at the phenolic liner/graphite nozzle interface.

## 3. Throttle Curve Evaluation

After the initial set of tests to determine the appropriate regulator set point, a series of 6 static throttle burns were performed at different ball-valve voltage command levels and using a newly fabricated helical fuel grain. To ensure that the fuel grain was fully cured, the newly printed grain was placed in a vacuum chamber and then left in front of a fan overnight before performing these tests. Each burn was set as 2 seconds in length using up ~12 seconds of total burn lifetime.

Figure 17 summarizes the test results where the achieved motor thrust, massflows, chamber pressure, and thrust coefficients are plotted as a function of the commanded ball valve voltage. Figure 16a also plots the required 200 lbf thrust full-throttle level. The effective range of the ball valve servo voltage command varies from 1.25 Volts (0% throttle) to 2.3 Volts (100% Throttle).



Figure 17. Helical motor response as function of commanded ball-valve voltage level.
Figures 18 and 19 plot the system performance parameters including combustion efficiency and specific impulse as a function of the commanded throttle level and the equivalent throttle actuator voltage command. At full throttle, the system achieves slightly better combustion efficiency (>80%) than was observed with the previous static tests, but this combustion efficiency and the associated specific impulse drops off significantly at the lower throttle levels. The plots of Figure 19 support the earlier assertion that the lower than predicted specific impulse for the system is a result of lowered combustion efficiency.



Figure 18. Combustion efficiency as a function of throttle command.



Figure 19. Specific impulse as a function of throttle command.

#### C. Closed-Loop Throttle Testing

The results of the static throttle tests were curve fit and used to generate look-up tables that correlate the servo command voltage to motor thrust and chamber pressure to motor thrust. Using these data tables, a proportional-gain closed-loop controller using chamber pressure feedback was developed and implemented within the LabVIEW code that resides on the controlling laptop computer. Measured feedback data and closed-loop commands are sent to and from the NI cDAQ-9174 Data Acquisition and Control Unit via an amplified USB extension.

The flow chart in Figure 20 shows the implemented control law. The control features chamber pressure feedback with either open-loop or recursive servo-voltage output commands. During recursive control, an option for smoothing the commanded voltage using a second-order Butterworth filter is included. Options for user-prescribed step, ramp, or arbitrary inputs are available. Values for the proportional gain  $k_p$  and the lowpass filter cutoff frequency  $\omega_p$  are user selectable.



Figure 20. Proportional closed-loop controller layout.

Figure 21 shows the result of a four hot-fire tests performed with various set values for  $k_p$  and  $\omega_p$ . For these tests, the motor was ignited with the actuator command set for the 25% thrust (50 lbf) level, and the commanded thrust was increased to 25% (100 lbf) 2 seconds into the burn. The lowest gain  $k_p = 0.3$  and a command filter rolloff frequency of  $\omega_p = 94$  radians/sec (15 hz) produced the response with the minimal overshoot. The observed response latency is primarily a function of the actuator response time.



Figure 21. Thrust profiles for various closed-loop step tests

#### 1. Closed Loop Throttle Filter Tuning

Using the controller command logic from the closed-loop step tests, a simulation was designed so that  $\omega_p$  and  $k_p$  can be "tuned" without requiring multiple trial-and-error hot-fire tests. The simulation decomposes the system dynamics into two concatenated components: 1) a model of the servo and control ball-valve dynamics, and 2) a model of the ballistic response of the motor combustion chamber. Both responses are modeled as second-order transfer functions. The ball servo and ball-valve transfer function is

$$\frac{\frac{96_{MVT}}{V_{ond}}}{\left[\frac{\tau_1}{2\cdot\varsigma_1}\right]^2\cdot s^2 + \tau_1 s + 1},$$
(6)

where { $\tau_1 = 0.52$ } and { $\zeta_1 = 0.85$ }. These values are based on Invenscience specifications (http://www.invenscience.com/index\_files/ torxis\_rotary\_servo.htm/) for the servo response properties. In Eq. (6),  $\mathcal{H}_{MVT}$  is the percentage of mean valve travel from fully closed to fully open— ~90°, and  $V_{cmd}$  is the servo command voltage level. The motor ballistics model transfer function is

$$\frac{T_{brain}}{\mathscr{H}_{NVT}} = \frac{1}{\left(\frac{\tau_2}{2 \cdot \varsigma_2}\right)^2 \cdot s^2 + \tau_2 s + 1},$$
(7)

where the values for  $\{\tau_2\}$  and  $\{\zeta_2\}$  are iterated to give the best fit between the simulator and measured response. Figure 22 shows the calculation sequence that was used to estimate the best-fit transfer functions. For a given control law parameter setting for  $\omega_p$  and  $k_p$ , the simulation was run multiple times, sweeping through the 2-dimensional parameter space for {  $\tau \cdot 2$ ,  $\zeta \cdot 2$ }. In this two-dimensional parameter space, the parameter set that produces the minimum root-sum-square (RSS) error between the measured system and simulation response is selected as the "best-fit" for the motor ballistics.



Figure 22. Finding the best-fit transfer function parameters.

Figure 23 shows the RSS best-fit comparison between the simulation and hot-fire test data for a prescribed throttle profile corresponding to a push-over–pull-up maneuver of the TGALS vehicle.<sup>14</sup>



**Figure 23.** Simulation compared with hot-fire test with ballistic parameters { $\tau_2 = 0.7$  and  $\zeta_2 = 1.9$ }, and motor controller parameter set { $\omega_p = 2\pi \cdot 15 \text{ rad/sec}, k_p = 0.3$ }.

Using this transfer function parameter set for the ballistic model, various values of the control law parameter set  $\{\cdot_p \text{ and } k_p\}$  were evaluated, which allowed the control law to be tuned for a best system response. Figure 24 shows a sample of various simulation runs and

illustrates the effects of the control law parameter set. These simulation results will be verified through follow-on hot-fire throttle tests.



Figure 24. Sample controller tuning simulation runs.

#### V. Conclusion

This document presents a status update of the design and integration of a throttled launch-assist hybrid rocket motor for an airborne nano-launch platform. Currently, NASA AFRC is developing a scaled prototype of a high L/D ratio glider designed as a flexible low earth orbit launch platform for nano-scale satellites (NanoSats). Because the high L/D platform is delivered to the launch altitude and airspeed using a high-efficiency air-breathing propulsion system, there is a significant reduction in the required  $\Delta V$  that must be delivered by the launch vehicle. Optimal  $\Delta V$  savings are achieved when the NanoSat launch vehicle is delivered to a high flight path angle that will approximate the condition that would be achieved along a ground launch trajectory at the same altitude and airspeed.

The glider platform itself is unable to achieve this flight condition, and launch-assist propulsion is required. A hybrid system was selected for the launch-assist motor because of the inherent safety, operational simplicity, and environmental friendliness of the propellants and because the hybrid system can be throttled and re-started on demand. This study establishes the requirements for this launch-assist propulsion system, develops the system design features, presents the end-to-end hardware layout, and develops the closed-loop throttle control law.

The control features chamber pressure feedback with either openloop or recursive servo-voltage output commands. During recursive control, an option for smoothing the commanded voltage using a second-order Butterworth filter is included. Options for user-prescribed step, ramp, or arbitrary inputs are available. Values for the proportional gain  $k_p$  and the lowpass filter cutoff frequency  $\omega_p$  are user selectable. A medium fidelity motor simulation is derived from preliminary ground test data and is used to tune the parameters of the closed-loop control law without having to perform multiple hot-fire tests.

Initial static tests were performed with a cylindrical fuel port to verify system functionality and establish a baseline for the propellant regression rate and optimal O/F ratio. Subsequent tests are performed using a helical fuel port to increase the volumetric efficiency of the system and allow operation near the optimal oxidizer-to-fuel condition. Multiple restarts of each system configuration are demonstrated. Results of both open- and closed-loop throttle tests are presented.

Static throttle tests were curve fit and used to generate look-up tables that correlate the servo command voltage to motor thrust and chamber pressure to motor thrust. Using these data tables, a proportional-gain, closed-loop controller using chamber pressure feedback was developed and implemented within the real-time code that resides on the controlling laptop computer. Currently, closed-loop system tuning is incomplete, and the testing campaign to tune the system before flight test will be completed over the next few months, with complete results to be reported in subsequent publications.

### References

1 Bartolotta, Paul, Wilhite, Alan W., Schaeffer Mark, Voland, Randall T., and Huebner, Larry, "Horizontal Launch: A Versatile Concept for Assured Space Access," NASA/SP-2011-21599, December 2011. https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20120000791.pdf.

2 Lanre, A. M., Fashanu, T. A., and Osheko, A. C., "Parameterization of Micro Satellite Launch Vehicle," J. Aeronautics Aerospace Eng., Vol. 3, No. 2 (2014), pp. 134-143, DOI: 10.4172/2168-9792.1000134.

3 Hague, N., Siegenthaler, Erika, Rothman, Julia, "Enabling Responsive Space: F-15 Microsatellite Launch Vehicle," IEEEAC paper #1102, 2003 IEEE Aerospace Conference Proceedings (Cat. No.03TH8652), IEEEXplore, Vol. 6-2703, DOI: 10.1109/AERO. 2003.1235195.

4 Chen, T. T., Ferguson, P. W., "Responsive Air Launch Using F-15 Global Strike Eagle," RS4-2006-2001, 4th Responsive Space Conference, Los Angeles CA, April 24-27, 2006, DOI:10.1.1.560.4885.

5 Whitmore, Stephen A., Walker, Sean D., Merkley, Daniel P., and Sobbi, Mansour, "High Regression Rate Hybrid Rocket Fuel Grains with Helical Port Structures," J. Propulsion Power, Vol. 31, No. 6 (2015), pp. 1727-1738, DOI: 10.2514/1.B35615.

6 Whitmore, Stephen A., Inkley, Nathan, and Merkley, Daniel P., "Development of a Power-Efficient, Restart-Capable Arc Ignitor for Hybrid Rockets," J. Propulsion Power, Vol. 31, No. 6 (2015), pp. 1739-1749, DOI: 10.2514/1.B35595.

7 Anonymous, "Department of Defense Interface Standard, Electromagnetic Environmental Effects Requirements for Systems," MIL-STD-464C, AMSC 9159, DECEMBER 1, 2010, Retrieved February 9, 2017, from https://snebulos.mit.edu/projects/ reference/MIL-STD/MIL-STD-464C.pdf.

8 Anonymous, "High Power 8C-30C Series, Single Output High Voltage DC/DC Modules," UltraVolt, Inc., Retrieved October 9, 2014, from http://www.ultravolt.com/uv\_docs/HP8C-30CDS.pdf.

9 Whitmore, S. A., Merkley, S. L., Tonc, L., and Mathias, S. D., "Survey of Selected Additively Manufactured Propellants for Arc Ignition of Hybrid Rockets", J. Propulsion Power, Vol. 32, No. 6 (2016), pp. 1494-1504. DOI: 10.2514/1.B36106.

<sup>10</sup> Whitmore, Stephen A., Peterson, Zachary W., and Eilers, Shannon D., "Comparing Hydroxyl Terminated Polybutadiene and Acrylonitrile Butadiene Styrene as Hybrid Rocket Fuels," AIAA J. Propulsion Power, Vol. 29, No. 3, May–June 2013, DOI: 10.2514/1.B34382.

11 Sutton, G. P., and Biblarz, O., Rocket Propulsion Elements, 7th ed., Chapters 4-5, Wiley, New York, 2001.

12 Gordon, S., and McBride, B. J., "Computer Program for Calculation of Complex Chemical Equilibrium Compositions and Applications," NASA RP-1311, 1994. https://www.grc.nasa.gov/www/ceaweb/rp-1311.pdf.

13 Karabeyoglu, Aruf, "Lecture 10 Hybrid Rocket Propulsion Design Issues," AA 284a Advanced Rocket Propulsion, AAE Department, Stanford University, May 14, 2012, p. 18, Retrieved January 1, 2016, from http://www.spg-corp.com/docs/Stanford\_AA284a\_Lecture10.pdf.

14 Whitmore, Stephen A., Merkley, Stephen L., Walker, Sean D., and Spurrier, Zachary S. "Throttled Launch-Assist Hybrid Rocket Motor for an Airborne NanoSat Launch Platform," 51st AIAA/SAE/ASEE Joint Propulsion Conference, Propulsion and Energy Forum, (AIAA 2015-3940), 28-30 July 2015, DOI: 10.2514/6.2015-3940.

## That's Rich: Survival at the Tudor Court

#### **Joshua Wiggins**

Southern Utah University

#### Abstract

Men and women at the Tudor court could garner breathtaking power and wealth so long as they remained in favor with King Henry VIII, but retribution for failure to please him was swift and brutal. Thomas Wolsey, Anne Boleyn, and Thomas Cromwell all fell victim to this cycle. One man spectacularly broke the trend by becoming more corrupt than the system he served and living to have it documented. This man died peacefully of old age, surrounded by the enormous fortunes he had gathered. This was Sir Richard Rich. This paper examines some significant moments in Sir Rich's life to document the motivation for everything he did: survival at court.

It has been well documented how merciless the Tudor court could be; men and women could garner breathtaking power and wealth so long as they were in the monarch's good graces, but could be swiftly destroyed when they (almost inevitably) fell out of favor. The objective of Tudor court politics was simple: stay in the favor of the monarch. The reality, of course, was far more complex. In the case of Henry VIII, remaining in his good graces was no simple matter: his capricious and volatile nature, especially later in his life, created an equally unstable environment at court. Cardinal Thomas Wolsey, for example, who had served his master well for fourteen years, lost everything—including his life—when he failed to secure the annulment from Catherine of Aragon that King Henry desired.

In addition to observing the ruler, an ambitious courtier would need to carefully gauge the relative power of the opposing factions at court—noble families against other nobles, or evangelists against Roman Catholics, for example—who competed for influence and the monarch's sympathies. To be caught on the losing side would invariably prove disastrous, as the Boleyn faction discovered during Queen Anne's fall from power. Thomas Cromwell, then Lord Chancellor, successfully implicated Anne's brother George in the resulting scandal.<sup>1</sup>

Such were the perils of being so close to the sovereign. Like Icarus of Greek lore, soaring to such great heights risked plummeting to ruination. It was a gamble that could yield enormous benefits; it was a sort of chess match between the various players at court to gain the most power without destroying themselves.

If Tudor court politics were a game, then Sir Richard Rich (1496?-1567) won decisively. Throughout his service to the Tudor crown, Rich had accumulated massive wealth with his role in the dissolution of the monasteries. Unlike so many of his contemporaries, Rich retained his fortune and retired in astounding luxury. He outplayed the Tudor power structure through his willingness to do whatever the monarch wanted, even when it meant behaving unethically. By examining a few key moments in this man's life, it becomes clear that the primary purpose of his infamous conduct was to survive at all costs.

Although most examinations of historical figures begin with the subject's childhood and upbringing, little was documented concerning Richard Rich's youth. John Sargeaunt, the academic and schoolmaster of Felsted School (the educational institution founded by Richard Rich later in his life) related in his work *A History of Felsted School with* 

<sup>&</sup>lt;sup>1</sup> British History Online, "Spain: June 1536, 6-10," in *Calendar of State Papers, Spain, Volume 5 Part 2, 1536-1538*, ed. Pascual de Gayangos (London: Her Majesty's Stationery Office, 1888), 137-162. http://www.british-history.ac.uk/cal-state-papers/spain/vol5/no2/pp137-162 (accessed April 15, 2016). The author of this letter attributes Cromwell as the mastermind behind the scandal: "He [Cromwell], himself had been authorised and commissioned by the King to prosecute and bring to an end the misstress's trial, to do which he had taken considerable trouble. It was he who, in consequence of the disappointment and anger he had felt on hearing the King's answer to me on the third day of Easter, had planned and brought about the whole affair."

Some Account of the Founder and His Descendants that Rich's father was "a wealthy mercer."<sup>2</sup> It was with this wealth that Rich was able to study common law and "pick up, in occasional fits of industry, enough law to pass muster with attorneys for his audience."<sup>3</sup> Rich's early career gains were steady yet impressive. Thomas Cromwell appointed him to be the "lesser hammer" in the dissolution of the monasteries,<sup>4</sup> and it was in this capacity that he mingled with the members of Henry VIII's court. He used his connections to bolster his own position at the expense of those around him. One such incident involved Bishop John Fisher.

Richard Hall, a professor during Elizabeth's reign, gives the account of Rich's betrayal of Fisher in his panegyric The Life and Death of That Renowned John Fisher Bishop of Rochester. According to Hall, King Henry secretly dispatched Rich to the Tower of London under the pretext of innocently discerning Bishop Fisher's opinion on the king's supremacy; Fisher had been sent to the Tower after refusing to acknowledge Henry's supremacy over the English church. Hall tells us that Rich first flattered Fisher, telling him "[T]hat there is no man within the Kings dominions, that he looks upon, as a more able man, or a man upon whose honesty and upright dealings he will more rely, than on your self."<sup>5</sup> He masked the true reason for his visit: if Fisher "should tell him plainly [the king's supremacy] were unlawfull, [the king] would never undertake it."<sup>6</sup> He then promised that his visit was "without any the least suspition of any fraud or guile intended to [Fisher], either by him or me," and that Fisher could "speak [his] minde boldly and freely unto [Rich], as to [the king] himselfe."<sup>7</sup>

The Tudor historian A. F. Pollard disputes some aspects of this account, arguing Hall's stated pretext of the visit (i.e., that Rich was sent in secret by the king) to be "improbable," but he agreed that Rich "came to the Tower and endeavored to ascertain the bishop's real views

<sup>&</sup>lt;sup>2</sup> John Sargeaunt, *A History of Felsted School with Some Account of the Founder and His Descendants* (Chelmsford: Edmund Durrant, 1889), 80.

<sup>&</sup>lt;sup>3</sup> Ibid., 81.

<sup>&</sup>lt;sup>4</sup> J. Hall Pleasants, "The Lovelace Family and Its Connections (Continued)," *The Virginia Magazine of History and Biography* 28, no. 4 (October 1920), 390. http://www.jstor.org/ stable/4243795 (accessed January 28, 2016).

<sup>&</sup>lt;sup>5</sup> Richard Hall, *The Life & Death of That Renowned John Fisher Bishop of Rochester* (London, 1655), 172. http://quod.lib.umich.edu/e/eebo/A45326.0001.001?view=toc (accessed March 7, 2016).

<sup>6</sup> Ibid.

<sup>&</sup>lt;sup>7</sup> Ibid., 172-3.

on the subject."<sup>8</sup> In any case, Fisher decided to trust him; he had no reason to doubt Rich's intentions. "As to the businesse of Supremacy," Fisher declared, "I must needs tell his Majesty, as I have often told him heretofore, and would so tell him, if I were to die this present hour, that it is utterly unlawfull; and therefore I would not wish his Majesty to take any such power or title upon him, as he loves his own soule, and the good of his posterity."<sup>9</sup> Fisher felt safe divulging this sentiment under the protection afforded by Rich's promise.

Rich, on the other hand, displayed no qualms using Fisher's confession to the king's (and thus his own) advantage. Fisher was tried and convicted on the evidence that Rich had manipulated out of him, though in vain the bishop decried the treacherous act.<sup>10</sup> This would become a common theme in Rich's life: he was willing to cross ethical lines to ensure his own survival, power, and position. In a similar fashion, Richard Rich stood at the center of the downfall and execution of one of Bishop Fisher's friends, Thomas More. Unlike Fisher, More declined to reveal his position on the supremacy issue. His deliberate and clever refusal to answer the question enraged the king, but as More would successfully argue, "[S]ilence itself is no crime. In treason an overt act must be proved."<sup>11</sup> By maintaining his silence, he could not be incriminated, yet that same silence was particularly damaging to the king. More was respected both abroad and in England, and his opinion was highly regarded.

In an effort to wear down More's resistance, the king, as in the case of Bishop Fisher, imprisoned More in the Tower of London, albeit with some degree of comfort commensurate with his social status. When More's relatively pleasant confinement failed to alter his stance, Sir Rich was dispatched to remove the luxuries More had enjoyed: his books and writings. While Rich's team set to the task, Rich struck up a conversation with More in an attempt to entrap the older man with a series of hypothetical scenarios. The objective was to trick him into revealing his position, thus breaking the stalemate between More and King Henry. Unlike Bishop Fisher, More knew of Rich's duplicitous nature from a young age. "And I, as you know, of no small while have been acquainted with you and your conversation, who have known you

<sup>&</sup>lt;sup>8</sup> A. F. Pollard, "Rich, Richard," in *Dictionary of National Biography*, ed. Sidney Lee (New York: MacMillan, 1896), 48, 123. https://archive.org/details/ dictionaryof-nati48stepuoft (accessed April 9, 2016).

<sup>&</sup>lt;sup>9</sup> Hall, John Fisher, 174.

<sup>&</sup>lt;sup>10</sup> Pollard, "Rich," 123.

<sup>&</sup>lt;sup>11</sup> J. Duncan M. Derrett, "The Trial of Sir Thomas More," *The English Historical Review* 79, no. 312 (July 1964), 459. http://www.jstor.org/stable/560988 (accessed November 21, 2015).

from your youth hitherto; for we long dwelled both in one parish together," More told Rich at his trial, "where, as you yourself can well tell (I am sorry you compel me so to say) you were esteemed very light of your tongue, a common liar, a great dicer and of no commendable fame."<sup>12</sup>

According to Rich's own testimony at More's trial, Rich succeeded in tricking More into stating his opposition to the king's supremacy over the church. The veracity of Rich's statement has always been suspicious, if not viewed as an outright fabrication. "If Rich is correct in his testimony," the philologist Edward Berry reasons, "then More not only broke his vow of silence but, uncharacteristically, refused to allow the game of 'putting the case' to end in a moot point. [...] The more likely possibility, however, is that More played the game with his characteristic skill and that Rich's forceful closure was his own treacherous invention."<sup>13</sup> More vehemently denied Rich's version of the story in a stinging rebuke:

In good faith, Master Rich, I am sorrier for your perjury than for mine own peril. And you shall understand that neither I, nor any man else to my knowledge, ever took you to be a man of such credit as in any matter of importance I, or any other, would at any time vouchsafe to communicate with you. [...] Can it therefore seem likely to your honourable Lordships that I would, in so weighty a case, so unadvisedly overshoot myself as to trust Master Rich (a man of me always reputed for one of so little trust, as your Lordships have heard) [...] that I would unto him utter the secrets of my conscience touching the King's Supremacy, the special point and only mark at my hands so long sought for, a thing which I never did, nor never would, after the Statute thereof made, reveal either to the King's Highness himself or to any of his honourable Councillors, as it is not unknown to your honours, at sundry several times sent from His Grace's own person unto the Tower to me for none other purpose? Can this in your judgements, my lords, seem likely to be true?<sup>14</sup>

<sup>&</sup>lt;sup>12</sup> William Roper and Nicholas Harpsfield, *Lives of Saint Thomas More*, ed. E. E. Reynolds (London: Aldine Press, 1963), 159. https://archive.org/details /livesofsaintthom000169mbp (accessed February 10, 2016).

 <sup>&</sup>lt;sup>13</sup> Edward Berry, "Thomas More and the Legal Imagination," *Studies in Philology* 106, no. 3 (2009), 338. http://www.jstor.org/stable/25656018 (accessed January 28, 2016).
 <sup>14</sup> Roper, *Thomas More*, 158-9.

The servants who had packed away More's books were called to verify Rich's story. Ominously for the prosecution, the witnesses deliberately refused to take sides. "Whereupon Master Palmer, upon his deposition, said that he was so busy about the trussing up of Sir Thomas More's books in a sack, that he took no heed to their talk. Sir Richard Southwell likewise, upon his deposition, said that because he was appointed only to look to the conveyance of his books, he gave no ear unto them."<sup>15</sup> Lying to appease the prosecution would not have sat well with their consciences, especially in a religious climate where the consequences in an afterlife were seen as very real; however, explicitly denying that Rich's version of the story was true would surely have endangered their own lives by embarrassing the king's case. Taking the middle ground was the only solution for their dilemma. Had the witnesses chosen to challenge Rich's testimony, this case would surely have cost Rich his career, if not his life. Fortunately for him, Rich's gamble paid off. More was convicted and executed, and Rich's status at the King's court remained exceedingly favorable-so favorable, in fact, that he was "appointed [...] overseer of liveries of lands, and chirographer of common pleas."<sup>16</sup>

These two incidents were not the only times Rich would conveniently possess the necessary evidence to destroy the targets of the king's displeasure. "[Rich] was always found on the winning side," Pollard noted, "and he had a hand in the ruin of most of the prominent men of his time, not a few of whom had been his friends and benefactors—Wolsey, More, Fisher, Cromwell, Wriothesley, Lord Seymour of Sudeley, Somerset, and Northumberland."<sup>17</sup> Yet as dangerous as he could be in the courtroom, Rich's sense of self-preservation extended beyond merely uttering whatever words would best serve the king's interests. Indeed, as an instrument in ruining these "prominent men," he observed firsthand how completely an individual could be destroyed if the king listened to the whisperings of his adversaries. He therefore extended his behavior beyond that of the sycophant; he was, in short, willing to get his hands dirty.

One particularly repellent incident demonstrated this aspect of Rich's character, and the unfortunate victim this time was Anne Askew, a brilliant, "highly educated" and biblically savvy young

<sup>&</sup>lt;sup>15</sup> Ibid., 160.

<sup>&</sup>lt;sup>16</sup> Pollard, "Rich", 124.

<sup>&</sup>lt;sup>17</sup> Ibid., 126.

woman.<sup>18</sup> Rich targeted the twenty-five-year-old because of her public evangelical and heretical views. Although ostensibly under scrutiny for her views regarding the Eucharist,<sup>19</sup> she was likely imprisoned in the Tower of London so that Rich could extract evidence to implicate Queen Catherine Parr as a Protestant, and thus a heretic. Historian Boyd Berry explained that "[Anne] was put on the rack in an effort to discredit Catherine Parr."<sup>20</sup> Ever the opportunist, Rich was determined to earn favor with Henry by deposing his latest out-of-favor wife. Askew, it appeared, was the key.

Remarkably, Askew withstood questioning as the king's agents attempted to discover other members of her group (and, by extension, the queen); she denied them the information they sought.<sup>21</sup> Historian James Gairdner wrote: "More than this even the rack could not get from her, which [...] was applied by Lord Chancellor Wriothesley himself and Sir Richard Rich, turning the screws with their own hands."<sup>22</sup> This highly unusual and unprecedented example—called "the darkest page in the history of any English statesman"<sup>23</sup>—demonstrates the extreme measures Rich employed to further his liege's agenda and thus preserve his own usefulness.

Perhaps nothing demonstrates his will to survive quite like his role in persecuting whatever religious faction found itself out of favor. Pollard's assertion that "[h]is religious predilections inclined to catholicism"<sup>24</sup> supports what historian J. Hall Pleasants describes: "Rich was a Roman catholic at heart, for although he aided Henry VIII in despoiling the monasteries, and in dispossessing bishops Bonner and Gardiner in the reign of Edward VI, when Mary was on the throne he founded a chaplaincy providing for the singing of masses in Felsted church."<sup>25</sup> Despite Rich's personal views on religion, both historians agree he was

<sup>21</sup> Gairdner, "Askew," 191.

<sup>&</sup>lt;sup>18</sup> James Gairdner, "Askew, Anne," in *Dictionary of National Biography*, ed. Leslie Stephen (New York: MacMillan, 1885), 2, 190. https://archive.org/details/ dictionaryof-nat02stepuoft (accessed May 23, 2016).

<sup>19</sup> Ibid.

<sup>&</sup>lt;sup>20</sup> Boyd Berry, review of *The Examinations of Anne Askew*, ed. Elaine V. Beilin, *Journal of English and Germanic Philology* 89, no. 2 (April 1999): 260. http://www.jstor.org/ stable/27711802 (accessed April 16, 2016). Berry admits, "Needless to say, she does not mention Parr; but her discussion of how she shuns human law to conform to divine law shows that she clearly understood her legal position" (260).

<sup>&</sup>lt;sup>22</sup> Ibid.

<sup>&</sup>lt;sup>23</sup> James Anthony Froude, *History of England from the Fall of Wolsey to the Death of Elizabeth*, Vol. 4, 3rd ed. (London: Longman, Green, Longman, Roberts, & Green, 1864), 504. https://books.google.com/books?id=dNE\_AAAAcAAJ (accessed May 23, 2016).

<sup>24</sup> Pollard, "Rich," 126.

<sup>&</sup>lt;sup>25</sup> Pleasants, "The Lovelace Family," 391.

"equally detested by catholic and protestant alike"<sup>26</sup> for his role in the persecution of both factions. This flexibility supplied him with the indispensability necessary to survive the reigns of four Tudor monarchs throughout the English reformation and counter-reformation periods.

How did Rich's influence finally dissipate? Sargeaunt relates the circumstances of Rich's retirement from court life:

There is, however, better evidence for another account. According to this[,] Rich, perceiving how Somerset was again rising to power, began again to court his favour. Even when Somerset had been betrayed and sent to the Tower there was a chance that his popularity might yet avail him. Lord Rich was therefore anxious to serve him, and sent a letter to the Tower to warn him of some design that Warwick had against him. He directed his letter in such haste that it was addressed merely "To the Duke." The Duke of Norfolk was also in the Tower, and the servant who bore the letter, knowing his master's acquaintance with Norfolk, and ignorant of his renewed communication with Somerset, handed the letter to the wrong Duke. Lord Rich heard of the mistake in time. He took to his bed and wrote to the King to say that he was exceeding sick and must resign the Great Seal forthwith: let officers be sent to fetch it, for he was too weak to move. The device saved him and he retired in peace to Leez.<sup>27</sup>

This narrative is a reassuring, but ultimately inaccurate, version of the events. It is comforting to think that someone so brazenly immoral was so close to losing everything he had worked for in one fell blow. It satisfies our sense of justice. However, Pollard dismisses this version, noting, "[I]t is improbable, however, that Norfolk, who made Rich one of his executors, would have betrayed him; at any rate, Rich did not resign the great seal to the king, but to Winchester, Northumberland, and D'Arcy, who were sent to his house for the purpose, and there can be no doubt of the genuineness of his illness."<sup>28</sup>

Perhaps this story perhaps tells more about human nature than about Rich. Surely his outrageous behavior elicited some punishment. However, Rich escaped any sort of retribution for his actions; he retired in splendid luxury and used part of his vast fortunes to found "a gram-

<sup>&</sup>lt;sup>26</sup> Ibid., cf. Pollard, "Rich," 126.

<sup>&</sup>lt;sup>27</sup> Sargeaunt, Felsted School, 85-86.

<sup>28</sup> Pollard, "Rich," 125.

mar school at Felsted".<sup>29</sup> Rich's actions cemented his legacy as "one of the most ominous names in the history of the age."<sup>30</sup> His family crest read "*Garde Ta Foy*":<sup>31</sup> "keep your faith." Though seemingly ironic given his background of persecution of the two major religious forces in England, his crest actually bespeaks his mindset quite accurately. He kept his faith—not in any external or spiritual force, but in his own ability to navigate the treacherous currents of an ever-shifting court dynamic. As he demonstrated repeatedly, Rich was willing to do whatever he deemed necessary to survive in the cutthroat world of Tudor politics.

## Bibliography

Berry, Boyd M. "Review of *The Examinations of Anne Askew*, edited by Elaine V. Beilin." *The Journal of English and Germanic Philology* 98 (April 1999): 260-64. http://www.jstor.org/stable/27711802 (accessed April 16, 2016).

Berry, Edward. "Thomas More and the Legal Imagination." *Studies in Philology* 106, no. 3 (Summer 2009): 316–40. http://www.jstor.org/ stable/25656018 (accessed January 28, 2016).

British History Online. "Spain: June 1536, 6-10." In *Calendar of State Papers, Spain, Volume 5 Part 2, 1536-1538*, ed. Pascual de Gayangos. London: Her Majesty's Stationery Office, 1888. 137-162. http://www.british-history.ac.uk/cal-state-papers/spain/vol5/no2/pp137-162 (accessed April 15, 2016).

Derrett, J. Duncan M. "The Trial of Sir Thomas More." *The English Historical Review* 79, no. 312 (July 1964): 449-77. http://www.jstor. org/stable/560988 (accessed November 21, 2015).

Dixon, Richard Watson. *History of the Church of England*. 2nd ed. 3 vols. London: George Routledge, 1983. https://archive.org/stream/ historyofchurcho03dixo#page/n9/mode/2up (accessed May 23, 2016).

<sup>&</sup>lt;sup>29</sup> Ibid., 126.

<sup>&</sup>lt;sup>30</sup> Richard Watson Dixon, *History of the Church of England*, 2nd ed. (London: George Routledge, 1893), 3, 211. https://archive.org/stream/historyofchurcho03dixo#page/n9/mode/2up (accessed May 23, 2016).

<sup>&</sup>lt;sup>31</sup> Sargeaunt, *Felsted School*, title page.

234 Letters-Humanities, Philosophy, and Foreign Languages

Froude, James Anthony. *History of England from the Fall of Wolsey to the Death of Elizabeth*. Vol. 4, 3rd ed. London: Longman, Green, Longman, Roberts, & Green, 1864. https://books.google.com/books?id=dNE\_AAAAcAAJ (accessed May 23, 2016).

Gairdner, James. "Askew, Anne." In *Dictionary of National Biography*, ed. Leslie Stephen, 2, 190-192. New York: MacMillan, 1885. https://archive.org/details/dictionaryofnat02stepuoft (accessed May 23, 2016).

Hall, Richard. *The Life & Death of That Renowned John Fisher Bishop of Rochester*. London, 1655. http://quod.lib.umich.edu/e/eebo/A45326.0001.001?view=toc (accessed March 7, 2016).

Pleasants, J. Hall. "The Lovelace Family and Its Connections (Continued)." *The Virginia Magazine of History and Biography* 28, no. 4 (October 1920): 375-92. http://www.jstor.org/stable/4243795 (accessed January 28, 2016).

Pollard, A. F. "Rich, Richard." In *Dictionary of National Biography*, ed. Sidney Lee, 123-26. New York: MacMillan, 1896. https://archive. org/details/dictionaryofnati48stepuoft (accessed April 9, 2016).

Roper, William and Nicholas Harpsfield. *Lives of Saint Thomas More*. Ed. E. E. Reynolds. London: Aldine Press, 1963. https://archive.org/ details/livesofsaintthom000169mbp (accessed February 10, 2016).

Sargeaunt, John. A History of Felsted School with Some Account of the Founder and His Descendants. Chelmsford: Edmund Durrant, 1889.

# **'Where Has Nabonidus Gone?** Where Can He Be?' A Synthesis of the Nabonidus Controversy

Spencer C. Woolley University of Utah

### Abstract

While the notions of kingship run deep in Mesopotamian history, one name stands out as unique from other monarchs—Nabonidus, the last king of Babylon. This paper examines the different steps that Nabonidus danced upon the well-trod floor of sacral regnality. On the one hand, Nabonidus fell into a similar pattern as his famed male ancestors, a sacred king ruling over a major cosmopolitan center. On the other hand, Nabondius took his kingly sanctity far more seriously and far more personally, as a gift bequeathed to him by his priestess mother. The archaeological evidence regarding the habitations of Nabondius, combined with the well-understood textual record, synthesize a new interpretation that combines the standard reconstruction with an understanding of Nabonidus as an itinerant ruler. "Nabonidus, the great king, the strong king, the king of Babylon, the king of the four corners"<sup>1</sup> eludes easy comprehension. According to both his own words and the denunciations of his enemies, he discarded the traditions of his kingly forebears and tread his own path.<sup>2</sup> But what his own path was, and where it led, and why he departed from precedent remains a puzzle. Nabonidus slips into and out of Mesopotamian history like a Euphrates carp, wriggly and difficult to pin down. This paper argues that while the archaeological and documentary evidence do not wholly support the standard reconstruction of Nabonidus's reign, the standard interpretation cannot be cast away. Nabonidus does not dwell at Babylon, but the evidence remains weak that he dwelt at Tayma. If he does reside at Tayma for the decade that the sources state, he does not reside in the usual fashion of a Mesopotamian potentate.

The standard reconstruction of the reign of Nabonidus, drawn from the three aforementioned ancient sources and others, has him taking the throne in 555-556 BCE.<sup>3</sup> Immediately thereafter, he diverts tribute and resources away from the city of Babylon and the temple of Marduk to the sanctuary of the moon god Sin, at Harran, called Ekhulkhul.4 He "both restored and contributed to the main temples of the principal cities, including the temples of Sin, Shamash, Bunene, and Anunit at Larsa, Sippar and Nippur."5 He installs his daughter as high priestess in the city of Ur,<sup>6</sup> and travels about Babylonia, renovating temples and digging up earlier inscriptions and statues. Many scholars see Nabonidus as the first archaeologist, as he not only rebuilds temples, but also installs the previous artifacts in a museum at Babylon.<sup>7</sup> While the Neo-Assyrian kings and Nebuchadnezzar rebuilt and rededicated temples in a similar manner, they did not seem to have a concern over preserving what came before. This is but the first instance of Nabonidus departing from the ways of his forebears.

Most significantly, the standard reconstruction views Nabonidus as absent from Babylon for most of his reign. As previously noted, he flits about Sumer, Babylonia, and Akkad, but he also spends a lot of time at Tayma, an oasis northwest of the Arabian Desert. Ostensibly,

<sup>&</sup>lt;sup>1</sup> The Nabonidus Cylinder from Sippar, taken from "The Nabonidus Cylinder from Sippar", http://www.livius.org/na-nd/nabonidus/cylinder.html, retrieved 11/15/14, last modified 1/28/07. Lines 5-7

<sup>&</sup>lt;sup>2</sup> Cyrus Cylinder, lines 1-8

<sup>&</sup>lt;sup>3</sup> D.J. Wiseman. "Nabonidus." in *The Cambridge Ancient History*. ed. John Boardman et.al. Cambridge: Cambridge University Press: 1991, 246

<sup>&</sup>lt;sup>4</sup> Wiseman, 244

<sup>&</sup>lt;sup>5</sup> Ibid.

<sup>6</sup> Ibid.

<sup>&</sup>lt;sup>7</sup> Wiseman, 245

Nabonidus lives at Tayma both to enhance a temple of Sin and to oversee the campaigns against the Arabs near the eastern coast of the Red Sea.8 Sin's temple at Tayma stood second in importance to his house in Harran, and as a devotee of the moon good, it makes sense for Nabonidus to restore this temple as well. However, Nabonidus breaks important monarchical and religious precedence by not celebrating the New Year's festival in Babylon, for ten years. Rather he delegates this task to his son Belshezzar. This neglect of his duties as king causes a rebellion in the cities, to the point where they refuse to aid any further the reconstruction of Sin's temple at Harran.9 Eventually, Nabonidus does return to Babylon, where the discontented population soon replaces him as king with Cyrus the Persian. Precisely how peaceful Cyrus's takeover was, and whether Nabonidus survived to live in exile or whether he was killed in the skirmishes for the city, remains disputed.<sup>10</sup> Nevertheless, Babylon becomes part of the Persian Empire, ruled by Achaemenid governors until Alexander the Great captures the city.

So goes the standard reconstruction. As it draws primarily on the documentary sources, it matches them in great detail. But the tale of history is not to be found in written words alone, but also in the material remains left by kings and priests, peasants, and slaves. Archaeology jams the tutelary spade of Marduk into the wheel of this interpretation, leaving serious gaps in its neatly spun story. But before material culture can be considered, the issues with the sources themselves must be examined. The weavers of the standard reconstruction have strung their looms with rotten threads.

Nabonidus speaks for himself in his inscriptions found at Harran, in southeastern Turkey. In 1956, Dr. D. S. Rice discovered Babylonian steles, turned facedown and used as paving stones in the mosque.<sup>11</sup> The steles have inscriptions written both by Nabonidus and by his mother, Adad-guppi.<sup>12</sup> Steles, curved tablets of carved stone, appear all over the Near East, set up by various Neo-Assyrian and Babylonian kings as announcements of their greatness and of their accomplishment of great deeds. Neo-Assyrian steles frequently recorded martial victories, while

<sup>&</sup>lt;sup>8</sup> Ricardo Eichmann, Hanspeter Schaudig, and Arnulf Hausleiter. "Archaeology and Epigraphy at Tayma (Saudi Arabia)" *Arabian Archaeology and Epigraphy* 17 (2006): 163

<sup>&</sup>lt;sup>9</sup> Wiseman, 246

<sup>&</sup>lt;sup>10</sup> Wiseman, 247

<sup>&</sup>lt;sup>11</sup> The Harran Inscriptions of Nebonidus and his Mother, designated H1.A, H1.B, H2.A, and H2.B Take from "Ancient Near East: Harran", http://www.bibliotecapleyades.net/sitchin/Adda\_Guppi\_Harran.htm, retrieved 1/14/15. Introductory material.
<sup>12</sup> Ibid.

the steles of Nebopolssar and Nebuchadnezzar focus on their achievements as builders. While not quite on the level of twentieth-century political propaganda, steles demonstrated the power and wealth of the king, as elaborate stone carvings cost much in effort and treasure. Such steles could be read by the literate and read to the non-literate populace. But their very presence belied the power of their founders. The steles spoke for themselves as intricately carved, decorated stones, the inscriptions being embellishments for the literate passing thereupon.

Adad-guppi introduces herself right away as a "votaress of the gods Sin, Nin-gal, Nusku and Sadarnunna, my deities who, from my childhood, have sought after their godheads."<sup>13</sup> Adad-guppi clearly did not fall asleep in the Babylonian equivalent of catechism or Sunday school. Nabonidus does slumber, but in slumbering, he hears the voice of the god: "in the night season he caused me to behold a dream (say-ing) thus 'E-hul-hul the temple of Sin which is in Harran quickly build."<sup>14</sup> Nabonidus and his mother obsess over the worship of Sin and over the restoration of his temples. The Harran stele inscriptions have them doing little else but attending to his worship. Nabonidus says that "the 'head' E-hul-hul I supported, I brought pleasure to the hearts of its people, I accomplished the command of Sin, king of the gods, lord of lords."<sup>15</sup> Adad-guppi intones, "The word of Sin, king of the gods, which he spoke to me I honored, and I myself saw (it fulfilled)."<sup>16</sup>

Of the ancient sources, the inscriptions at Harran are the least problematic. They speak in the voice of Nabonidus and his mother, and they support the standard narrative that Nabonidus and Adad-guppi held Sin in a position superior to the other gods.<sup>17</sup> The steles date from Nabonidus's reign, and the textual evidence strongly indicates that they were set up at his royal command. To the *barru*-priests in Babylon, such maniacal devotion to a few minor members of the pantheon, at the expense of the prestige of the god of the capital, would come as grave insult. That insult applied to the god, to his priests, and to the city devoted to his worship. The steles also indicate that Nabonidus goes to "the lands of the Arabs"<sup>18</sup> at the command of Sin and remains there for ten years.<sup>19</sup> The Babylonian Chronicle places Nabonidus in Tayma as

<sup>&</sup>lt;sup>13</sup> Harran Inscriptions, Nabonidus H1.B Column 1, 3-6

<sup>&</sup>lt;sup>14</sup> Harran Inscriptions, Nabonidus H2.A Column 1 11-12

<sup>&</sup>lt;sup>15</sup> Harran Inscriptions, Nabonidus H2.B Column 3 27-29

<sup>&</sup>lt;sup>16</sup> Harran Inscriptions, Nabonidus H1.A a Column 2 11-12

<sup>&</sup>lt;sup>17</sup>The inscriptions also make reference to Nin-gal, Nusku and Sadarnunna, but in much lesser frequency than to Sin himself.

<sup>&</sup>lt;sup>18</sup> Harran Inscriptions, Nabonidus H2.A Column 1 43

<sup>&</sup>lt;sup>19</sup> Harran Inscriptions, Nabonidus H2.A Column 2 11

well; confirmation from two sources cannot be ignored.<sup>20</sup> But the Harran inscriptions lack precision on what Nabonidus does at Tayma, besides various forms of honoring Sin and subjugating the local Bedouins. Nabonidus may have dwelt at Tayma; but what did he do there?

The Babylonian Chronicle seems to focus much more about what Nabonidus did not do, rather than what he did; the Chronicle's authors find more significance in Nabonidus's absences than in his achievements. To be sure, "the king called up his army and marched against the country Hume,"<sup>21</sup> and he recovers from an illness,<sup>22</sup> but the key phrase, repeated at least five times,<sup>23</sup> is "the king stayed in Tayma."<sup>24</sup> More significantly, "the crown prince, his officials and his army were in Akkad."25 It would be quite difficult to conquer Arabs, even if they are camel-bound Bedouins, if the army stays home in Babylonia. This then contradicts the Harran inscriptions- Nabonidus, personally powerful as he might have been, needed an army to conquer territory.<sup>26</sup> Equally important to the barru-priests who wrote the Babylonian Chronicle would be that "the king did not come to Babylon for the [New Year's] ceremonies of the month of Nisannu."27 The New Year's festival was the pinnacle of the year, when Marduk would inhabit the king's body so that the proper rituals and propitiations could be made. By his absence at the New Year's festival, Nabonidus denied Marduk his homage at the head of the Babylonian pantheon. He also managed to offend all of Marduk's priests, and their clients, the citizens of Babylon itself. It should be noted that the veneration of Marduk came with numerous gifts, tribute, taxes, and baksheesh. Nabonidus diverts the offerings from Marduk to Sin, which would have an effect on the priests' and the city's economic health. Moreover, the usual god-to-god visits in which the Babylonian pantheon engaged are curtailed by the ninth year of his reign.<sup>28</sup> The Chronicle remains silent as to the specific reasons, whether at the command of Nabonidus or because of turmoil in the kingdom caused by the Persians. Either way, the keepers of the

<sup>&</sup>lt;sup>20</sup> A. Leo Oppenheim, "The Babylonian Chronicle." in *Ancient Near Eastern Texts Relating to the Old Testament*. Princeton: Princeton University Press, 1950, 4

<sup>&</sup>lt;sup>21</sup> Oppenheim, "The Babylonian Chronicle," 2

<sup>&</sup>lt;sup>22</sup> Ibid

<sup>&</sup>lt;sup>23</sup> And probably more with the lacuna.

<sup>&</sup>lt;sup>24</sup> Oppenheim, "The Babylonian Chronicle," 3

<sup>25</sup> Ibid.

<sup>&</sup>lt;sup>26</sup> Harran Inscriptions, Nabonidus H2.A Column 1 43

<sup>&</sup>lt;sup>27</sup> Oppenheim, "The Babylonian Chronicle," 2

<sup>&</sup>lt;sup>28</sup> Ibid.

Babylonian Chronicle have ample religious, economic, and political reasons for disliking the Chaldean theoretically occupying the throne.

The Babylonian Chronicle keepers served as the omen-reading barru-priests of Marduk. The Chronicle was vital in their ability to interpret signs and wonders. As a chronographic text, it indicated precisely which events had happened when and where, so that precedent could be consulted for current marvels and signs. It was first and foremost a reference book, a list of the names of the years and the notable events that occurred therein. It certainly was distributed for public consumption, possibly for reading at religious festivals and national holidays, but that was not its primary function. Much of the Chronicle's reliability concerning Nabonidus depends on whether the accounts are contemporaneous to him or if they have been redacted at a later day. For most of its history, the Chronicle seems to be composed year by year, with each year matching certain identifiable patterns while describing different events. The Chronicle clearly sees Cyrus as a superior ruler: "Cyrus entered Babylon, [unidentified objects] were filled before him— the state of peace was imposed on the city."29 The Chronicle further emphasizes that the Persians stayed out of the sacred precincts and that the ceremonies continue uninterrupted.<sup>30</sup> This is directly contrasted with Nabonidus's neglect of the worship of Marduk, and of his cowardice; he flees from Sippar and gets arrested in Babylon.<sup>31</sup> Because of their clear antipathy toward them, the barru-priests may well have distorted Nabonidus's reign. One possible instance has the army remain in Babylonia, rather than going to Arabia. But by virtue of their role as interpreting the signs and omens, they work under a strict commission to tell the truth, otherwise the religious and chronographic value of the Chronicle disappears. It is possible that Nabonidus's actions may have been so reprehensible, and Cyrus's victory so total, the barru-priests may have felt justified in adjusting their record to make the former king seem weaker and stranger.

If the Babylonian Chronicle seems prejudiced against Nabonidus, the Cyrus Cylinder attacks him with the force of a battering ram. Cyrus replaces him; the Persian king has no need to make Nabonidus look good, and has every reason to portray him as an immoral, bumbling fool:

> An incompetent person was installed to exercise lordship over his country, and [...] he imposed upon them. A counterfeit of

31 Ibid.

<sup>&</sup>lt;sup>29</sup> Oppenheim, "The Babylonian Chronicle," 4

<sup>&</sup>lt;sup>30</sup> Ibid.

Esagila he made ... for Ur and the rest of the cultic centers, a ritual which was improper to them, an [unholy] di[splay offering x x x without] fear he daily recited. Irreverently, he put an end to the regular offerings and he interfered in the cultic centers. By his own plan, he did away with the worship of Marduk, the king of the gods. <sup>32</sup>

Cyrus then goes further; not only has Nabonidus lost the mandate of heaven to rule, Marduk has bestowed it upon him, the worthy foreign adventurer: "he [Marduk] searched everywhere, and then he took a righteous king, his favorite, by the hand, he called out his name: Cyrus, king of Ansan, he pronounced his name to be king all over the world."<sup>33</sup> Cyrus goes to Babylon with Marduk "like a companion and a friend"<sup>34</sup> beside him and enters the city peacefully as a liberating hero.

Cyrus composes his Cylinder from a position of great strength, as the victorious hero of both gods and men. The Cylinder itself may have been placed in a temple foundation, although the text of the Cylinder would have been disseminated over all the former Babylonian Empire, read by the literate, and read to the uneducated. Cyrus picks the divine sovereignty that Nabonidus drops during his desert sojourn. The Cylinder text either ignores as irrelevant or distorts as blasphemy Nabonidus's efforts to conquer Arabia. Nabonidus loses the approval of Marduk through both his incompetence and his ill handling of religious ritual.<sup>35</sup> While other things do occur during the reign of Nabonidus that may have inflamed the urbanites or invited foreign occupation, Cyrus chooses to make his argument for supremacy an entirely religious one. Of the three ancient documentary sources, the Cyrus Cylinder must be relied upon the least. Cyrus wants to keep ruling, and he uses the religious arguments in the Cylinder to position himself as the most divinely appropriate one to do so. Moreover, the archaeological record at Babylon indicates remnants of a fairly nonviolent siege, but a siege nonetheless, including a siege wall. Given that siege wall, it does not seem likely that Cyrus rode into Babylon on the back of an ass with palm branches waving. If he did so, his troops may have been shaking the arms that held the palms.

<sup>&</sup>lt;sup>32</sup> The Cyrus Cylinder, Fragment A. Taken from "Cyrus Cylinder (2)," http://www. livius.org/ct-cz/cyrus\_I/cyrus\_cylinder2.html, retrieved 1/14/15, last modified 2/5/10. Lines 2-7.

<sup>&</sup>lt;sup>33</sup> Cylinder, line 12

<sup>&</sup>lt;sup>34</sup> Cylinder, line 15

<sup>&</sup>lt;sup>35</sup> Cyrus Cylinder, lines 2-7

Archaeology highlights the discrepancies in the Cyrus Cylinder and also indicates what was going on at Tayma. The site of Tayma lies on the main north-south trade route between the incense groves of Arabia Felix and the ports of the eastern Mediterranean. It grew quite wealthy from the aromatic gum trade, but the site lacked the water necessary for urban expansion. In addition, the surrounding tribes could waylay the caravans and attack whatever settlement the oasis possessed. Tayma was a frontier fort and a way station, much like Fort Bridger and Fort Hall were during the Westward Expansion period in the United States. What Tayma does not seem to be is the luxurious residence of a Babylonian monarch.

Strengthening the credibility of both the Harran inscriptions and the Babylonian Chronicle, a temple to Sin at Tayma exists and was uncovered by the German archaeological team.<sup>36</sup> In comparison to the grand sanctuary at Harran, the temple at Tayma seems a mediocre shrine, at best. Yet there is also a stele, contemporaneous to Nabonidus's reign that declares he set it up and that he dedicated various offerings of gold and lapis lazuli to Sin and his divine groupies.<sup>37</sup> What it does not say is whether Nabonidus dwelt at Tayma or how long he stayed there. A bored guard provided some graffiti stating that he was sent by Nabonidus to Tayma, but again, no information exists in the archaeological record that Nabonidus lived in and around Tayma nor that he wars against the native Arab tribes. To be sure, Tayma has been occupied for millennia, including after the reign of Nabonidus. If Nabonidus did build a royal residence, it may yet be unexcavated or be under inaccessible modern buildings or sites.38 But the lack of any evidence of Nabonidus ten-year odyssey in the desert is telling. Where did Nabonidus go? Where could he have been, and what was he doing?

The documentary sources agree that: First, Nabonidus behaved in ways unusual, unexpected, and religiously divergent from what his subjects expected and wanted. Second, Nabonidus does not spend much time in Babylon, except at the very beginning and ending of his reign. Two of the sources, the Harran inscriptions and the Babylonian Chronicle, have him spend ten years in Tayma. The Cyrus Cylinder argues for his heresy and does not mention his location. Third, in both the Babylonian Chronicle and the Cyrus Cylinder, Cyrus justifies his usurpation of Nabonidus's throne on religious grounds; Marduk told him to do it. In addition to the above points, the archaeological record at Tayma gives one mediocre temple of Sin, one stele of Nabonidus, and one line

38 Ibid.

 <sup>&</sup>lt;sup>36</sup> Eichmann et al., "Archaeology and Epigraphy at Tayma," 163.
 <sup>37</sup> Eichmann et al., "Archaeology and Epigraphy at Tayma," 170.

of the Babylonian equivalent of 'Kilroy was here.' Any formulation of a narrative concerning the life and times of Nabonidus must take into account all of the above information. Although the standard reconstruction embraces much of the above, it ignores critical biases and inconsistencies in the documentary sources and does not address at all the findings at Tayma. To be fair, the Tayma excavations are only ten years old. Nevertheless, a more accurate telling of the reign of Nabonidus could read as follows:

Nabonidus becomes king of Babylon in 556-556 BCE. His mother taught him to honor the moon god Sin above all others. Upon assuming the kingship, he begins to divert resources away from the tradition urbanite deities and towards Sin. He begins to reconstruct the sanctuary of Sin in Harran and also moves around Sumer and Akkad, visiting temples, excavating their older remains and preserving the interesting tidbits. Rather than stay in Babylon and celebrate the native urban festivals, he goes across his kingdom in a Royal Progress, honoring diverse gods. He may frequently visit Tayma, restore the temple there, post some guards to ward off the caravan raiders and put up a stele, but it is unclear whether he lives there on a permanent or semi-permanent basis. Perhaps he contemplates the glories of the moon underneath a date palm. Regardless of what he does or does not do at Tayma, his lack of residence in Marduk's sacred city enrages the populace to the point that when a foreign invader appears, they offer but token resistance before delivering the city into the hands of the Persians.

The key to the above narrative is the theme of Nabonidus-aswandering-religious-mystic. This ambulatory mystic theme brings together the point where all three of the ancient documentary sources agree, that Nabonidus rebuilds many temples and honors many gods. The Babylonian Chronicle indicates that in the seventh year of Nabonidus's reign, "The image of the god Nabu did come to Babylon, the image of the god Bel did not go out of Esagila in procession."<sup>39</sup> The standard reconstruction suggests that this lack of movement may be due to Persian raiding, but Cyrus does not threaten Babylon for three more years.<sup>40</sup> The wandering king may be having a party with Bel and Nabu in their own temples, rather than at Babylon itself. The Harran inscriptions list the gods Samas, Nusku, Istar, Adda, Nergal, Nin-gal, and Sadarnunna as being worthy of worship at the command of Sin.<sup>41</sup> These gods have their own temples in their own cities, in addition to their townhouses in Babylon. Finally, the Cyrus Cylinder indicates that

<sup>&</sup>lt;sup>39</sup> Oppenheim, "The Babylonian Chronicle," 2

<sup>&</sup>lt;sup>40</sup> Oppenheim, "The Babylonian Chronicle," 3

<sup>&</sup>lt;sup>41</sup> Harran Inscriptions H2.B Column 3 line 31

Nabonidus changed cultic worship from its traditional form into something new, something that angered Marduk and probably angered Marduk's priest. Cyrus could play off of that anger as a justification for taking the city. Nabonidus attempted to change the state religion. That much is clear. But whether he did so by dwelling at Tayma remains uncertain.

The evidence at Tayma stands currently inconclusive. A valid syllogism: the absence of evidence is not evidence of absence. But in either case an absence exists, needing filling with educated guesses and reasonable estimates. Based on current knowledge, Nabonidus did not live at Tayma in a similar manner to other Mesopotamian kings. He may have had a cot in the temple of Sin located in Tayma. He may have built a reed hut like his Chaldean ancestors, either using local reeds or importing them from the swamps. He may have adopted Bedouin practice and lived in a tent. He could have eaten figs under a palm tree—any definitive placement of a long-term residence for Nabonidus at Tayma is wholly speculative. If a yet-undiscovered palace of Nabonidus appears out of the sand at Tayma, then this narrative will need to be revised. But the most likely scenario, the scenario that fits all of the evidence, sees Nabonidus, king of Babylon, as a pilgrim, going from shrine to shrine and temple to temple, ignoring the state cult of Marduk in Babylon for the less well-known and perhaps unrecognized gods that Sin tells him to venerate. His distinctive worship and honor to the others of the Babylonian pantheon would bring the wrath of the city and its *barru*-priests upon him, so that when Cyrus descends from the Zagros Mountains, it is Cyrus who becomes the blessed of Marduk, rather than the Chaldean who seems to have preferred the solitude of the desert to the bustle of the big city.

## Bibliography

#### Ancient Sources

Oppenheim, A. Leo. "The Babylonian Chronicle." in *Ancient Near Eastern Texts Relating to the Old Testament.* ed. James B. Pritchard. Princeton: 1950.

The Harran Inscriptions of Nebonidus and his Mother, designated H1.A, H1.B, H2.A and H2.B. Taken from "Ancient Near East: Harran," http://www.bibliotecapleyades.net/sitchin/Adda\_Guppi\_ Harran.htm, retrieved 1/14/15.

The Cyrus Cylinder, Fragment A, lines 1-19. Taken from "Cyrus Cylinder (2)," http://www.livius.org/ct-cz/cyrus\_I/cyrus\_cylinder2. html, retrieved 1/14/15, last modified 2/5/10.

The Nabonidus Cylinder from Sippar, taken from "The Nabonidus Cylinder from Sippar," http://www.livius.org/na-nd/nabonidus/ cylinder.html, retrieved 11/15/14, last modified 1/28/07. Lines 1-7.

#### Modern Sources

Eichmann, Ricardo, Hanspeter Schaudig and Arnulf Hausleiter. "Archaeology and Epigraphy at Tayma (Saudi Arabia)." *Arabian Archaeology and Epigraphy* 17 (2006): 163-176.

Wiseman, D.J. "Nabonidus." in *The Cambridge Ancient History*. ed. John Boardman, I.ES. Edwards, N.G.L. Hammond, E. Sollberger, and C.B.F. Walker. Cambridge: Cambridge University Press, 1991. pg. 243-251.

# "Thy Placeless Power": Melville, Mobility, and the Poetics of Placelessness

### **Todd Goddard**

Utah Valley University

### Abstract

This paper investigates Melville's unease with the erosion and absence of abiding places, which in turn are linked to proliferating spatial mobilities in the first half of the nineteenth century. Although critics have pointed to Melville's celebration of travel and its attendant freedom of movement, few have commented on his concern for the consequences of mobility on the integrity of place or the resultant implications of placelessness for identity and authorship alike. What we see in Melville's works, I argue, and particularly in Moby-Dick, is a place-anxiety or place-panic that derives in part from the accelerating velocities of modernity and a sense of the increasing loss of stable, bordered, and bounded places. Melville's is thus an early voice antedating current cultural debates and ecocritical concerns over the role of place, the synergies and antagonisms between the local and the global, and the tensions between the rooted and the mobile. Melville's novels describe the consequences of the velocities of modern life and envision how its ever-increasing matrix of non-places may be navigated.

Writing for *Harper's New Monthly Magazine* in June 1865, author Robert Tomes ascribed a uniquely mobile temperament to Americans: "The American is a migratory animal. He changes places with such facility that he never seems so much at home as when leaving it" (57). Tomes naturalizes movement here as an essential characteristic of national identity, with places serving as way-stations or fixed nodes in networks of traffic. In this formulation, the idealized noun "home" is set aside for the predicate verb phrase. The American is located or positioned as most "at home" in the temporal act of departure from it. Tomes continues:

With miles of railroad and length of navigable river more than those of all the rest of the world together, a citizen of our vast republic passes with ease and rapidity from the Lakes to the Gulf of Mexico, or from the Atlantic to the Pacific. Thus readily moving over a great continent, within the limits of his own country, he becomes almost unconscious of space, and so habituated to travel that he thinks no more of counting the hundreds of miles of his frequent journeys, by railway and steamer, than the steps of his daily walk. (57)

Here travel becomes less an option than a necessity: space demands movement for the American, and movement becomes a necessary condition of being American. The Great Lakes and the Gulf of Mexico, like the vast and physically differentiated landscapes between New York and California, become little more than a medium through which travelers pass, as particularity gives way to a universal and everywhere-similar medium for movement. The generic traveler, like the landscape itself, also becomes abstracted. Unconscious of space and always *en route*, bodily steps are replaced by the passivity of mechanized travel, the convenience of ample facilities for comfort and consumption, and the ease of speed over vast distances. Tomes's screed on American movement appears almost 15 years after the appearance of Herman Melville's *Moby-Dick*, yet the vast changes wrought on the young nation during the so-called "Great Acceleration" were equally dizzying in the 1850s.

This paper investigates Herman Melville's unease with the erosion and absence of abiding places, which in turn are linked to proliferating spatial mobilities in the first half of the nineteenth century. While

critics have pointed to Melville's celebration of travel and its attendant freedom of movement, few have commented on his concern for the consequences of mobility on the integrity of place or the resultant implications of placelessness for identity and authorship alike. What we see in Melville's works, I argue, and particularly in Moby-Dick, is a place-anxiety or place-panic that derives in part from the accelerating velocities of modernity and a sense of the increasing loss of stable, bordered, and bounded places.<sup>1</sup> In extreme forms, such placelessness figures forth in his works as the void, a boundless and eternal "Nothing" akin to infinity and zero wherein all distinction of identity and place disappear. Melville registers other permutations in the loss of rooted and delineated places, in the slide away from centered particularity toward decentered universality, in natural erosion and ecological decay, and in the rapid building up and wearing out of place. While we can see Melville wrestling with these ideas throughout his works, his most pronounced response comes in Moby-Dick, where he struggles with the metaphysics of placelessness and the consequences of placelessness on identity. In this response, Melville imagines an internalized topography of the self that manages to retain the integrity of distinct borders, if only within the self, and of heterogeneous identity. Melville therein challenges the prevailing myth of progress in the nineteenth century as tethered to and defined by mobility, and he simultaneously proffers the chance for a renewal of distinctiveness and stability.

### "Thy Placeless Power"

In many of Melville's novels, the concern we see with mobility has less to do with the act of incessant movement *per se* than with the consequent instability or diminishment of the stopping place. For Melville, any vision of progress linked to mobility carried serious implications for place. The vertiginous mobility erupting about him at midcentury, he shows, exacerbated the disappearance of richly varied and distinct experiences of place and the deterioration of bordered, bounded, sound, and solid places. In making such connections, Melville anticipates the contemporary voices that connect the proliferation of placelessness in the modern world with incessant movement and

<sup>&</sup>lt;sup>1</sup> My use of the terms "place-anxiety" and "place-panic" is informed by the work of philosopher Edward Casey, particularly as those terms are deployed in *Getting Back into Place: Toward A Renewed Understanding of the Place-World*, (Indiana University Press: Bloomington, 2009). For Casey, we are apt to experience place-anxiety or place-panic when "we confront the imminent possibility of there being no place to be or to go. We feel not so much displaced as without place." (ix-x). An extreme example, he suggests, is the horror of the nuclear annihilation of places.

speed. As geographer Nigel Thrift points out, "In a world of everintensifying mobility it becomes increasingly difficult to imagine space as bounded," given that localities lose their spatial individuality and autonomy to become diminished points in the circulation of traffic (222).<sup>2</sup> Thrift's description of the transformation of place into "almost/not quite spaces"-that is, into nodes that channel, serve, or direct accelerating flows, vectors, and velocities-resonates with Robert Tomes's observations of the American who readily sheds places and feels most "at home" in movement (222). In both instances, place (even home) is reduced to the status of a kind of way-station in networks of movement. For Melville, the reduction of place to a mere resting space includes the increasing loss of stability and integrity. Indeed, he foreshadows the (quite Melvillian) observation by Deleuze and Guattari that "Between things does not designate a localizable relation going from one thing to the other and back again, but a perpendicular direction, a transversal movement that sweeps one and the other away, a stream without beginning or end that undermines its banks and picks up speed in the middle" (25). In Deleuze and Guattari's figuration, the "banks" stand in for place and the "stream" for the ever-accelerating mobility fueled by the erosion of place. Melville employs similar images of flowing water and eroding soil as metaphors for the erosive relationship between mobility and place in Moby-Dick, where he develops his expression of placelessness most fully.

The ocean provides Melville with perhaps the best location to contemplate place in its most elemental forms. It compels him to ponder questions about the importance and absence of place and ultimately the relation between place and the divine. In *Moby-Dick*, the sublime immensity of the ocean conjures a sense in Ishmael of the infinite: "In landlessness alone resides the highest truth, shoreless, indefinite as God" (97). To go to sea brings with it the chance to trace the world back to its prehistoric origins, back to the very creation of the world. Upon descending into its "unwarped primal world," Pip discovers "God-omnipresent," while on the ocean floor he finds "God's foot upon the treadle of the loom" (347). From out of the ocean's "firmament of waters," God weaves the earth's places into existence (347). Melville's

<sup>&</sup>lt;sup>2</sup> For an interesting discussion of the relation between mobility and place, see also anthropologist Marc Augé's *Non-Places: An Introduction to Supermodernity*, trans. Stuart Gilbert (London: Verso, 2008), vii-xxii. Augé's study examines "non-places," or the malls, motorways, and airports that comprise what he terms supermodernity. For Augé, non-places are "unprecedented extension[s] of spaces of circulation, consumption and communication corresponding to the phenomenon we identify as 'globalization.''' For an extended treatise on the fate of place in the modern world, see Edward Casey's *Getting Back Into Place: Toward a Renewed Understanding of the Place-World*.

use of the word "firmament," moreover, references what is arguably the first act of creation (the separation of earth and sky) and in turn the first creation of place. In the Judeo-Christian tradition, central to Moby-Dick's symbolic architecture, Genesis begins with God's establishment of the horizon, distinguishing light from darkness and sea from sky: "And God said, 'Let there be a firmament in the midst of the waters, and let it divide the waters from the waters.' And God made the firmament, and divided the waters which were under the firmament from the waters which were above the firmament; and it was so."3 As Melville recognizes, the creation of the horizon in Genesis provides the first act of spatial distinction.<sup>4</sup> Philosopher Edward Casey points out the importance of such distinction in his landmark study on place, where he explains that we need the horizon if the world is to retain its identity. Without it, Casey argues, we would be lost in "a primal mist of indifferentiation, a perceptual morass, a 'slush' of indetermination" (The Fate of Place 11). Thus, the ability to posit oneself in one place as opposed to another begins with God's separation of sea from sky. In this cosmological framework, Melville suggests, God and emplacement are inextricable. But just as the presence of God can explain the creation of distinct and differentiated places, God's absence can also engender a terrible sense of placelessness.

If the source of Melville's unease with mobility lay in the diminishment of place, his place-anxiety or place-panic derives from the possibility of total placelessness or *atopia*, with the threat of the "void" at the extreme end of the placeless spectrum. The fear of an infinite, vacant expanse of space haunts Melville's *Moby-Dick* and helps to explain Melville's personal fear of the loss of distinct and delineated places. In this loss, Melville sees a frightening slide away from particularity toward decentered universality, from bounded place to infinite space, and from individuality to the diffusion (and possible destruction)

<sup>&</sup>lt;sup>3</sup> Gen. 1: 1-7. The issue of place inevitably comes up in most creation myths. The narratives of cosmic creation often must involve movement along some temporal sequence (i.e., Before/After) from placelessness to place-filled creation. The act of creation becomes increasingly complex when we consider that the initial creative act requires some place from which to occur. In other words, the act of creation must have its own place. For a fascinating discussion of the role of place in creation narratives, see Edward Casey, *The Fate of Place: A Philosophical History* (Berkeley: University of California Press, 1998), 3-22.

<sup>&</sup>lt;sup>4</sup> Gen. 1:2 retells the story of creation in ways that perhaps more robustly emphasize place. Here God forms man from "the dust of the ground, and breathed into his nostrils the breath of life; and man became a living soul" in an act of fusion that arguably melds the physical properties of place with the spiritual. God also creates unique, geographically oriented, and purposed places: "And the Lord God planted a garden eastward in Eden; and there he put the man whom he had formed."

of the self. The ideas of universalism and infinite space were only tolerable for Melville if in fact a divine spirit infused the universe. One might forfeit particularity (of self, of place) with some comfort at the prospect that God's essence infuses the empty spaces of the universe, as Christian scholars believed. The claim by Andrew Delbanco, that Melville could never escape the suspicion that "his feelings for divine immanence were illusory-and at the next moment, the luminous world fades away and he finds himself in chill and darkness" (282), echoes Nathaniel Hawthorne's observation that Melville "can neither believe, nor be comfortable in his unbelief" (432). Ishmael's confrontations with "the heartless voids and immensities of the universe" are likewise indissociable from the question tersely posed by Starbuck, "Great God, where are you?" (Moby-Dick 422). As we see in Moby-Dick, placelessness is a tolerable notion only in a God-infused universe; in the absence of the divine, the icy vacancies of the universe inspire the deepest anxietv.

The void appears throughout Moby-Dick in various shades and permutations, yet its sharpest edge is figured forth in the ever-moving white whale, whose hieroglyphic nature and impenetrable whiteness is redolent with infinities and voids (Adams 144). The terror of the whale's "placeless power" (Moby-Dick 417) is prefigured in Melville's *Redburn* when the eponymous character gazes from the ship's rigging at night into the "black gulf [of the ocean], hemmed in, all round, by beetling black cliffs. I seemed all alone; treading the midnight clouds; and every second, expected to find myself falling-falling-falling, as I have felt when the nightmare has been upon me" (113). Redburn's vision occasions the deepest anxiety because he gazes upon the vista of the cosmic abyss. Indeed, there is no *place* to catch him from falling. The passage echoes Pascal's pronouncement uttered two centuries earlier that "[t]he eternal silence of these infinite spaces terrifies me" (201).<sup>5</sup> The whiteness of the whale, of course, elicits a similar reaction (Adams 144).<sup>6</sup> Ishmael queries,

> Is it that by [whiteness's] indefiniteness it shadows forth the heartless voids and immensities of the universe, and thus stabs us from behind with the thought of annihilation, when behold-

<sup>&</sup>lt;sup>5</sup> Pascal's fragment no. 201 reads in French as "Le silence éternal de ces espaces infinis m'effraie."

<sup>&</sup>lt;sup>6</sup> I agree with Robert Martin Adams that to assign the whale any specific signification, even the void, reduces the book's import. I also agree with his expansive assessment that the whale becomes "a multiple metaphoric window opening on Nothing." See Adams, *NIL*, 144.
ing the milky way? Or is it, that as in essence whiteness is not so much a color as the visible absence of color, and at the same time the concrete of all colors; is it for these reasons that there is such a dumb blankness, full of meaning, in a wide landscape of snows—a colorless, all-color of atheism from which we shrink? (Melville, *Moby-Dick* 169)

The threat here that so troubles Ishmael and that conjures the void is the absence of specificity or differentiation in the blankness of whiteness. If whiteness is all colors, it is simultaneously no color. To achieve the universal is to forfeit unique personality and the particulars that make up the individual and the author alike. The white whale, of course, succeeds in destroying the *Pequod*, whose crew is ripped into the void-like "yawning gulf" of the sea, with the concluding image of the vortex neatly merging mobility—mechanical, boundless, endless circular motion—with destruction (Leonard 13-25).

Subtler variations of the void appear in romantic efforts, simultaneously alluring but no less threatening, to paper it over with Pantheism. Ishmael succumbs to this temptation when, in a moment of heightened sensation brought about by squeezing the gelatinous sperm, he desires to forego his own particularity to join in with the universal: "I squeezed that sperm till I myself almost melted into it; [...] I found myself unwittingly squeezing my co-laborers' hands in it, mistaking their hands for the gentle globules [...]. Come; let us squeeze hands all around; nay, let us all squeeze ourselves into each other; let us squeeze ourselves universally into the very milk and sperm of kindness" (Moby-Dick 349). One could hardly imagine a more democratic sentiment. The moment transcends its obvious homoeroticism to evoke an orgiastic merging of selves, wherein Ishmael embraces a benevolent pantheism-akin to a kinder, gentler void-that promises brotherhood and eternal fellowship, if not union with the divine. As with the strict void, however, individual identity here is forfeited, along with the distinctiveness of material objects and place. Where the crystallized sperm ends or his fellow shipmates' hands begin becomes perceptually uncertain, akin to the "colorless all-color of atheism." The whiteness of the "very milk and sperm of kindness" serves as an optimistic rebuke to the chilly voids and vacancies shadowed forth by the whiteness of the whale.

The consequences of self-forfeiture or self-diffusion are broached again in the often-quoted chapter, "The Mast-Head." Once again the reverie seems to infuse the universe with a kind of benevolence, yet it belies the reality of physics and the ocean waters below. The youth in the masthead imagines himself dissolving into the undulating motion of the waves beneath him:

[...] lulled into such an opium-like listlissness of vacant, unconscious reverie is this absent-minded youth by the blending cadence of waves with thoughts, that at last he loses his identity; takes the mystic ocean at his feet for the visible image of that deep, blue bottomless soul, pervading mankind and nature; [...] But while this sleep, this dream is on ye, move your foot or hand an inch; slip your hold at all; and your identity comes back in horror. Over Descartian vortices you hover. And perhaps, at mid-day, in the fairest weather, with one halfthrottled shriek you drop through that transparent air in to the summer sea, no more to rise for ever. Heed it well, ye Pantheists. (*Moby-Dick* 140)

Notable here are the recurring images of repetition and duplication with merger and union: the ceaseless cadence or rhythm of the waves blends in the vacancy of his thoughts. The sentences themselves move forward softly punctuated with commas and semi-colons until the hard stop of the period, as the youth's identity returns briefly before he meets his unfortunate end. Vacancy lingers here as both text and subtext, while Descartes' vortices again suggest ceaseless movement without the promise of a final rest or a final meaning (Leonard 13-25). The threat of the void as it is figured at sea in *Mohy Dick*, appears in

The threat of the void, as it is figured at sea in Moby-Dick, appears in other permutations on the land itself, where we find it suggested in the aura of placelessness and in the restless velocities and mobilities of modernity. The void's vacancies echo in the lack of abiding places and in the erosive instability of the land, as we see in Melville's mythical representation of the island of Nantucket (Figure 1). Prior to his writing Moby-Dick, Melville had never visited the island, and its role as a major whaling port had long been superseded by New Bedford. Yet Melville's representation revives its fabled past as the primary port and profit center for the world's whale fisheries. With a short chapter devoted to it, Melville portrays its culture as one of almost pure mobility. We are told that Nantucketers exceed the imperial nations in global reach and spatial ambition: "Let America add Mexico to Texas, and pile Cuba upon Canada; let the English overswarm all India, and hang out their blazing banner from the sun; two thirds of this terraqueous globe are the Nantucketer's" (Moby-Dick 62-63). Nantucketers are all movement: they go, issue forth, "take to the sea," "overrun," wade out, push off, launch, and "conquer" (61-63). In addition to their economic pursuit of whale oil, their commitment to mobility may stem from their island's own lack of determinacy: it seems forever poised on the verge of watery disintegration. In a curious aside, Melville requests that readers take out a map and peruse the island, as if its cartographic (and literary) representation might help to stem its erosion or give the island a more stable shape or form.

> Nantucket! Take out your map and look at it. See what a real corner of the world it occupies; how it stands there, away off shore, [...] Look at it—a mere hillock, and elbow of sand; all beach, without a background. There is more sand there than you would use in twenty years as a substitute for blotting paper. Some gamesome wights will tell you that they have to plant weeds there, they don't grow naturally; that they import Canada thistles; that they have to send beyond seas for a spile to stop a leak in an oil cask; that pieces of wood in Nantucket are carried about like bits of the true cross in Rome; that people there plant toadstools before their houses, to get under the shade in summer time; that one blade of grass makes an oasis, three blades in a day's walk a prairie; that they wear quicksand shoes, something like Laplander snow-shoes; [...], that to their very chairs and tables small clams will sometimes be found adhering, as to the backs of sea turtles. (61-62)



Figure 1. Nantucket Island, Massachusetts, 1869, by F. C. Ewer. Map reproduction courtesy of the Norman B. Leventhal Map Center at the Boston Public Library.

With the absence of trees and grass, the island lacks a "background," foundation, or any sense of rootedness and adheres only to planted weeds and imported thistles—these alone allow the island to abide. A model of seashore erosion, an absorptive and permeable "anthill," Nantucket seems always at risk of washing away into an oceanic placelessness.

In Moby-Dick, Melville's response to these place-anxieties, all of which implicate the integrity of the self, is to articulate a space within the individual that is impregnable and sovereign, secure and bounded, particular and unique. In an 1851 letter to Nathaniel Hawthorne, Melville describes the ideal author as "the man who, like Russia or the British Empire, declares himself a sovereign nature (in himself) amid the powers of heaven, hell, and earth" (qtd. Moby-Dick 536). The parenthetical emplacement "(in himself)" posits a kind of bordered and secure topography within the author, as well as in the "man." Such a declaration is an act of will and an assertion of individual and authorial independence akin to the creation of nationhood. Autonomous, independent, and self-governing. Melville's ideal author-individual claims a jurisdiction of the self that is spatialized and impregnable. Even amid the "powers of heaven, hell, and earth," and thus presumably against the void itself, the individual retains distinction and is free to move about the land or "watery part of the world" secure and intact (Moby-*Dick* 18).<sup>7</sup>

At stake for Melville is nothing less than the autonomy and survival of the individual whose very being is premised on the existence of place. As Casey (via Heidegger) remind us, "To be is to be in place" (Casey, Getting Back Into Place 14). Although certainly not Melville's mouthpiece in Moby-Dick, Father Mapple and his pulpit, a "selfcontaining stronghold," suggest the kind of emplaced author-individual that Melville envisions. Ishmael observes "Father Mapple after gaining the height, slowly turn round, and stooping over the pulpit, deliberately drag up the ladder step by step, till the whole was deposited within, leaving him impregnable in his little Quebec" (Melville, Moby-Dick 43). Bounded and impregnable, Mapple's self-sovereignty and distinct identity is anchored to and defined by a stabilized geography that distinguishes him from the members of his congregation. While we might even say that his proprietary relationship with the land (his little Quebec) lends him the Lockean grounds for selfhood, his "little Quebec" provides security from the increasing placelessness of the world and the

<sup>&</sup>lt;sup>7</sup> I agree with Wai-Chee Dimmock's analysis in *Empire of Liberty* that for Melville "Authorship is almost exclusively an exercise in freedom, an attempt to proclaim the self's sovereignty over and against the world's." Wai-Chee Dimmock, *Empire of Liberty: Melville and the Poetics of Individualism* (Princeton: Princeton University Press, 1989). I also read Melville's spatialization of the self as an assertion of bounded particularity and self-sovereignty.

consequent loss of identity. A similar spatialization of the self occurs with the "*Isolatoes*" aboard the *Pequod* who do not acknowledge "the common continent of men, but each Isolato liv[es] on a separate continent of his own" (Melville, *Moby-Dick* 108). The challenge of the "*Isolatoes*" mirrors the challenge of the young republic to retain the distinctiveness of place and of each individual.

Melville's concerns about mobility and placelessness reflect recent cultural debates and ecocritical concerns over the role of "place" in today's world. As critic Lawrence Buell suggests, citing theorist Marc Augé, "environmental criticism must confront the proposition that 'non-places are the real measure of our time'" (The Future 69). Melville provides us with not only a genealogy and early cartography of placelessness in the United States, a phenomenon akin to Augé's "nonplaces," but also, perhaps, with a potential roadmap for navigating this terrain. He contemplates place, as we have seen, with an eve to metaphysics and along a continuum with the "void" (or total placelessness) at one end and distinct and delineated locales at the other. While Augé reminds us that absolute placelessness does not actually exist (at least not on earth and in everyday life), Melville's novels register degrees of place and placelessness that derive from subjective perceptions, interpretative perspectives, and emotional responses. In so doing, he affords us new understandings of how we live in, experience, and move about the world-and of consequences of that movement.

## Works Cited

Adams, Robert Martin. NIL: Episodes in the Literary Conquest of Void During the 19th Century. New York: Oxford University Press, 1966.

Augé, Marc. Non-Places: An Introduction to Supermodernity. Translated by John Howe. London: Verso, 2008.

Buell, Lawrence. The Environmental Imagination: Thoreau, Nature Writing, and the Formation of American Culture. Cambridge MA: Belknap Press of Harvard University Press, 1995.

---. The Future of Environmental Criticism: Environmental Crisis and Literary Imagination (Malden MA: Blackwell, 2005).

Casey, Edward. The Fate of Place: A Philosophical History. Berkeley: University of California Press, 1998.

---. Getting Back Into Place: Toward a Renewed Understanding of the Place-World (Studies in Continental Thought). Bloomington: Indiana University Press, 2009.

Delbanco, Andrew. Melville: His World and Work. New York: Vintage, 2005.

Deleuze, Giles and Félix Guattari. A Thousand Plateaus: Capitalism and Schizophrenia. Translated by Brian Massumi. Minneapolis: University of Minnesota Press, 1987.

Dimmock, Wai-Chee. Empire of Liberty: Melville and the Poetics of Individualism. Princeton: Princeton University Press, 1989.

Hawthorne, Nathaniel. The English Notebooks. Edited by Randall Stewart. New York: Modern Language Association of America, 1941.

Leonard, David Charles. "Descartes, Melville, and the Mardian Vortex." South Atlantic Bulletin 45: 2 (1980): 13-25.

Melville, Herman. Melville, Herman. Moby-Dick. Edited by Harrison Hayford and Hershel Parker. New York: W.W. Norton & Company, 1967.

---. Redburn: His First Voyage. New York: Penguin, 1986.

Pascal, Blaise. Pensées. Edited by Louis Lafuma. Paris: Editions de Luxembourg, 1951.

Thrift, Nigel. "Inhuman Geographies: Landscapes of Speed, Light, and Power." In Writing the Rural: Five Cultural Geographies, edited by Paul Cloke, Marcus Doel, David Matless, Martin Phillips, and Nigel Thrift, 191-248. London: Paul Chapman, 1994.

Tomes, Robert. "The American on Their Travels." Harper's New Monthly Magazine 31: 181 (June 1865): 57.

# Out-printing the Crown: Richard Carlile's "Leaking" of *The Age of Reason* during the Regency Period

**Andrew Doub** 

Brigham Young University

#### Abstract

Between 1818 and 1824, radical printer and publisher Richard Carlile made a determined effort to disseminate copies of Thomas Paine's banned text The Age of Reason in England. Despite strict censorship laws and harsh legal penalties used to curtail previous publishers of this title, Carlile employed a number of creative techniques that kept Paine's deistic writings in print and in circulation during the Regency period. These included republishing public domain court documents and reading The Age of Reason in its entirety into testimony during his trial for seditious libel, making it part of the public record. Copied from trial transcripts and reprinted in cheap pamphlet form, Carlile's editions of The Age of Reason would sell an impressive 20,000 copies in these formats. He managed to provide widescale access to a work that had been successfully suppressed by the British government since its original publication in 1794. My paper argues that Carlile's approach to subverting censorship of The Age of Reason should be viewed as an early instance of what is today known as the mass "leaking" of a restricted document. His unconventional publishing methods, network of collaborators, and use of technology to distribute this work place Carlile at the vanguard of document leaking, making his contribution an important but often overlooked episode in Regency print culture. My study places Carlile's work in a new context and contributes to recent scholarship, including work put forth by Andrew Franta and Russ Castronovo, that seeks to identify parallels between modern surveillance state controversies and similar events from the Romantic period. Like the massive document leaks of our day, Carlile's publication set a new standard for law-breaking innovation in the pirate publication industry.

In the early fall of 1819, radical journalist, publisher, and printer Richard Carlile was ordered to stand trial for illegally reprinting-as a series of pamphlets-Thomas Paine's banned text, The Age of Reason. His trial commenced in October with the defendant himself offering the first exhibit for the court's consideration. Over the course of 12 long hours, Carlile stood before the jury and read aloud, from cover to cover, his edition of Paine's banned text (Bald 8). Following this epic testimony, the court entered into evidence on behalf of the defense the illegally produced set of pamphlets printed at Carlile's Fleet Street shop earlier in the year. This had been Carlile's objective from the moment he was brought to trial: to make the content of Paine's controversial text a matter of public record (Bald 8). Even though Paine's Deist writings had been vigorously suppressed by the Crown for two decades, Carlile creatively identified a loophole that he could exploit in the midst of his own legal misfortunes, much to the chagrin of those who put him on the docket.

With Carlile's trial testimony, Paine's text entered the public domain—and henceforth *The Age of Reason* could legitimately be reprinted as documentation of the court's proceedings. Carlile's compatriots began producing pamphlets of his testimony en masse. The irreverent work would go on to sell an impressive 20,000 copies in this format, and the crown could do nothing to stop it (Bald 8). Ironically, abridged editions of *The Age of Reason* consisting of passages that the prosecution itself specifically excerpted and deemed the most blasphemous were simply copied from his indictment and mass-produced. Instead of halting the publication of *The Age of Reason*, the trial would ensure the widespread promulgation of its most "dangerous" ideas, thanks to Carlile's ingenuity.

This and other pirate-publication methods of Richard Carlile are worth examining more closely. Although brief descriptions of Carlile's subversive work have found their way into scholarship written about early nineteenth-century radical print culture, few have discussed his Age of Reason exploits in detail or have evaluated his technique. I argue that Carlile's publication of The Age of Reason should actually be recognized as an early instance of mass "leaking."1 The strategies Carlile developed to ensure the successful dissemination of Paine's text parallel many of those employed by internet leakers today. Like the founders of The Pirate Bay or Julian Assange, Carlile possessed a relentless philosophical dedication to free speech and in finding new ways to distribute ideas freely. My paper reveals that the methods of subverting state censorship employed by Carlile anticipate modern subversive techniques, including the use of legal loopholes and technological advances to disseminate information. Thus, Carlile's fully legal distribution of The Age of Reason also anticipates current strategies for effectively "leaking" restricted content.

To understand the significance of Carlile's work, one must first understand the legal odds against him, together with the nature of state actions taken to prevent The Age of Reason from being circulated in England. The suppression of Paine's text serves as a valuable case study of the repressive conditions that publishers and booksellers endured during the Regency. Of the censorious climate of this period, William St. Clair notes that "many in authority feared the growth of reading as such, irrespective of textual content" (308). A literate, informed, and intellectually diverse citizenry was very undesirable to those seeking to maintain the power of the aristocracy. Andrew Franta's monograph on the "emergence of a mass reading public" in the uncertain years after Napoleon's defeat at Waterloo details the reaction of a government gravely concerned with the effects of press freedoms in a new era of broad political interest (1). Concurrent with the conception of this mass public, Franta finds that "the growing scope of the press's influence ... prompted a record number of prosecutions for libel in postwar England" and a drastic expansion, on legal grounds, of what libel constituted (12).

Libel charges would give the government a new legislative weapon against dangerous reading material that publishers unleashed. Crucially, prosecutors linked the effects of libelous speech to sedition. Franta explains that "[b]ecause the law of criminal libel defined politi-

<sup>&</sup>lt;sup>1</sup>I use the term "leaked" here (and elsewhere in my argument) in the modern sense of a massive dissemination of proscribed information to the general public.

cally dangerous expression in terms of a text's potential for inciting a breach of peace, libel trials focused on a publication's consequences" (12). Prosecutors only had to prove the possibility that some type of treasonous behavior could be inspired by an interpretation of the language in a text, whether it was explicit or not. This ensured that even the most mundane or veiled political and religious statements could be construed as criminally treasonous.

Carlile and members of his profession, especially those who shared his ideological convictions, came under attack from Tory supporters of the Prince Regent. William Wilberforce's Society for the Suppression of Vice (founded in 1802) saw the publishing community as a serious threat that needed to be neutralized or made to toe the party line. As Don Herzog emphasizes, anything even remotely controversial was officially banned or prohibitively taxed to prevent it from "poisoning the minds of the lower orders" (100). And if some books were poison, Tom Paine was intellectual strychnine. The authorities were worried that his so-called liberal and anti-establishment assertions "would cause a resurgence of the political philosophies and egalitarian ideals that had inspired the revolution in France" (St. Clair 308). Thus, the beginning of the Regency period saw the 1799 Seditious Societies Act—and its broad censorship applications—renewed and strengthened (St. Clair 308). In 1811, Parliament reaffirmed the act, and the charge of "seditious libel" would be used to ensure a virtual blackout of Paine's writings in England (St. Clair 311).

The fate of Carlile's predecessors in the business of pirating Paine should probably have discouraged him from having anything to do with Paine's writings. From the moment that the first copies of The Age of Reason came off the press in 1794, the pamphlet was denounced as blasphemy against the Church of England and an affront to God's anointed monarchy. Censors and government lawyers worked hard to make printing or selling Paine untenable. In 1797, Thomas Williams was found to have printed 2,000 copies of the first British edition of the text in response to the interest it garnered following its debut. The copies that remained in his possession were confiscated, and Williams was indicted for producing seditious and blasphemous material. He was convicted, fined a hefty £1,000, and sentenced to one year of hard labor (Bald 8). The court concluded that "by subverting the truths of Christianity, [Paine's writings] undermined the government and the Constitution," both of which rested on the authority of the Anglican Church (Bald 8). As a result of the judge's verdict, no edition of Paine's text "was sold openly in England after 1797" (Weiner 108).

In 1812, plebeian journalist Daniel Isaac Eaton patched together "some minor theological writings by Paine, which he described as 'Part Three' of The Age of Reason" and began selling them to book stalls (Weiner 108). The authorities were quick to respond. He was convicted of printing blasphemies, sentenced to stand in the pillory, and served 18 months in Newgate Prison for what was essentially a phony bootleg of the text (Bald 8). Right after his release, he was discovered to have yet another edition in the works and was swiftly convicted again (Bald 8). Eaton's trial and news about Paine's continued suppression was, as Kelly Walters puts it, "all very good for publicity" and created a "surge in demand for the book" (Walters 19-20). However, Eaton was simply unable to meet it, given the effectiveness of government censors at containing The Age of Reason using the expanded prosecutive powers given them by Parliament. The publishers who produced any editions of the work, even faked ones, were quickly jailed and their inventory burned. Thus, it was not until the late 1810s, when Richard Carlile's publications of The Age of Reason began to be "leaked" to the public, that Paine's deistic writings were known to British readers on any significant scale. Carlile's personal persistence, the dedication of his network of associates, and the unconventional tactics he used to keep The Age of Reason in print were key factors in his success.

An important difference between Carlile and his predecessors was his remarkable dedication to everything Paine wrote. Carlile's profile is remarkably similar to that of the modern "leaker." Like Edward Snowden or Julian Assange, Carlile's devotion to his principles drove him to overcome personal and practical difficulties standing in his way. An established radical and freethinker from a young age, Carlile first read a smuggled copy of The Age of Reason in December 1817, and, as Joel H. Weiner suggests, "the impact was literally enough to last him a lifetime" (106). Paine's ideas had a profound effect on Carlile, and the devotion of the latter was so complete that he named his son Thomas Paine Carlile in the writer's honor (Weiner 111). Carlile biographer G. D. H. Cole remarks that, regarding Paine, Carlile "had the endurance and the persistence of the fanatical believer" (11). Publishing The Age of Reason became a kind of sacred mission for Carlile to advance a free press and to secure the legacy of a writer whom he esteemed above all others. Williams's reasons for publishing earlier editions of Paine's text were, for the most part, purely financial. Eaton's motivations are harder to ascertain, but to whatever extent his ideologies were aligned with Carlile's, Eaton lacked the cunning and endurance that Carlile would later exhibit. Like Snowden and Assange, Carlile's ideology-even his personality-was inextricably bound to planning and executing the dissemination of Paine's words.

The publication process that Carlile employed would ensure that his editions of *The Age of Reason* could not be suppressed as others had

been. Although he obviously lacked a means of distribution as convenient or available as the internet, Carlile created a proto-analog network through which he would deliver the leak. Carlile's first move in March 1817 was to establish a business relationship with fellow Paineite and printer William T. Sherwin. Carlile took over operation of Sherwin's presses and got to work, initially reissuing only Paine's political writings as a sort of trial run. Then, working from original editions, he began reprinting Paine's theological texts, most especially including The Age of Reason, beginning in December 1818 (Weiner 108). Knowing he had precious little time to spare, Carlile and his shopmen churned out Paine's religious texts in a variety of forms and in large print runs. Cole writes that "[h]e sold these works both in expensive bound volumes and in cheap parts, at prices which poor people could afford to pay" (3). Carlile secured the necessary materials needed for physical production of the texts, and he used his specialized business acumen as a bookseller to ensure maximum distribution.

The first 1,000 copies of his edition of The Age of Reason sold out within a month. The event was "companied by a wave of excitement" from the reading public (Weiner 109). A second edition of 3,000 copies was quickly ordered; it sold at an even faster rate. Carlile then published an anthology entitled Theological Works (1818) that contained, among other writings, the highly controversial and extremely hard-tofind Part 3 of The Age of Reason (Weiner 109). Carlile organized a complex and secretive network of booksellers and intermediaries who moved the copies. By the time the authorities had enough evidence to bring him to trial in the autumn of 1819, the damage was done. His editions of The Age of Reason were very much in circulation and were causing what Weiner describes as a "revival" of interest in Thomas Paine's writings (Weiner 106). The forbidden nature of reading a censored work added to the allure of acquiring the text, a motivation that easily finds its parallel among modern individuals who get excitement from downloading, copying, and redistributing large PDF files of confidential documents on Wikileaks or peer-to-peer file-sharing sites.

While the previous prosecutions of Williams and Eaton ended their publication of *The Age of Reason*, Carlile used his own trial, and the trials of his wife and shopmen, as the means of continuing his publication of the banned text. The popularity and availability of *The Age of Reason* in trial transcript format would result in the government's eventually lifting the ban. Nevertheless, in the short run Carlile was convicted of libel and, in late October 1819, was sentenced to six years in Dorchester Gaol for blasphemy and publishing seditious material. Immediately after the trial, Carlile's wife—Jane Carlile—began printing cheap copies of the trial reports she acquired from the magistrate. These copies contained inflammatory parts of *The Age of Reason* and other deistic writings that Carlile had read into his court testimony (Cole 11). The practice of publishing documents related to one's own legal proceedings can be traced to Daniel Defoe, whose *Hymn to the Pillory* (1703) was circulated among the crowd during his time in the stocks for publishing reformists' writings (Wood 26). But Carlile's persistence was unique, and, over the next decade and more, edition after edition of trial transcripts were published, editions incorporating *The Age of Reason* and other "banned" texts. As family members and workers in the shop subsequently faced legal prosecution, including courageous women who kept Carlile's doors open during his darkest years,<sup>2</sup> each trial resulted in another series of pamphlets produced by those who were lucky enough to not be in jail at the moment. This exploitation of public domain laws left the hands of government censors tied behind their backs.

The most interesting method the shop used to sell Paine's works involved Carlile's own invention-a banned books "vending machine" that protected his employees from prosecution. Pairing technological innovation with print distribution, Carlile's machine utilized a large circular dial displaying the titles of banned texts that could be purchased. The customer turned a clock arm to the title he or she wanted to buy and deposited the appropriate amount of money in a designated drawer. The book, appropriately wrapped in brown paper, was then delivered through a hole in the shop's ceiling or slipped through a veiled slot underneath the dial (St. Clair 313). The machine served as an un-prosecutable middle man: no one could observe that actual booksellers were involved in the transaction, so no individual could be definitively charged. This anonymity also made the patron feel more comfortable with the sale. As with the modern leak, members of the distribution network had no need to meet each other in dark back alleyways, passing a manila envelope in a rolled-up newspaper. Like the internet, Carlile's machine helped create a distance between individuals involved in illicit-text transactions.

Carlile's popular trial transcripts flooded Britain with tens of thousands of copies of Paine's theological writings. Other publishers would bring out pirated reprintings of *Carlile's* editions, truly placing *The Age of Reason* beyond the control of government censors. Cole notes that during the time Jane Carlile herself was not in jail, she "ran the shop with considerable success, profiting by the large demand

<sup>&</sup>lt;sup>2</sup>See Angela Keans' excellent essay on this subject, "Richard Carlile's Working Women."

which the trials had produced for a type of publication which the Government and the Vice Society were uniting their best efforts to suppress" (19). By 1824, Paine's writings were so readily available that the government withdrew its suppression of them, and "Paine's writings were never again the object of a prosecution in England" (Weiner 123). Aside from his convictions and the time he spent in prison, Carlile was victorious over those who sought to control what he could print and to whom he could sell it.

The Regency-era publication of The Age of Reason by Richard Carlile is, above all, an example of how one small group of individuals banded together to effectually out-print government censorship. Like the mass-torrenting, downloading, and internet file sharing of today, Carlile's methodology relied on simple principles to create a leaking of information to a widespread populace. Like Carlile and his print shop personnel, Assange's Wikileaks team procured originals of banned documents, copied them, and found means to disseminate them as widely as possible. Like Assange, Carlile thought of himself first and foremost as an idealistic journalist tasked with the responsibility of ensuring a free intellectual marketplace. In both cases, teams of dedicated assistants continued their leaders' work. And, most importantly, the success of both Carlile and Assange was measured by the availability of the texts they reproduced and by the number of people who read those texts. Carlile's resourcefulness and dedication brought about genuine change in press freedom in Great Britain and anticipated the techniques that would be used to spread information in a later age.

## Works Cited

Bald, Margaret. Banned Books: Literature Suppressed on Religious Grounds. New York: Facts on File, Inc. 1998. Print.

Cole, G.D.H. Richard Carlile, 1790-1843. London: V. Gollancz and the Fabian Society, 1943. Print.

Franta, Andrew. Romanticism and the Rise of the Mass Public. New York: Cambridge UP, 2007. Print.

Herzog, Don. Poisoning the Minds of the Lower Orders. Princeton: Princeton UP, 1998. Print.

St. Clair, William. The Reading Nation in the Romantic Period. Cambridge: Cambridge UP, 2004. Print.

Weiner, Joel H. "Collaborators of a Sort: Thomas Paine and Richard Carlile." Citizen of the World: Essays on Thomas Paine, Ed. Ian Dyck. London: Christopher Helm, 1987. 104-128. Print.

Walters, Kelly. Introduction. The Age of Reason. By Thomas Paine. Toronto: Broadview Editions, 2011. Print.

Wood, Marcus. Radical Satire and Print Culture, 1790-1822. Oxford: Oxford UP, 1994. Print.

## Patterning Supported Lipid Bilayers Using Magnetic Tweezers

Travis M. Bulloch, T. Charles Argyle,<sup>1</sup> Madeline Gleave Parson,<sup>1</sup> and Christopher F. Monson

Southern Utah University

#### Abstract

Cell membranes, which are essential to living cells, consist of a lipid bilayer with membrane proteins and other membrane components integrated into the membrane. The cell membrane separates the fluid inside the cell (the cytoplasm) from the fluid surrounding the cell. The arrangement of components within the cell membrane is also extremely important, particularly in cell communications. Supported lipid bilayers consist of a lipid bilayer a few nanometers above a solid support (often glass) and are a convenient platform with which to study membrane patterning. Previous work with supported lipid bilayers has explored several methods by which various grid-based patterns can be formed in the cell membrane in an attempt to interact with cellular communication at a molecular level. The previously explored grid-

<sup>&</sup>lt;sup>1</sup>These two authors contributed equally to this work.

<sup>\*</sup>This paper was presented at the 2015 UASAL Annual Conference.

based patterns do not provide a way to generate the type of circular pattern that would be required for cell synapses. We generated circular patterns using supported lipid bilayers containing a small amount of fluorescein-labeled lipids. We then introduced superparamagnetic beads coated with anti-fluorescein antibodies, causing the aggregation of fluorescein-labeled lipids into a circular pattern under each bead. The beads were removed from the surface with a magnet, and pattern formation was confirmed by fluorescence microscopy.

## Introduction

Cell membranes are complex macromolecular aggregates composed of phospholipids, proteins, cholesterol, and other macromolecules arranged in a bilayer. Membranes are essential for cell function, communication, and transport and thus are of interest to multiple scientific disciplines. The lipids that make up the majority of a bilayer have polar heads and nonpolar tails. When a collection of such lipids is placed into aqueous solution, the nonpolar tails will aggregate, as will the polar heads.<sup>1</sup> The result is a structure resembling that in Figure 1, in which lipids are oriented such that the polar (hydrophilic) head groups (represented by spheres) are in contact with aqueous solution while the nonpolar (hydrophobic) hydrocarbon tails (represented by wavy lines) are sequestered from the aqueous solution.



**Figure 1.** A phospholipid bilayer cross-section (in vesicle form). Each lipid is represented by a filled circle and two lines. The filled circles represent the polar head groups and make contact with aqueous solution. The two lines represent the nonpolar tails and are sequestered from aqueous solution.

The arrangement of components within a cell membrane plays a crucial role in cell communications, immune responses, and cellular signaling in T-cells and other biological processes.<sup>2</sup> One important example is the immunological synapse between a lymphocyte and an antigen-presenting cell. As illustrated in Figure 2, the center of the synapse on a lymphocyte displays T-cell receptors, which interact with the major histocompatibility complex (MHC) displayed on the antigenpresenting cell. Around the edges of this interaction, lymphocyte function-associated antigen 1 (LFA1) molecules on the lymphocyte interact with intercellular adhesion molecule 1 (ICAM1) on the antigenpresenting cell. If any of these interactions is unsatisfactory or absent, the antigen-presenting cell is targeted for destruction. This forms the basis of much of the immune response to viral infections and cancers, and defects in this system have been linked to many harmful diseases in humans.<sup>3-5</sup> For a synapse to form properly, the membrane components must arrange themselves correctly, with the MHC in the central region of the synapse and the LFA1 in a ring around the edges. This pattern is referred to as a "bulls-eye" pattern. The ability to generate this pattern in a synthetic bilayer would be invaluable to medical research and drug therapy. As such, the bulls-eye pattern has been the focus of much research.6,7



**Figure 2.** The immunological synapse. (A) Proteins on both the immune cell (lymphocyte, left) and the body cell (antigen-presenting cell, right) must interact and arrange themselves correctly for the successful formation of an immunological synapse, with specific proteins on the inside (T cell receptors and MHC) and the outside (LFA1 and ICAM1) of the synapse. When viewed perpendicularly to the plane of the bilayer, these form a "bulls-eye" pattern (B).

An attractive platform to create this sort of synapse in vitro is the supported lipid bilayer (SLB). A SLB forms when spherical lipid vesicles encounter a smooth glass, quartz, or silicon surface.<sup>8</sup> A lipid vesicle, upon encountering such a surface, flattens and ruptures, and SLB growth continues as additional vesicles encounter the free bilayer edge.9 Within a SLB, lateral diffusion of membrane components remains nearly unchanged when compared with free bilayers,<sup>10</sup> making the SLB a valuable platform for investigating interactions of immunological systems.<sup>11</sup> Such techniques could be applied to living cells to control biological targeting. Simple patterns such as parallel rows<sup>12</sup> and square grids<sup>13,14</sup> have been prepared via photolithography, electrophoresis, and soft lithography, but these designs are limited by the rigidity of the patterning method and have to date been unable to interact successfully with living systems. Because of the complexity of shapes and patterns in the surface membranes of living cells, a more robust, dynamic method of patterning would likely be required if SLB patterning techniques were to be applied in vivo.<sup>15</sup> Magnetic tweezers have shown promise in this regard. Magnetic tweezers are superparamagnetic microbeads, with an iron oxide core that allows magnetic control, and a polymer coating that allows conjugation with biological molecules. These features, along with their microscopic size, make them attractive for dynamic patterning experiments.<sup>16</sup> In this paper, we used magnetic tweezers conjugated with anti-fluorescein to create a circular pattern in a SLB with fluorescein-labeled lipids. This study may open the door to more complex patterning in SLBs using magnetic tweezers.

## **Experimental Methods**

### Magnetic Tweezers Calibration

The magnetic tweezer apparatus was calibrated by placing beads in glycerol solutions and measuring bead velocity at various magnet– bead distances. A stack of permanent rare-earth disc magnets 9/16-inch in diameter and 1/8-inch thick (K&J Magnetics) were used to produce the magnetic field. After determining bead velocity, it was possible to determine the force of the magnet stack at each distance using the Stokes equation:

#### $F_{Bead} = 6\pi\eta rv$

where  $\eta$  is the viscosity of the solution, *r* is the radius of the bead, and *v* is the velocity of the bead.<sup>17</sup> A plot of magnet distance vs applied force is shown in Figure 3. In our experiments, the bead-magnet distance was around 4 mm, thus the applied force was greater than 5 pN. This force

is sufficient to disrupt a small number of antibody-antigen interactions.  $^{18}\,$ 



**Figure 3.** Force vs. magnet distance. Values of force were calculated using the Stokes equation. Known values of viscosity for 60% and 80% glycerol solutions were used (10.8 cP and 60.1 cP, respectively). The bead diameter,  $5.8 \mu m$ , was obtained from the manufacturer.

#### Cleaning Slides

Slides were immersed in a boiling 7X solution (MP Biomedicals; approximately 1:10 dilution with deionized water) on a slide rack for 30 minutes. They were rinsed with copious amounts of flowing deionized water and blown dry with air before being annealed at 530°C overnight. Slides were stored upright in a clean plastic box until use.

#### Lipid Preparation

Solutions of fluorescein-labeled 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine (DOPE) in chloroform (Avanti) and 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) in chloroform (Avanti) were used in vesicle preparation. Appropriate amounts of each lipid solution were measured and mixed to form the desired lipid composition. This was typically 0.5 mole % fluorescein-labeled DOPE and 99.5 mole % POPC. The chloroform was dried with air, and the lipids were placed in a vacuum (~4 Torr) for 1–3 hours. The mixture was then hydrated with 1 mL of a 0.01 M Tris(hydroxymethyl)-aminomethane (Fisher), 0.1 M NaCl, pH 7.4 (Tris) buffer solution while mixing vigorously. This usually eliminated the need for freeze-thaw cycles before extrusion. The lipid solution was extruded with a mini-extruder

(Avanti) using a 100-nm-pore polycarbonate membrane until the solution was clear (usually about 30 times). The solution was then diluted to approximately 1 mg/mL total lipid concentration using Tris buffer.

#### **Bead Preparation**

Superparamagnetic microspheres (Bangs Labs, 5.8-µm diameter) were rinsed several times with deionized water and then suspended in basic phosphate buffered saline (PBS, 0.1 M NaCl, 0.01 M phosphates, pH  $\approx$  9). 1-Ethyl-3-(3-Dimethylaminopropyl)carbodiimide, Hydrochloride (EDAC), and anti-fluorescein were added to the solution, and the beads were allowed to react for one hour with gentle shaking. They were then rinsed with Tris buffer several times and stored at 4°C until use.

#### **Bilayer** Formation

Previously cleaned glass slides were rinsed with deionized water and blown dry with air before use. Polydimethylsiloxane (PDMS) wells, fabricated as previously described,<sup>9</sup> were rinsed with methanol or acetone and deionized water and then cleaned with transparent office tape before being applied to the cleaned slides. Between 30 and 60 µL of lipid solution was pipetted onto the clean surface of the slides inside the well and left for 30 minutes. The slide was then thoroughly rinsed with deionized water. After rinsing, the slide was scratched in an xpattern using a dissection probe to permit the fluorescent bilayer to be differentiated from the non-fluorescent scratch. It also served as a zero reference point when examining the slide under the microscope and assisted in focusing the microscope on the supported lipid bilayer. A drop of Tris buffer was added to the water in the well and then a few uL of bead solution was added to the well. The beads were allowed to interact with the bilaver for 30 minutes, after which the bilaver was imaged both fluorescently and with ambient light. A stack of 6 magnets was placed 4 mm above the slide to remove the beads. The device was imaged with a fluorescent microscope several times each minute during bead removal

#### Results

The bead-bilayer interaction resulted in fluorescent aggregations of fluorescein-labeled lipids (Figure 4). As fluorescein-labeled lipids diffused laterally throughout the SLB and came into contact with antifluorescein antibodies bound to the beads, they were held in place by antibody-antigen interactions. Over time (typically around 30 minutes in our experiments), this caused the fluorescein-labeled lipids to accumulate beneath the beads and form a circular pattern in the bilayer directly under each bead. After sufficient time had passed for pattern formation, magnets were placed 4 mm above the beads. The force of the magnets was sufficient to pull a fraction of the beads away from the bilayers, and each removed bead left behind the pattern that had formed beneath it.



**Figure 4.** Illustration of magnetic tweezers with antibody conjugates interacting with the SLB on a glass slide. The lipids are freely diffusing throughout the bilayer and thus initially randomly distributed (left). The labeled lipids are immobilized underneath the magnetic tweezers by interactions between the antifluorescein antibodies of the magnetic tweezers and the fluorescein tags on the lipids.

On bead removal, pattern formation was confirmed using a fluorescence microscope. Because of the relatively poor detection limit of our microscope and magnetic tweezer setup, we were unable to directly visualize the spots formed but had to rely on image analysis to verify pattern formation. Prior to the application of the magnets, a picture was taken with the fluorescence microscope using ambient lighting. After the application of the magnets, pictures were taken several times each minute of the green fluorescence emitted by the fluorescein-labeled lipids in the SLB. The fluorescence images were overlaid with images taken before placement of the magnets (Figure 5). The first picture (taken before placement of the magnets) shows the initial position of the beads. The second picture (taken with fluorescent light) shows the positions at which beads are still present after placement of the magnets. By combing these two pictures, we were able to determine the positions at which a bead had been present before, but not after, placement of the magnets. These sites (the sites that had beads before but not after magnet placement) are the sites at which pattern formation and confirmation was possible. After identifying these spots on the overlaid image, we located the same spots on non-overlaid fluorescence pictures. Light intensity values were measured at such sites and compared with light intensity values obtained from positions at which no beads

were found (and therefore no patterns could have formed). Because nearly all light detected was emitted by the fluorescein, which had been concentrated under the beads, a difference in light intensity between the sites of pattern formation and the sites where no pattern formation occurred was expected. The difference in light intensity at these sites was followed over time, which allowed us to evaluate the intensity drop as function of time. This is associated with the diffusion of fluoresceinlabeled lipids from the initial patterned state into a nonpatterned, random state. The process we monitored is essentially the reverse of the fluorescence recovery after photobleaching (FRAP) experiment, in which lipids are bleached in a spot and then the diffusion of unlabeled lipids into the bleached region increases the fluorescence of the spot. In FRAP, the fluorescence signal increases as a function of time, while in our experiments a similar diffusion-controlled process causes the fluorescence to decrease as a function of time.



**Figure 5.** Merged false-color images of the magnetic tweezer positions before and after magnets were added. The ambient light image from before the beads were removed was colored red (beads appear light grey in image) and merged with a fluorescent image colored blue (beads appear as medium grey spots with a darker ring around them) taken after the magnets were in place and the beads were removed. Red dots on top of blue dots (lighter grey spots on top of a darker ring) indicate a bead that was not removed. Isolated red dots (light grey) indicate beads that were removed by the magnets and are the sites of possible pattern formation. Isolated blue dots (dark rings) show beads that drifted to another position in the magnetic field without being removed from the bilayer and were not analyzed.

As can be seen in Figure 6, the first few minutes are the most important in the decay of intensity patterns. This is due to the very small size of the patterns formed—although the beads themselves are 5.8  $\mu$ m in diameter, the actual area of the beads in contact with the SLB is small because of the curvature of the beads—perhaps as small as tens of nanometers. Unlike other patterning techniques, patterns formed by this method have no barriers to lateral diffusion within the bilayer. As a result, the patterns formed decayed quickly, on the order of a few minutes. Thus, images were taken at a 1-second exposure every 15 seconds for the first 2 minutes. Efforts were made to capture more than 6 fluorescent images per minute, but with our setup this resulted in such a degree of photobleaching that the decay was indistinguishable. The quick decline of the initial signal is consistent with previously reported diffusion constants for supported lipid bilayers.

While performing the experiments, we noted that in all cases where binding occurred, a significant fraction of the beads were not removed. These immobile beads likely come from one of several sources. Those on a scratch or other area where the bilaver exhibited irregularity could result from strong nonspecific binding between the antibody conjugates and exposed glass. There were also many immobile beads observed in regions of high-quality SLB. These beads were probably bound by a large number of antigen-antibody interactions that resulted in a stronger bead-bilayer bond than could be disrupted by the force supplied by the magnets. Indeed, this behavior has been observed before and was expected, particularly as the force we applied was relatively small (5-15 pN).<sup>18,19</sup> Incidentally, had these beads been removed they would have produced the strongest signals as they concentrated the largest number of fluorophores. Finally, a few beads drifted under the magnetic field as shown by isolated dark rings in Figure 5. Only those beads that appeared to have been removed without drifting were analyzed. Finally, it should be noted that the light grey fluorescence image spots in Figure 5 often seem to be offset from the beads that did not move. This is an artifact of the method by which the image was taken-the lighting was largely from one side and in the initial image the beads exhibited a bright spot at their top right edge (from the point of view of the camera). This bright spot is highly visible in the merged image; the slight offset allowed beads that had not been removed to be identified and the alignment of the images to be verified.

A control was prepared as above with the exception that Texas Red-labeled lipids were used in place of fluorescein-labeled lipids and the fluorescent images were captured using a Texas Red light filter. In this case, no antibody-fluorophore interaction—and therefore no pattern formation—was expected. The Texas Red-labeled lipids failed to show aggregation in conjunction with the anti-fluorescein–labeled magnetic tweezers.



**Figure 6.** Time vs. intensity profile for fluorescein-labeled lipid patterns (diamonds) and a Texas Red–labeled lipid control (squares). The pattern formed by the beads (initial high value in the blue fluorescein data) diffuses away over 1– 2 minutes and is absent in the Texas Red control. The error bars represent the standard deviation of the normalized average intensity value at all points for each time interval.

### Conclusions

We have demonstrated the formation of patterns in SLBs using magnetic tweezers. The patterns formed were beyond the capabilities of our setup to directly image, but pattern formation was verified by image analysis. The use of a magnetic tweezers setup granted the ability to interact with the dynamic bilayer in a way that has been difficult with other techniques, which could be valuable as many biological complexes are similarly dynamic.<sup>20-22</sup> Fluorescence microscopy allowed us to detect and follow the decay of very small patterns. Future work will seek to remove more strongly bound beads and thus generate patterns that are directly visualizable, and eventually to incorporate proteins into SLBs, allowing for the study of protein interactions similar to recently reported results.<sup>23</sup>

## Acknowledgments

We thank Susan Harris for making the ceramic rack we used to anneal glass slides. We thank the SUU Faculty Scholarly Support Fund and the L.S. and Aline W. Skaggs Research Foundation for funding this project. M.G.P. thanks the Walter Maxwell Gibson Summer Research Fellowship program for its support.

## References

[1] Voet, D.; Voet, J.G.; Pratt, C.W. Fundamentals of Biochemistry: Life at the Molecular Level; John Wiley & Sons: Hoboken, 2013; 255-285.

[2] Monks, C.R.; Freiberg, B.A.; Kupfer, H.; Sciaky, N.; Kupfer, A. Three-dimensional segregation of supramolecular activation clusters in T cells. *Nature* **1998**, 395, 82-86.

[3] Dejaco, C.; Duftner, C.; Grubeck-Loebenstein, B.; Shrimer, M. Imbalance of regulatory T cells in human autoimmune diseases. *Immunology* **2005**, 17, 289-300.

[4] Shevach, E.M. Regulatory/suppressor T cells in health and disease. *Arthritis Rheum* **2004**, 50, 2721–2724.

[5] Longhi M.S.; Ma Y.; Bogdanos D.P.; Cheeseman P.; Mieli-Vergani G.; Vergani D. Impairment of CD4+ CD25+ regulatory T-cells in autoimmune liver disease. *J Hepatol* **2004**, 41, 31–37.

[6] Grakoui, A.; Bromley, S.K.; Sumen, C.; Davis, M.M.; Shaw, A.S.; Allen, P.M.; Dustin, M.L.; The immunological synapse: a molecular machine controlling T cell activation. *Science* **1999**, 285, 221-227.

[7] Stevens, M.M.; George, J.H. Exploring and engineering the cell surface interface. *Science* **2005**, 310, 1335-1338.

[8] Tamm, L.K.; McConnell, H.M. Supported phospholipid bilayers. *Biophys. J.* **1985**, 47, 105-113.

[9] Cremer, P.S.; Boxer, S.G. Formation and spreading of lipid bilayers on planar glass supports. *J. Phys. Chem.* **1999**, 103, 2554-2559.

[10] Smith, B.A.; McConnell, H.M. Determination of molecular motion in membranes using periodic pattern photobleaching. *Porc. Natl. Acad. Sci. U. S. A.* **1978**, 75, 2759-2763.

[11] Mossman, K.D.; Campi, G.; Groves, J.T.; Dustin, M.L. Altered TCR signaling from geometrically repatterned immunological synapses. *Science* **2005**, 310, 1191-1193.

[12] Van Oudenaarden, A; Boxer, S.G. Brownian ratchets: molecular separations in lipid bilayers supported on patterned arrays. *Science* **1999**, 285, 1046-1048.

[13] Kung, L.A.; Kam, L.; Hovis, J.S.; Boxer, S.G. Patterning hybrid surfaces of proteins and supported lipid bilayers. *Langmuir* **2007**, 16, 6773-6776.

[14] Mossman, K.; Groves, J. Micropatterned supported membranes as tools for quantitative studies of the immunological synapse. *Chem. Soc. Rev.* **2007**, 36, 46-54.

[15] Lenne, P.F.; Lecult T. Cell surface mechanics and the control of cell shape, tissue patterns, and morphogenesis. *Nat. Rev. Mol. Cell Biol.* **2007**, 8, 633-644.

[16] Gosse, C.; Croquette, V. Magnetic tweezers: micromanipulation and force measurement at the molecular level. *Biophys. J.* **2002** 82, 3314-3329.

[17] Haynes, W.M. Glycerol viscosity values. In *CRC Handbook of Chemistry and Physics*, 77<sup>th</sup> edition; Chemical Rubber Publishing: Cleveland, 1996.

[18] Shevkoplyas, S.S.; Siegel, A.C.; Westervelt, R.M.; Prentiss, M.G.; Whitesides, G.M. The force acting on a superparamagnetic bead due to an applied magnetic field. *Lab Chip* **2007**, *7*, 1294-1302.

[19] Monson, C.F.; Driscoll, L.N.; Khoie, S.K.; Bennion, E.; Miller, C.J.; Majda, M. Antibody-antigen equilibria in a field of magnetic forces: design of reagentless biosensors. *ECS Trans.* **2009**, 19, 79-91.

[20] Lippincott-Schwartz, J.; Snapp, E.; Kenworthy, A. Studying protein dynamics in living cells. *Nat. Rev. Mol. Cell Biol.* **2001**, 2, 444-456.

[21] Adessi, C.; Miege, C.; Albrieux, C.; Rabilloud, T. Twodimensional electrophoresis of membrane proteins: a current challenge for immobilized pH gradients. *Electrophoresis* **1997**, 18, 127-35.

[22] Wilkins, M.R.; Gasteiger, E.; Sanchez, J.C.; Bairoch, A.; Hochstrasser, D.F., Two-dimensional gel electrophoresis for proteome projects: the effects of protein hydrophobicity and copy number. *Electrophoresis* **1998**, 19, 1501-1505.

[23] Manford, A.G.; Stefan, C.J.; Yuan, H.L.; MacGurn, J.A.; Emr, S.D. ER-to-plasma membrane tethering proteins regulate cell signaling and ER morphology. *Dev. Cell* **2012**, 21, 2048-2052.

## How Much Less Is More? The Predictive Consequences of Overfitting

## Brian Knaeble,<sup>1</sup> Bill Bynum,<sup>2</sup> Gano Hasanbegovic,<sup>2</sup> Garret Wilcox<sup>2</sup>

<sup>1</sup>Utah Valley University, <sup>2</sup>Westminster College

#### Abstract

The law of parsimony, or Occam's Razor, states that given competing hypotheses that equally explain the phenomena at hand, one should choose the hypothesis with the fewest assumptions. Statistician R. A. Fisher reminds scientists each data set contains a natural amount of information that cannot be increased through ingenious statistical models. Mathematician John von Neumann claims that with four parameters he can fit an elephant, and with five he can wiggle its trunk. In statistics, a model with more explanatory variables than justified by the sample size is said to be overfit, violating the law of parsimony. Overfit models can predict poorly because they have modeled noise rather than underlying relationships between variables. To avoid overfitting, there is a rule of thumb for regression stating that the number of observations must be at least 10 times the number of explanatory variables. To evaluate this claim with respect to prediction accuracy as measured with mean square error (MSE), we have conducted statistical simulations. The simulations have been carried out under a variety of assumptions and the results graphically displayed.

## Introduction

Cochran (1965) quotes Fisher as saying, "Make your theories elaborate," in apparent defiance of Occam's razor until it is realized, as explained by Rosenbaum (2005), that elaborate theories lead to many more specific predictions. If any one of these predictions repeatedly fails to agree with observed data then the theory can be rejected (Popper, 1959). A theory lacking in detail may be too easily made to fit quite diverse observed data, as some claim happened with Freud (Popper, 1965).

Overfitting has occurred when a multiple regression model (Seber and Lee, 2003) uses irrelevant explanatory variables to obtain a better fit (Hawkins, 2004). Much like a line can be drawn through any two points, a (hyper) plane of sufficiently high dimension can be made to fit through a large number of points in high-dimensional space. This fit may be better than expected when the dimension of the plane approaches the number of points from below, and the result is an overfit model. The slope estimates of an overfit model can have inflated variance (Seber and Lee, 2003, p. 231). To avoid overfitting, there is a rule of thumb for multiple regression that recommends at least 10 observations for each explanatory variable in the model (Babyak, 2004). Here we run simulations to test this rule of thumb.

We are testing the extent to which overfitting contributes to prediction error within the context of multiple regression. We are not considering the extent to which adjustment for confounding variables impacts estimates for causal effects (see McNamee, 2005). We have found overfitting to be a valid concern, but mainly when the sample size is only slightly larger than the number of explanatory variables. The previously stated rule of thumb for overfitting (as we have defined it) seems overly conservative, under our somewhat restrictive assumptions. The details of our methodology are described in Section 2. Graphical results are shown in Section 3. A discussion occurs in Section 4. Computer code is provided in the appendix.

#### **Materials and Methods**

Our simulations were carried out using R. We assumed independent, normally distributed, random, explanatory variables and an independent normal error term, along with independent observations. For a given sample size n, we generated data on s explanatory variables,  $x_1,...,x_s$  and an error term,  $\epsilon$ , and determined the response via  $y=1x_1+...+1x_s+\epsilon$  using slope coefficients of 1 for each explanatory variable. We also generated data on an additional p-s explanatory variables,  $x_{s+1},...,x_p$ .

We fit a model of y in terms of  $x_1$  through  $x_s$  and refer to it as our scientific model. We fit a model of y in terms of  $x_1$  through  $x_p$  and refer to it as our overfit model. The performance of both models was evaluated by generating an additional 1000 points, determining the true response values ( $y=1x_1+...+1x_s+\epsilon$ ) as well as the predictions of the scientific model and the overfit model, and using mean square error (MSE) to assess prediction performance.

Values for s were set at 1, 2, 5, and 10. Values for the sample size n were set at 10, 30, 100, and 500. For each (n, s) combination, p ranged from s to min{max{n/5,20},n-1} with the exception of an omitted case for s=10 and n=10. In the case where the sample size was n=500, multiples of 10 were selected across the p range as an adequate selection with a reasonable simulation run time. The specific values of n and s were prospectively chosen so as to cover a range of plausible values. The values for p were prospectively chosen to produce evidence relating to assessment of the rule of thumb described in the introduction.

For each n, s, and p combination tested, and for both scientific and overfit models, the above simulation process was run 1000 times and the mean of the 1000 generated MSEs reported.

#### Results

The series of plots below (Figures 1-4) illustrate the mean MSE for different numbers of scientific model predictors (s) and sample sizes (n), across the range of tested number of predictors (p) for the overfit model. In each plot, the horizontal dashed line represents the mean MSE for the corresponding scientific model while the solid line shows the mean MSE for the corresponding overfit models illustrating the increase in mean MSE as the number of overfit variables increase. For the given sample size in each plot, the vertical dashed line represents the rule of thumb for the number of acceptable predictor variables which recommends p < n/10.



Figure 1. Plots of mean MSE for s=1 Scientific Predictors



Figure 2. Plots of mean MSE for s=2 Scientific Predictors



Figure 3. Plots of mean MSE for s=5 Scientific Predictors



Figure 4. Plots of mean MSE for s=10 Scientific Predictors

### Discussion

The results are displayed for (n,s) combinations with p on the horizontal axis. This is to benefit those analyzing an existing, large data set. The results could have been displayed for (p,s) combinations with n on the horizontal axis, so as to benefit those designing a study, or for (p,n) combinations with s on the horizontal axis, so as to benefit those interpreting the results of an existing study.

It is important to note that MSE is a function of not just n, s, and p, but also depends on distributional and other assumptions. Recall how we assumed independence between observations, independence and normality for the explanatory variables, as well as normality for the error terms. Departures from these assumptions could change MSE. For example, with n=100, s=5, and p=10, simulations with normal error terms produce a mean MSE of 1.11 for overfit models and a mean MSE of 1.05 for scientific models. Simulations with the error terms distributed according to the t-distribution with five degrees of freedom produce a mean MSE of 1.85 for overfit models and a mean MSE of 1.77 for scientific models.

Ideally, all parameters and assumptions could be specified as arguments to a new R function written and programmed to assess concerns about overfitting. Modelers should compare their assumptions with our assumptions before relying on our conclusions. Our R code (Appendix) can be modified to reflect differing assumptions if needed.

Our results do not indicate, even for (n,s,p) triplets where the overfit mean MSE is acceptable, that the model is adequate. The results only indicate that poor prediction due to overfitting, specifically, the use of too many unrelated explanatory variables, is not likely to be a large concern in this setting. We have only shown that the mean MSE is tolerable. In his paper "The Problem of Overfitting," Hawkins (2004) distinguishes between prediction problems (which we have addressed here) and effect quantification, relating to confounding, specifically bias in estimation of causal effects (Greenland and Morgenstern, 2001). Chatfield (1995, p. 423) warns against using statistical information to select a model. McNamee (2003) recommends model selection that is instead based on (scientific) subject matter knowledge and statistical information.

Our results show that prediction error does not always drastically increase when irrelevant explanatory variables are used within multiple regression models. By examining the mean of the MSE, our graphical results indicate that the rule of thumb suggesting at least 10 observations per predictor variable is overly conservative. However, examining other measures, such as the max MSE and 95% MSE, or visualizing the
distribution of MSE across all individual models could provide a more nuanced picture of the consequences of overfitting. The simulations here for reasons of simplicity dealt only with this one rule of thumb while additional studies could interrogate other heuristics present in the literature (Green 1991). Here, simulations were carried out under specific assumptions. Further simulations could focus on departures from normality, allow for collinearity, account for dependencies between explanatory variables, account for dependencies between observations, and use different distributions for the error term. In addition, future simulations could allow selection or even definition of extraneous explanatory variables to be made by researchers based on optimizing some measure of goodness of fit.

# References

Babyak, M. (2004). What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. Psychosomatic Medicine, 66, 411–421.

Chatfield, C. (1995). Model uncertainty, data mining and statistical inference. J.R. Statist. Soc. A, 158(3), 419–466

Cochran, W.G. (1965). The planning of observational studies of human populations (with Discussion). Journal of the Royal Statistical Society, Series A. 128, 134–155.

Green, S. B. (1991). How many subjects does it take to do a regression analysis? Multivariate Behavioral Research, 26, 499–510.

Greenland, S. and Morgenstern, H. (2001). Confounding in health research. Annu Rev Public Health, 22, 189–212

Hawkins, D. M. (2004). The problem of overfitting. Journal of Chemical Information and Computer Sciences, 44, 1-12

McNamee, R. (2003). Confounding and confounders. Occupational & Environmental Medicine, 60, 227–234.

McNamee, R. (2005). Regression modelling and other methods to control confounding. Occupational & Environmental Medicine, 62, 500–506.

Popper, K. R. (1959). The logic of scientific discovery. London: Hutchinson.

Popper, K R. (1965). Conjectures and Refutation: The Growth of Scientific Knowledge. New York: Harper Torchbooks. pp. 34–37.

Rosenbaum, P.R. (2005). Observational Study, Encyclopedia of Statistics in Behavioral Science, Volume 3, pp. 1451–1462.

Seber, George A. and Lee, A. (2003). Linear Regression Analysis. Hoboken, NJ: Wiley.

# Appendix

```
# Code executed in RStudio Version 0.99.491 on Mac OS X
10 10 5
library(ggplot2)
# Simulations run for s=1,2,5,10 and n=10,30,100,500
# Here case s=1 and n=10 is illustrated
s <- 1
n <- 10
trialdata <- data.frame(p var=integer(),</pre>
meansMSE=numeric(),
                          meanpMSE=numeric())
scientific MSE <- 0
overfit MSE <- 0
index <- 1
p cases <- seq(s,min(max(n/5,20),n-1))</pre>
#For n=500 instead use 'p_cases<-c(s,</pre>
seq(10*round((s+9)/10),100,10))'
for(p in p_cases) {
  for(1 in 1:1000) {
    M <- matrix(rnorm(n*p), nrow=n)</pre>
    betas <- c(rep(1,s),rep(0,p-s))</pre>
    y <- M%*%betas+rnorm(n)</pre>
    scientific_s_model <- lm(y~M[,1:s])</pre>
    overfit p model <- lm(y~M)
    new obs <- matrix(rnorm(1000*p), nrow=1000)</pre>
    new y <- new obs%*%betas+rnorm(1000)</pre>
    scientific y.hat <-</pre>
new obs%*%c(scientific s model$coefficients[-1], rep(0,p-
s))
    overfit y.hat <-
new obs%*%overfit p model$coefficients[-1]
    scientific MSE[1] <- sum((scientific y.hat -</pre>
new y)^2/1000)
    overfit MSE[1] <- sum((overfit y.hat - new y)^2/1000)</pre>
  }
  trialdata[index,] <- c(p, mean(scientific MSE),</pre>
mean(overfit MSE))
  index <- index +1</pre>
}
cases <- 2*nrow(trialdata)</pre>
plotdata <- data.frame(p_var=integer(cases),</pre>
Model=character(cases),
                         MSE=numeric(cases))
plotdata$p_var <- c(trialdata$p_var, trialdata$p_var)</pre>
plotdata$Model <- c(rep("Scientific",</pre>
```

292 Physical Sciences

```
cases/2), rep("Overfit", cases/2))
plotdata$MSE <- c(trialdata$meansMSE, trialdata$meanpMSE)</pre>
my white <- theme_bw() +</pre>
theme(plot.title=element_text(size=rel(1.8)),
panel.grid.major.x=element_blank(),
panel.grid.minor.x=element_blank(),
panel.grid.major.y=element blank(),
panel.grid.minor.y=element blank(),
                                axis.title =
element text(size = rel(1.5)),
                                axis.text =
element text(size=rel(1.25)),
legend.text=element_text(size=rel(1.5)),
legend.title=element text(size=rel(1.5)))
mse plot <- ggplot(plotdata, aes(x=p_var, y=MSE,</pre>
linetype=Model, shape=Model)) +
  geom_point() + geom_line() + my white +
  ggtitle(paste("Scientific Predictors s =", s, ", Sample
Size n =", n)) +
  labs(x = "Variables in Overfit Model (p)", y = "Mean
MSE") +
  geom hline(yintercept=0) +
  geom_vline(xintercept=n/10, linetype="dashed") +
  scale x continuous(breaks=min(s,n/10):p) +
  ylim(c(0,min(10, round(max(trialdata$meanpMSE+1)))))
print(mse plot)
#For n=500 instead use
'scale x continuous(breaks=c(s,seg(10,p,10)))'
```

# Density Functional Theory Investigation of Polycyclical Peroxide Stability

# Charles Joseph Simon, Don R. Davies, and H. Laine Berghout

Weber State University

## Abstract

Polycyclic peroxide compounds have been of interest for their antimalarial activity. The synthesis of stable peroxide compounds can be challenging, thus making preliminary computational determination of the stability of promising compounds advisable. Density functional theory is used to gauge the stability of one such peroxide, 2,3,10trioxabicyclo[5.2.1]decan-4-ol. An intramolecular hydrogen bond between the two rings of this bridged bicyclical compound is anticipated to contribute to the stability of the molecule. Based on coordinate scans of the hydroxyl dihedral bond angle at the B3LYP/6-311+G(2d, p) level of theory, we estimate the strength of this intramolecular hydrogen bonding interaction at 8.6 kcal/mol, considerably above the 5 kcal/mol typical for R-O-H---O=C-R'. The mechanism of action of peroxide antimalarial compounds is likely initiated via homolytic bond cleavage of the O-O bond. The energy of activation ( $\Delta G^{\sharp}$ ) for this homolysis process is calculated to be approximately 40.6 kcal/mol in the gas phase, similar to the literature bond dissociation value of hydrogen peroxide at 43.7 kcal/mol, which makes it likely that this process is fairly slow with an estimated uncatalyzed first-order rate constant of  $1 \times 10^{-17} \text{ s}^{-1}$ .

# Introduction

Polycyclic peroxide compounds such as Artemisinin and Yingzhaosu A, shown in Figure 1, have been of interest because of their antimalarial activity.<sup>1</sup>



Artemisinin

Yingzhaosu A

Figure 1: Artemisinin and Yingzhaosu A structures.

Although the exact mechanism of action for these peroxides is not known, the presence of the peroxide bridge must be required to initiate antimalarial activity because similar compounds lacking the peroxide group do not show any antimalarial activity.<sup>2</sup> Further investigation into the mechanism of action suggests that ferrous ions found in the heme portion of hemoglobin bind to the peroxide bridge and induce a homolytic fragmentation of the bond, resulting in an alkoxy radical. This alkoxy radical, or a rearranged carbon radical, is believed to bind to a calcium ion pump called PfATP6, preventing the phosphorylation of nucleosides and resulting in cell death of the parasite.<sup>3</sup> It has also been found that structures having polar functionalities near the peroxide bridge show increased activity. The polar group is thought to assist the peroxide trigger at the binding site of the drug. Steric hindrance of the peroxide has been shown to negatively impact its activity.

The current study is focused on the polycyclic peroxide 2,3,10trioxabicyclo[5.2.1]decan-4-ol, which shares structural properties with other known antimalarials. As a class, organic peroxides can be challenging to synthesize. For this reason, it is advantageous to investigate the stability of these compounds using computational methods before investing laboratory resources in their synthesis. The goal of this research was to gauge stability using electronic structure calculations. 2,3,10-Trioxabicyclo[5.2.1]decan-4-ol and its suspected decomposition pathway are shown in Figure 2.



Figure 2: 2,3,10-trioxabicyclo[5.2.1]decan-4-ol and its decomposition

One factor that makes 2,3,10-trioxabicyclo[5.2.1]decan-4-ol interesting is that it possesses an internal hydrogen bond that potentially enhances its overall stability. The likely hydrogen-bonding interaction is between the hydroxyl hydrogen and the bridgehead oxygen as indicated by a dashed line in Figure 2.

Density functional theory (DFT) calculations have become the standard because of their generally good accuracy and inclusion of electron exchange and correlation energies combined with relatively low computational costs similar to Hartree-Fock (HF) methods.<sup>4,5</sup> DFT calculations use electron density to describe a system instead of the wavefunction used in traditional *ab-initio* calculations.<sup>4</sup> DFT calculations produce fairly accurate energies for intermolecular hydrogen bonds because of their inclusion of electron correlation.<sup>6,7</sup> In particular, the B3LYP/6-311 G(d, p) model chemistry produces accurate results with reasonable economy.<sup>7</sup> This study applied the B3LYP/6-311+G(2d, p) model chemistry with a diffuse function and an extra polarization function added to heavy atoms.

Peroxide molecules such as the one studied in this experiment tend to decompose by homolytic bond cleavage of the O-O bond.<sup>8</sup> For the peroxide studied in this experiment, this process is shown as the first step in Figure 2. The rate at which this bond decomposes can be estimated by finding a transition state structure with its accompanying energy and applying transition state theory calculations.<sup>4,9</sup> DFT calculations at the level of theory used for the hydrogen bond calculations in this experiment have been shown to produce fairly accurate energies with good economy for transition state calculations of similar molecules as well.<sup>10</sup> This is important since the same model chemistry must be used for all calculations for computed values to be comparable. Therefore, the B3LYP/6-311+G(2d, p) level of theory was chosen again.

Obtaining transition state structures can be challenging especially in moderate to large molecules where similar conformations can connect multiple possible local minima on the potential energy surface. Finding a minimum energy structure is not sufficient. The structure calculation must be followed by a frequency calculation to show that it is in fact a first-order saddle point and produces only one imaginary frequency.<sup>11</sup> Further calculations must be done to show that the saddle point connects the structures of the reactants and products on a reaction coordinate.

#### **Computational Procedures**

All of the calculations performed to measure hydrogen bond strength were run at the B3LYP/6-311+G(2d, p) level of theory using the Gaussian 09 electronic structure program.<sup>12</sup> The geometry of 2,3,10-trioxabicyclo[5.2.1]decan-4-ol was optimized, and potential energy surface (PES) dihedral angle coordinate scans, both rigid and relaxed, were carried out to gauge the stability contribution of potential intramolecular hydrogen bonding and other steric and torsional factors in the molecule. A rigid scan varies the coordinate of interest and performs single-point energy calculations at each step while the rest of the molecular coordinates are frozen.<sup>11</sup> A relaxed scan is the same with regard to the coordinate of interest, but the rest of the molecular energy.<sup>11</sup> These scanning calculations were performed by incrementally adjusting the dihedral angle between the hydrogen of the hydroxyl group, labeled "b" in Figure 3, in 10° increments for 36 steps.



**Figure 3:** Dihedral angle pictorial representation with the hydroxyl hydrogen atom labeled "a" and the dihedral reference oxygen atom labeled "b."

This scan produced a potential energy well from which the stability provided by the intramolecular hydrogen bond could be estimated. The lowest, most stable part of the potential corresponded to the conformation with the hydroxyl hydrogen facing the bridgehead oxygen and in closest proximity to its lone pair electrons; and the highest, least stable part of the potential corresponded to the conformation with the hydroxyl hydrogen facing away from the bridgehead oxygen and furthest from its lone-pair electrons. Transition state calculations for this rotation process were run to confirm that the higher energy conformation was indeed a saddle point in the hydroxyl rotation coordinate and to provide an accurate estimate for its energy.

Optimizations and transition state calculations were all run at the same level of theory as the hydrogen bond calculations. Frequency calculations were run on the lowest energy optimized structure of the molecule to determine thermodynamic quantities of the molecule including the zero-point correction, thermal correction for a temperature of 298.15 K, and Gibbs energy of formation from its electrons and noninteracting nuclei. The diradical product of homolytic bond cleavage was optimized by specifying a multiplicity of three, reflecting two unpaired electrons, and requesting that the force constants be calculated. The singlet state with two paired electrons was attempted first but would not converge. Without imposing these parameters, it is difficult to obtain an optimized diradical structure as the molecule loses much of its rigidity upon cleavage of the peroxide and becomes a difficult optimization case.<sup>11</sup> This optimization was followed by a frequency calculation to determine thermodynamic quantities. Once the optimized diradical structure was obtained, the Synchronous Transit-Guided Quasi-Newton (STQN) Method was used, specifically QST2, with the optimized geometries of the molecule before and after homolytic cleavage to obtain an optimized transition state structure.<sup>11</sup> In addition to the optimization that was done, a frequency calculation was performed to determine whether this structure corresponded to a first-order saddle point and yielded a single imaginary frequency. To determine whether or not this structure corresponds to the cleavage process, the reaction path was followed with an intrinsic reaction coordinate (IRC) calculation.<sup>11,13</sup> With this calculation, it was necessary to request that force constants be generated.<sup>11</sup> As 10- and 30-step calculations in both directions failed to unambiguously connect the optimized reactant and product structures, a 100-step calculation was run. Results from the IRC calculation of the OST2 structure still appeared suspect so an alternative approach to obtaining a transition-state structure, a general transition-state optimization, using the Berny algorithm, was run on the optimized transition-state structure provided by the OST2 method.<sup>14</sup> Frequency calculations were run on this structure as well to confirm that it was a first-order saddle point. An IRC calculation was run on this structure as well in an attempt to connect the reactants and products through this transition structure.

# **Data and Results**

The lowest energy optimized geometry of 2.3.10trioxabicyclo[5.2.1]decan-4-ol is shown in Figure 4 on the left. The values for the energy,  $\Delta E_{r, elec}$ , at each optimization step in the dihedral angle, relaxed PES scan in hartree and kcal/mol relative to the most stable optimized form are included alongside their corresponding angle in Table 1. The dihedral angle was measured from the hydroxyl hydrogen labeled 12 to the peroxide oxygen labeled 9 in Figure 4. The energy at the angle of -57.2° corresponded to the energy of the optimized structure in Figure 4. The graph of the potential energy well produced by this calculation is included in Figure 5. The strength of the hydrogen-bonding interaction between the bridgehead oxygen and the hydroxyl hydrogen was estimated by subtracting the optimized energy from the dihedral angle scan at -57.2°, where the hydrogen-bonding interaction was maximized, from the optimized energy at  $162.8^{\circ}$ , where there was expected to be no hydrogen-bonding interaction. This value of 8.58 kcal/mol was recorded in Table 2 along with the literature value for similar, O-H.: O, hydrogen bonds for comparison.<sup>15</sup> The value calculated using the rigid scan is also included in Table 2; however, as will be elaborated on in the discussion section, this value is expected to be less accurate



**Figure 4:** Lowest energy optimized structure of 2,3,10-trioxabicyclo[5.2.1] decan-4-ol before (left) and after (right) homolytic bond cleavage

Figure 4 shows the optimized structures of 2,3,10trioxabicyclo[5.2.1]decan-4-ol before and after homolytic bond cleavage. The structures of the suspected transition states calculated using the Berny algorithm and the QST2 calculation are shown in Figure 6.

Table 1: Relaxed dihedral angle PES scan data				
Coordinate	Energy	Energy relative to reactant		
-177°	-574.933 hartree	8.49 kcal/mol		
-167°	-574.933 hartree	8.40 kcal/mol		
-157°	-574.933 hartree	8.15 kcal/mol		
-147°	-574.934 hartree	7.83 kcal/mol		
-137°	-574.935 hartree	7.21 kcal/mol		
-127°	-574.936 hartree	6.41 kcal/mol		
-117°	-574.938 hartree	5.45 kcal/mol		
-107°	-574.940 hartree	4.23 kcal/mol		
-97.2°	-574.942 hartree	3.00 kcal/mol		
-87.2°	-574.943 hartree	1.78 kcal/mol		
-77.2°	-574.945 hartree	0.75 kcal/mol		
-67.2°	-574.946 hartree	0.11 kcal/mol		
-57.2°	-574.946 hartree	0.00 kcal/mol		
-47.2°	-574.946 hartree	0.43 kcal/mol		
-37.2°	-574.945 hartree	1.12 kcal/mol		
-27.2°	-574.944 hartree	1.39 kcal/mol		
-17.2°	-574.944 hartree	1.64 kcal/mol		
-7.20°	-574.943 hartree	1.93 kcal/mol		
2.80°	-574.943 hartree	2.13 kcal/mol		
12.8°	-574.943 hartree	2.37 kcal/mol		
22.8°	-574.942 hartree	2.48 kcal/mol		
32.8°	-574.942 hartree	2.55 kcal/mol		
42.8°	-574.942 hartree	2.62 kcal/mol		
52.8°	-574.942 hartree	2.74 kcal/mol		
62.8°	-574.941 hartree	3.06 kcal/mol		
72.8°	-574.941 hartree	3.50 kcal/mol		
82.8°	-574.940 hartree	4.11 kcal/mol		
92.8°	-574.939 hartree	4.88 kcal/mol		
103.8°	-574.937 hartree	5.65 kcal/mol		
113.8°	-574.936 hartree	6.45 kcal/mol		
122.8°	-574.935 hartree	7.18 kcal/mol		
132.8°	-574.934 hartree	7.77 kcal/mol		
142.8°	-574.933 hartree	8.21 kcal/mol		
152.8°	-574.933 hartree	8.43 kcal/mol		
162.8°	-574.933 hartree	8.58 kcal/mol		
172.8°	-574.933 hartree	8.55 kcal/mol		



Figure 5: Potential energy well for the relaxed dihedral angle PES scan

Table 2: Computational results and literature comparisons				
Computational estimate of hydrogen Bond	11.5 kcal/mol			
Strength via the rigid scan				
Computational estimate of hydrogen Bond	8.58 kcal/mol			
Strength via the relaxed scan				
Literature value <sup>7</sup> of the typical $O-H$ :O	5.0 kcal/mol			
interaction				
$\Delta G^{\ddagger}$ trioxabicyclo[5.2.1]decan-4-ol for the	19.6 kcal/	81.9 kJ/		
QST2 method	mol	mol		
$\Delta G^{\ddagger}$ trioxabicyclo[5.2.1]decan-4-ol using	40.6 kcal/	172 kJ/		
the Berny algorithm	mol	mol		
Literature value <sup>8</sup> of $\Delta G^{\ddagger}$ of $H_2O_2$	43.7 kcal/	182 kJ/		
	mol	mol		
Literature value <sup>8</sup> of $\Delta G^{\ddagger}$ of RO-OR bonds	31.8 kcal/	133 kJ/		
	mol	mol		
<i>k</i> trioxabicyclo[5.2.1]decan-4-ol for the	$0.03 \text{ s}^{-1}$			
QST2 method				
<i>k</i> trioxabicyclo[5.2.1]decan-4-ol using the	$1*10^{-17} \text{ s}^{-1}$			
Berny algorithm				
Literature estimate <sup>8</sup> of $k$ of O-O bond	$6*10^{-12}s^{-1}$			
cleavage				

The frequency calculations carried out on these structures revealed that one and only one imaginary frequency was present for each method. The existence of this imaginary frequency indicated that these structures were both first-order saddle points consistent with them corresponding to transition-state structures. The imaginary frequency for

both structures corresponded to a stretching motion between the two oxygen atoms of the peroxide-the expected motion leading to cleavage of this bond. Although the existence of one imaginary frequency showed that the two structures in Figure 6 were both first-order saddle points, IRC calculations did not conclusively connect structures similar to the reactant and product of the cleavage for the QST2 method. The IRC for the Berny algorithm transition state structure showed a reactant similar to the lowest energy conformer found beforehand on the reactant side, but it did not find a minimum similar to the product that was expected. The reactant structure found by the IRC calculation appeared to be the same as that of the suspected structure but with the hydroxyl rotated so that the hydrogen was interacting with the peroxide oxygen as shown in Figure 7. This apparent rotation is consistent with a twostep process: the first step is the rotation of the hydroxyl to stabilize the cleavage process and the second step is the cleavage process. This result supports that this structure connects the reactant to the transition state. The molecular motion proceeding in the direction of the product appears to show the appropriate stretching of the peroxide bond, but an unexpected stretching of the hydrogen bonded to the oxygen in the hydroxvl also occurred. This stretching was taken as a potential indication of another step in the reaction where the hydroxyl moved for a more favorable reaction for the cleavage. Since transition-state theory assumes a quasi-equilibrium between the reactant and the transition state, the actual structure of the product generated by the IRC calculation is not necessary for rate determination.<sup>4</sup>



**Figure 6:** Transition state calculated with QST2 (left) and the Berny algorithm (right).



Figure 7: Optimized reactant found at the endpoint of the IRC calculation for the Berny Algorithm

The calculated change in Gibbs free energy of formation,  $\Delta G_{f,}$ , for the suspected reactants and products of the homolytic cleavage process shown in Figure 3, the  $\Delta G_f$  for the optimized reactant structure found by the IRC calculation for the transition state calculated by the Berny optimization shown in Figure 7, and the  $\Delta G_f$  for the transition states shown in Figure 6 are listed in Table 3. This formation is not what is traditionally referred to as a molecule's formation from its elements in their reference states but rather as calculated by the Gaussian software<sup>12</sup> and corresponds to the formation of reactant, transition state, and product from separate, noninteracting electrons and nuclei. To make the numbers more manageable, the  $\Delta G_f$  of the most stable optimized structure of 3,10-trioxabicyclo[5.2.1]decan-4-ol before cleavage was defined as zero and all other  $\Delta G_f$  values were defined relative to this reactant energy.

The activation barrier,  $\Delta G^{\ddagger}$ , was calculated by subtracting the free energy of the reactant from the free energy of each of the transition states. Because the Gibbs energy is a state function, this calculation

Table 3: Gibbs free energy of formation values in milihartree					
and kcal/mol relative to the optimized reactant					
Reactant $\Delta G_{f}$	0 mhartree	0 kcal/mol			
IRC Reactant $\Delta G_{f}$	0.5 mhartree	0.3 kcal/mol			
Product $\Delta G_f$	23.1 mhartree	14.5 kcal/mol			
Berny algorithm transition	65.3 mhartree	41.0 kcal/mol			
state $\Delta G_{f}$					
QST2 transition state	31.2 mhartree	19.6 kcal/mol			

gives the same result as calculating this value from traditional formation values. In the case of the Berny algorithm, the reactant used was the optimized structure at the end of the IRC calculation. The values of  $\Delta G^{\ddagger}$  are included in Table 2 with a comparison to accepted values of  $\Delta G^{\ddagger}$  for hydrogen peroxide and organic R-O-O-R peroxides calculated from literature values  $\Delta S^{\ddagger}$  and  $\Delta H^{\ddagger}$  at 298 K with the relation,  $\Delta G^{\ddagger} =$  $\Delta H^{\ddagger} - T\Delta S^{\ddagger.8}$  The first-order rate constant, k, was determined for the cleavage process using the Eyring–Polanyi equation and the  $\Delta G^{\ddagger}$  values for both suspected transition states.<sup>9</sup> A literature estimate of the k value typical of the homolytic cleavage process for stable peroxides is also included in Table 2.

#### Discussion

The calculated value of the strength of the intramolecular hydrogen bond between the bridgehead oxygen and the hydroxyl group of 8.58 kcal/mol appeared to be much more significant than the literature value for similar hydrogen bond interactions between hydroxyl groups and lone pair electrons on oxygen atoms of 5.0 kcal/mol.<sup>15</sup> However. this calculation is most likely an overestimate. This discrepancy arises because the potential energy scan measures molecular energy, and this energy changes based on the locations of all of the atoms in the molecule. Steric and torsional contributions to the energy must be considered. As the hydroxyl dihedral angle is changed, there is potential for overlapping electron domains. The relaxed scan minimizes—but does not eliminate-potential steric and torsional clashes by allowing the rest of molecule, besides the coordinate that is scanned, to optimize to its lowest energy. Thus, the actual hydrogen-bond energy is likely much closer to the typical 5 kcal/mol value with the additional 3 kcal/mol arising from steric factors. These steric effects are also likely why the strength of the hydrogen bond predicted by the rigid scan, 11.5 kcal/mol, is much larger than was predicted with the relaxed scan and much further away from the literature value for similar hydrogen bonds. The typical error in density functional theory measurements hydrogen bond strength is 1 kcal/mol.<sup>6</sup> This error is for intermolecular hydrogen bonds not intramolecular hydrogen bonds so it may not accurately represent the error of the estimate in the experiment.

The transition state for homolytic bond cleavage predicted by the Berny algorithm is suspect in light of the suspiciously small rate constant. The first-order rate constant at 298 K was predicted to be  $1 \times 10$ -17 s-1. This is significantly slower than the literature estimate by Bach, et. all, of  $6 \times 10$ -12 s-1 for organic peroxides.<sup>8</sup> This calculated value is similar to the first-order rate constant of the decomposition of uranium-

238 based on its half-life of 4.5 billion years is  $5 \times 10-18 \text{ s}-1$ .<sup>16</sup> This suspiciously slow rate constant is potentially due to the fact that the Eyring-Polanyi equation only takes into account the height of the activation barrier without consideration of densities of state.<sup>9</sup> The activation energy of the peroxide calculated using the Berny algorithm of 40.6 kcal/mol appears comparable to that of hydrogen peroxide with an activation energy of 43.7 kcal/mol; however, for peroxides with nonhydrogen substituents, this activation energy is significantly higher than the typical literature values of about 32 kcal/mol.<sup>8</sup> The activation energy of 19.6 kcal/mol calculated using the QST2 calculation has a much worse agreement with the literature values of other peroxides. The value for the first-order rate constant determined using this method is significantly farther from the literature value estimate of 6×10-12s-1.8 This observation combined with the IRC calculations that did not show any correspondence to the expected stretching cast serious doubt on the legitimacy of the QST2 calculated transition state structure. Based on these comparisons and that the IRC of the Berny algorithm structure connected the reactant to the transition state, it seems the Berny algorithm structure is more likely. Based on these results, it appears that this peroxide will undergo homolytic bond cleavage at a rate slower than typical RO-OR peroxides and is comparable to hydrogen peroxide, a relatively stable peroxide.

None of the calculations performed in this experiment took into account solvent effects. This absence of a solvent could contribute significant error to the calculations performed. If the molecule has a polar protic solvent, there is a possibility that the intramolecular hydrogen bond strength may be much less. In addition, solvents often play roles in stabilizing transition states and lowering activation energies. The solvent could also potentially catalyze the cleavage. Another source of error in this experiment is that not every possible conformation of this molecule was considered in calculation. The structure shown in Figure 3 before homolytic bond cleavage was the minimum found by the optimization calculation based of the initial geometry it was given, but it is possible that there are other conformations at lower energies.

## Conclusion

Based on the calculations performed here 2,3,10trioxabicyclo[5.2.1]decan-4-ol appears to be a relatively stable peroxide. An intramolecular hydrogen bond between the hydroxyl group and the bridgehead oxygen appears to contribute 8.58 kcal/mol to the stability, which is more significant than the expected value for similar interactions of 5.0 kcal/mol. The rate of homolytic bond cleavage in the gas phase of this peroxide is expected to be similar to that of hydrogen peroxide based on transition state calculations. Attempting to synthesize this compound appears worthwhile based on these estimates of stability, though further studies of solvated stability would be informative.

# References

(1) Park, B. K.; O'Neill, P. M.; Maggs, J. L.; Pirmohamed, M. Safety Assessment of Peroxide Antimalarials: Clinical and Chemical Perspectives. *Br. J. Clin. Pharmacol.* **1998**, *46* (6), 521–529.

(2) Luo, X.-D.; Shen, C.-C. The Chemistry, Pharmacology, and Clinical Applications of Qinghaosu (Artemisinin) and Its Derivatives. *Med. Res. Rev.* **1987**, 7 (1), 29–52.

(3) Shandilya, A.; Chacko, S.; Jayaram, B.; Ghosh, I. A Plausible Mechanism for the Antimalarial Activity of Artemisinin: A Computational Approach. *Sci. Rep.* **2013**, *3*.

(4) Atkins, P.; de Paula, J. *Atkins' Physical Chemistry*; Oxford University Press: Oxford [etc.], 2006.

(5) Mole, S. J.; Zhou, X.; Liu, R. Density Functional Theory (DFT) Study of Enthalpy of Formation. 1. Consistency of DFT Energies and Atom Equivalents for Converting DFT Energies into Enthalpies of Formation. *J. Phys. Chem.* **1996**, *100* (35), 14665–14671.

(6) Ireta, J.; Neugebauer, J.; Scheffler, M. On the Accuracy of DFT for Describing Hydrogen Bonds: Dependence on the Bond Directionality. *J. Phys. Chem. A* **2004**, *108* (26), 5692–5698.

(7) Plumley, J. A.; Dannenberg, J. J. A Comparison of the Behavior of Functional/Basis Set Combinations for Hydrogen-Bonding in the Water Dimer with Emphasis on Basis Set Superposition Error. *J. Comput. Chem.* **2011**, *32* (8), 1519–1527.

(8) Bach, R. D.; Ayala, P. Y.; Schlegel, H. B. A Reassessment of the Bond Dissociation Energies of Peroxides. An Ab Initio Study. J. Am. Chem. Soc. **1996**, *118* (50), 12758–12765.

(9) Ochterski, J. W. Thermochemistry in Gaussian. *Gaussian Inc* **2000**, 1–19.

(10) Guner, V.; Khuong, K. S.; Leach, A. G.; Lee, P. S.; Bartberger, M. D.; Houk, K. N. A Standard Set of Pericyclic Reactions of Hydrocarbons for the Benchmarking of Computational Methods: The Performance of Ab Initio, Density Functional, CASSCF, CASPT2, and CBS-QB3 Methods for the Prediction of Activation Barriers, Reaction Energetics, and Transition State Geometries. *J. Phys. Chem. A* **2003**, *107* (51), 11445–11459.

(11) Foresman, J. B.; Frisch, Ae. *Exploring Chemistry with Electronic Structure Methods*; Gaussian Inc.: Pittsburgh, PA, 1996.

(12) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, USA, 2009.

(13) G09 Keyword: IRC http://www.gaussian.com/g\_tech/g\_ur/k\_irc.htm (accessed Oct 20, 2016).

(14) G09 Keyword: Opt http://www.gaussian.com/g\_tech/g\_ur/k\_opt.htm (accessed Oct 20, 2016).

(15) Emsley, J. Very Strong Hydrogen Bonding. *Chem. Soc. Rev.* **1980**, *9*(1), 91–124.

(16) Jaffey, A. H.; Flynn, K. F.; Glendenin, L. E.; Bentley, W. C.; Essling, A. M. Precision Measurement of Half-Lives and Specific Activities of 235U and 238U. *Phys. Rev. C* **1971**, *4*, 1889–1906.

# Lyman-α Emission from an Infant Black Hole in the Early Universe

# Brandon K. Wiggins,<sup>1,2</sup> Joseph M. Smidt,<sup>2</sup> Jarrett L Johnson<sup>2</sup>

<sup>1</sup>Southern Utah University, <sup>2</sup>Center for Theoretical Astrophysics, Los Alamos National Laboratory

#### Abstract

The COSMOS survey recently discovered an exotic young galaxy, COSMOS Redshift 7 (CR7), in the early universe (1 billion years after the Big Bang), which is devoid of evidence of elements heavier than hydrogen and helium. Whereas some believe this might be the first galaxy discovered with stars made only from these elements, others think CR7 may be powered by a newborn supermassive black hole. In this paper, we summarize for a general academic audience our efforts to model the creation of this galaxy through cosmological simulations. These state-of-the-art calculations include primordial chemistry and cooling and the interaction of x-rays from the black hole with surrounding gas. We simulate the process of light escaping this object with Monte Carlo Lyman-a transfer and compare our calculations with observations of CR7. Our work demonstrates the viability of the black hole interpretation for this intriguing object in the early universe.

# 1. Introduction

At the heart of virtually every large galaxy resides a supermassive  $(>10^6$  solar masses) black hole. Their presence in their host galaxies can be manifest by the orbits of stars, in brilliant, high-energy emission near a galaxy's center, or in the presence of spectacular, kiloparsecscale jets that can rival the extent of the host galaxy itself. Large black holes can be the *brightest* objects in the universe; as black holes "eat" matter (a process called *accretion*), the infalling material heats up and glows. High accretion rates of massive black holes power ultraluminous quasars that are visible from across the cosmos.

The impressive size of massive black holes bespeaks their ancient origin. The now-supermassive black holes that reside in the hearts of galaxies must have begun life much smaller in the distant past, gradually accreting material, growing through cosmic time to become the monsters observed today. The question of where they came from may initially seem obvious. One might assume that a very large star in the early universe ended its life with the collapse of its core, leaving behind a stellar-sized (with a mass from about 1 to 100 suns) black hole. Such a star may have served as a *massive black hole seed* that may have subsequently accumulated material through billions of years to emerge in the modern universe as a supermassive black hole.

This seemingly reasonable origin story is fraught with its difficulties, however. If a black hole seed was born in the cataclysmic collapse of a massive star, the resulting supernova would have likely pushed away the young black hole's gas supply, and the object is liable to have been "born starving." Indeed, some primordial halos in which these stars were born may have even been completely destroyed in the explosion (Whalen et al. 2008). It is further unclear on what timescale a sizeable gas supply would be available to fuel the black hole's growth. The second major issue is that black holes cannot feed arbitrarily fast. Infalling material will heat up and gives off light as it falls toward the black hole and interacts with additional infalling matter. If the accretion rate becomes sufficiently large, the radiation from the infalling material will become so substantial that it will begin to impart outward momentum to the accretion flow. This radiation pressure can ultimately impede the infall of matter onto the black hole. Thus, there exists a well-defined accretion rate limit (called the *Eddington Accretion Rate*)<sup>1</sup> where the

<sup>&</sup>lt;sup>1</sup> The Eddington Accretion Limit technically describes spherical accretion, but the general principle, that radiation can limit accretion rates, is applicable generally. Thick (radiation-dominated) accretion disk theory and accompanying numerical simulations do not place a firm upper limit on accretion rates with some authors achieving rates orders of

rate of infalling matter and outward radiation pressure balance. It can be shown that this terminal accretion rate is related to the accretor's mass, with larger accreting objects being capable of accreting faster. Nevertheless, black holes are limited in the rate at which they can grow even in the presence of a large gas supply.

While these two lines of reasoning may strongly suggest that the possibility of growing a supermassive black hole from a small stellarsized black hole is difficult, data from perpetually deeper surveys render it all but impossible. In 2011, a young quasar, ULAS J1120+0641, now known to the community as the *Mortlock Quasar*, was observed at redshift z = 7.1 (roughly 1 billion years after the Big Bang) with an estimated mass of 1 billion suns (Mortlock et al. 2011). Last year, the supermassive black hole in SDSS J010013.02+280225.8, a quasar observed from redshift z = 6.1 (about 1.2 billion years after the Big Bang), weighed in at 12 billion solar masses (Wu et al. 2015). The scenario where a stellar-sized black hole produced around redshift z = 15, or 500 million years after the Big Bang, accretes continuously to achieve masses of millions or billions of suns struggles strongly to explain these observations. There must surely be alternate pathways to growing massive black holes.

The origin of supermassive black holes isn't merely interesting on its own merits. Understanding the histories of supermassive black holes sheds light on processes of large-scale structure formation in the early universe in general. How the universe makes very large things is still an active area of research, and the growth of massive black holes appears to be intimately connected to the formation of larger structures, including galaxies. Cryptic but robust correlations between black hole masses and the velocities of orbiting stars in galaxies strongly suggest that massive black holes and their host galaxies have evolved together. Understanding the history of massive black holes may offer hints as to the processes involved in assembling large-scale structure in the cosmos.

#### 1.1 Making a Monster in the Early Universe

The universe in which black hole seeds, the forerunners of supermassive black holes, would likely have first appeared differed strongly from the cosmos we observe today. This was a period after the universe awoke from general cosmic blackness (formally known as the Cosmic Dark Ages) and well after it began to be illuminated with the first generation of stars. No large-scale structure, no grand-design spiral galax-

magnitude larger than that of Eddington (e.g., Jiang, Stone & Davis 2014). Nevertheless, the Eddington Limit appears to be respected in nature, by and large.

ies or super clusters of galaxies existed during this period; gravitational perturbations had not yet grown to give rise to cosmological halos much larger than  $10^6$  solar masses. If one could visit this period in the universe, they would only see collections of very bright stars gradually collecting in stellar communities under the influence of gravity.

The material from which this first generation of stars was created was chemically simpler than the gas and dust observed throughout the universe today. Before stars formed, the baryonic universe consisted only of hydrogen, helium, and trace amounts of lithium. Structures in the early universe would have formed via collapse of this "pristine" gas. To collapse, however, halos must cool, converting heat to light, which ultimately escapes the system. This collapse process would have differed from analogous collapses of molecular cores in the modern universe. Clouds of pure hydrogen and helium do not cool well compared with clouds of gas enriched with heavier elements. This is because hydrogen and helium have very simple electron structures compared with heavier elements on the periodic table, which represent a comparatively small number of energy states that can be collisionally excited. Because the halos cannot cool well, they are on average larger before collapse, a trend that could give rise to stars more massive than those found in the modern universe. In practice, one expects such halos to grow until virial temperatures reach ~1000 K. In the meantime, there is formation of trace amounts of molecular hydrogen, which has a series of infrared vibrational states that are collisionally excited at these temperatures within the halo. The molecular hydrogen, even in its small amounts, acts as an effective coolant, and the halo cools catastrophically and collapses. The stars emerging from this process are called "population III" stars as they are formed from materials devoid of elements heavier than hydrogen and helium. Thus far, there has been no observational evidence for population III structures. They have eluded surveys and deep searches because they lie literally at the edge of the observable universe, when the universe was only a few hundred million vears old.

Halos with masses of  $10^6$  solar masses are large, but scenarios can be contrived where halos grow much larger before collapse. If a halo in the early universe was kept free from heavy elements even as surrounding halos underwent star formation, these surrounding halos could provide a background of Lyman-Werner photons (or photons with the energy required to dissociate molecular hydrogen), which could keep the metal-free halo also free of molecular hydrogen. If this takes place, nothing is to stop the halo from growing to a mass of ~ $10^8$  solar masses corresponding to virial temperatures of 10,000 K. At these temperatures, the Lyman series becomes collisionally excited, and the massive halo cools through Lyman- $\alpha$ , triggering a catastrophic collapse.

The subsequent evolution of the gargantuan collapsing halo has been the subject of numerous studies (e.g. Rees 1984, Eisenstein & Loeb 1995, Begelman et al. 2006, Begelman 2010, Ferrara et al. 2014), some suggesting that the collapsing cloud gives rise to a supermassive  $(10^5 \text{ solar mass})$  star (e.g., Johnson et al. 2013) that quickly (on the timescale of Myr) implodes to form a similarly massive black hole. Some work suggest that such a supermassive star could in rare circumstances undergo a general-relativistic instability and explode with 10,000 times the energy of canonical core collapse events (Chen et al. 2014). Other studies have examined the collapse of the cloud directly to a black hole without assuming an intermediate stellar-like stage of activity (e.g., Ferrara et al. 2013). Whether such a star forms and subsequently explodes to become a black hole or whether only a black hole forms in the collapsing cloud, a very large, 10<sup>5</sup> solar mass black hole emerges from the process, which could serve as a very effective seed for supermassive black holes observed 1 Gyr after the Big Bang.

The scenario we have just proposed is called the *direct collapse hypothesis*. Alternate scenarios have been proposed, including collisional run-aways in stellar clusters or the growth of primordial black holes (for summaries, see Latif & Ferrera 2016, Regan and Haehnelt 2009). In each proposed process, the goal is the same: create larger black hole seeds so that the black hole has sufficient time to grow to masses observed in quasars in deep surveys.

Until recently, there has been no direct observational evidence for population III structures or the direct collapse scenario. Population III stars are too distant to observe and direct collapse black holes have existed only as a theoretical construct to achieve desired quasar masses. But recent surveys (Sobral et al. 2015) have brought to light an exotic object, one that has offered theorists the very first glimpse into this exotic era.

## 1.2 Luminous Lyman-α Emitter Cosmos Redshift 7

It is a somewhat rare thing when observations arise to confirm the musings of theoreticians. In 2015, Sobral et al. reported the finding of a young, luminous galaxy at z = 6.6 (about a billion years after the Big Bang), which contains no evidence for elements other than hydrogen and helium. The object was a brilliant Lyman- $\alpha$  emitter ( $L_{lya} > 8.44e43$  erg s<sup>-1</sup>), evidence for very hot or strongly ionized hydrogen. It was also particularly luminous in the HeII 1640 Angstrom line ( $L_{He} = 2.0e44$  erg s<sup>-1</sup>) emission, which was further indicative that the luminosity of the

object was powered by a particularly hot light source, emitting photons well in the ultraviolet and x-rays. The Lyman- $\alpha$  line is 160 km s<sup>-1</sup> offset redward of the HeII emission and has a width of ~260 km s<sup>-1</sup>. The object has been designated "Cosmos Redshift 7," or "CR7," after the COSMOS survey in which it was discovered. This unprecedented find, a young galaxy with no evidence for heavy elements, could suggest that CR7 is powered by luminous population III stars, which would herald the very first pristine stellar population ever detected. Yet a substantial portion of the community wonders if CR7 could represent a different type of population III structure. As with stars in the local cosmos, the lifetime of massive stars in the early universe is highly dependent upon their masses (massive stars living much shorter lives), with some estimates putting population III stars at up to hundreds of solar masses (Hirano et al. 2014). Following life, such stars undergo spectacular supernova explosions that seed their surroundings with heavy elements (see Wiggins et al. 2015 for background). If these constraints on population III stellar lifetimes exist, it becomes unlikely that observers fortuitously observed CR7 during a brief phase when it was free from enrichment from supernovae. A lack of heavy elements (which indicates a lack of supernovae) also places constraints on how fast star formation must have taken place: All the stars powering CR7 must have formed during the few million years before the first supernova, a surprisingly rapid star formation rate (Hartwig et al. 2015). The strong helium emission exposes another problem. HeII 1640 Angstrom emission signals ionized or collisionally excited helium, which is indicative of an ionizing spectrum much more energetic than what could be provided by any stellar population, no matter how massive (Hartwig et al. 2015). These together appear to point to an alternate power source for CR7: an accreting black hole. A very bright central source of such an object could heat the cloud, forbidding instances of protostellar cooling and collapse. In this way, CR7's strong radiation field may be keeping it metal free by forbidding star formation to take place. There appears to be additional evidence, however, that CR7 may be the first observed instance of a direct collapse black hole asserting its influence on the halo in which it was born.

CR7 consists of three major "subclumps." Its central, major clump appears to be most luminous in Lyman- $\alpha$ , with a neighboring clump more luminous in the infrared. An understood prerequisite for direct collapse is the production of a Lyman-Werner background, which could be provided by a neighboring halo with massive population III stars. CR7's smaller subclumps seem fortuitously positioned to provide such a background, which would have kept the central halo of CR7 H<sub>2</sub> free, allowing it to grow to the large masses required for direct collapse black hole formation. The combination of observables—the conveniently positioned subclumps, the brilliant HeII 1640 Angstrom emission that indicates a hard x-ray source at its center, and the metal free gas lend themselves to a convenient interpretation: CR7 may be powered by a direct collapse black hole.

In this paper, we summarize for a broad audience our recent efforts to model emission from the CR7 system assuming it hosts a direct collapse black hole. For an abbreviated, technical treating of this study with a slightly different focus for our own community, we encourage the reader to consult Smidt, Wiggins & Johnson (2016). The current paper contains details that we do not have the luxury of expanding on in this former work. Numerous additional studies have attempted to model emission from this system (e.g., Smith et al. 2016; Dijkstra et al. 2016), however, such studies have been limited to spherical symmetry or have highly idealized models for the distribution of gas in CR7. Here we combine state-of-the-art cosmology simulations with Monte-Carlo Lyman- $\alpha$  transport to demonstrate the viability of the direct collapse black hole interpretation of CR7.

This paper is structured as follows. In section 2, we detail our method, including the methodology in our cosmological calculations and our code for calculating Lyman- $\alpha$  line profiles. In section 3, we present our simulation results, and we discuss implications for our work in section 4. We conclude in section 5.

# 2. Method

Our calculation is broken into two major portions. We first carry out a cosmological simulation, evolving a 3e10 solar mass halo down to redshift z = 6.6. We then post-process the calculations to determine emissivity in Lyman- $\alpha$  and He 1640 Angstrom emission. We discuss these stages in detail through the remainder of this section.

## 2.1 Cosmological Calculations

Our cosmological simulations were carried out in ENZO (Bryan et al. 2013) a publically available cosmological code. Initial conditions were generated in MUSIC (Hahn & Abel 2011) using the PLANCK parameters (Planck Collaboration). ENZO is an adaptive mesh refinement (AMR), grid-based fluid dynamics code. ENZO solves the Eular fluid equations with self-gravity on a box with periodic boundary conditions. Our simulation box is 4 comoving Megaparsecs on a side with a root grid of 256<sup>3</sup> cells and 11 levels of adaptive refinement. In our calculation, ENZO is coupled with MORAY (Wise & Abel 2010), a ray-tracing radiation package to model radiative feedback from the

central black hole, and GRACKLE (Bryan et al. 2013), a primordial chemistry package that follows the formation and reactions of H, H+, H-, H2, H2+, He, He+, He++, e- in addition to tracking radiative cooling from collisional excitations of these species.

In our work, we do not model all three clumps of the CR7 complex. Rather, we follow the evolution of the black hole in the central clump only and impose a global Lyman-Werner flux to simulate the radiative effects of star formation in the neighboring clumps. We also do not model self-consistently the formation of the black hole itself. When the simulation has reached z = 15, we insert a black hole sink particle with a mass of ~3000 solar masses that is allowed to accrete gas and grow down to redshift z = 6.6.

Material infalling into the black hole would undoubtedly form an accretion disk. Although a disk of sorts forms around the black hole in our model, the formal accretion disk would appear on scales much smaller than the scales resolved in our simulation. This is unfortunate as the radiative feedback, the brilliant flaring of material on the accretion disk, which is sufficiently powerful to alter the flow of infalling gas on galactic scales, cannot be modeled self-consistently. This is an unavoidable shortcoming of virtually every cosmological calculation involving a black hole. In our calculation, the luminosity of the black hole is tied to its current accretion rate via a subgrid alpha disk (e.g., DeBuhr et al. 2010). The black hole then deposits this energy into the gas by shooting a large number of rays from the point–particle into the surrounding space.

Recall that our purpose is not so much to recreate every aspect of CR7, but to demonstrate the viability of the hypothesis that it may be powered by a black hole instead of a primordial stellar population. This motivates our somewhat simple approach in modeling major characteristics of this object.

# 2.2 Calculating Emission

For our newly obtained three-dimensional model of CR7, we wish to now determine the emissivity in Lyman- $\alpha$  and HeII 1640 Angstrom line. Both types of emission can be either collisionally excited or appear during recombination, as the recombining electron cascades to the ground state. The emissivity *j* of a parcel of gas of volume *V* in Lyman- $\alpha$  integrated over frequency from recombination is just



where  ${}^{\mathbf{L}_{\mathbf{V}\mathbf{R}}}$  is the fraction of recombinations resulting in Lyman- $\alpha$  emission (~0.68),  ${}^{\mathbf{n}_{\mathbf{g}},\mathbf{n}_{\mathbf{p}}}$  are the number densities of electrons and protons, respectively,  ${}^{\mathbf{n}_{\mathbf{g}\mathbf{e}},\mathbf{n}_{\mathbf{p}}}$  is the case-B recombination coefficient, and  ${}^{\mathbf{E}_{\mathbf{L}_{\mathbf{V}\mathbf{R}}}}$  is the energy of a Lyman- $\alpha$  photon (10.2 eV). The emissivity due to collisional excitation  ${}^{\mathbf{l}_{\mathbf{c}\mathbf{e}\mathbf{l}}}$  is

# $\mathbf{j}_{coll} = \mathbf{n}_{e} \mathbf{n}_{H} \mathbf{q}_{1s2p} \mathbf{E}_{Lya} \mathbf{V},$

where  $n_{\rm H}$  is the number density of neutral hydrogen and  $q_{1520}$  is the collisional coefficient between the 1s and 2p levels. The relations provide Lyman- $\alpha$  emissivity as a function of position in our domain at redshift z = 6.6. Similar relations can be derived for HeII 1640 emission. To calculate these, we use the method in Yang et al. (2006).

Lyman- $\alpha$  is a resonant line, which means that after its production, it is liable to quickly be absorbed and re-emitted again in a random direction. In optically thick environments, or so-called damped Lyman- $\alpha$  systems, Lyman- $\alpha$  photons perform something of a random walk through the cloud before escaping. As the photons are absorbed and reemitted by various atoms, each with their own velocity with respect to the bulk motion of the gas, they become Doppler shifted, giving rise a second random walk or diffusion in frequency space. To capture this effect, we carry out a Monte Carlo Lyman-α post-process with a peeloff method (e.g., Zheng & Miralda-Escude 2002) to calculate line profiles. This attempts to model the absorption-remission "scattering" of photons through the cloud and their eventual arrival to a virtual camera situated at the edge of the simulation domain. We detail this process for the interested reader in our Appendix. For simplicity, we do not consider processes that destroy Lyman- $\alpha$  including absorption by dust, which can be reasonably ignored in an environment with very few heavy elements.

#### **3** Results and Discussion

Our CR7 object forms on a major filament that spans nearly the entirety of the simulation domain (Figure 1). At redshift z = 15, a ~3000 solar mass black hole is inserted at the location of highest density. As the black hole eats, it radiates, pushing back on the flow of infalling material. By redshift z = 6.6, the black hole has attained a mass of a few tens of millions of solar masses and feeds at rates of 90% Eddington for much of its lifetime. As one might expect, gas flows along the major filaments, feeding the central halo. The accretion scenario is far from spherically symmetric and in the ensuing turbulent

accretion, an oblate, rotating cloud forms around the central black hole (Figure 2).



**Figure 1.** Far-field slice plots of the CR7 model as calculated in our ENZO cosmological run. The CR7 object appears in the center of each plot along a filament that spans much of the simulation domain.

The black hole shines in x-rays (200 eV to 1 keV photons). This energetic radiation field ionizes gas immediately in the vicinity of the black hole, but also gas at very large distances (Figure 1, lower righthand panel). Some of this ionized hydrogen and helium will recombine giving rise to Lyman- $\alpha$  and HeII 1640 emission. Far from the black hole in the largely neutral medium, a single high-energy photon might ionize a hydrogen atom, resulting in a very energetic electron, which might go on to ionize additional hydrogen atoms. These "secondary ionizations" can be an important source of Lyman-a, and our cosmological code takes these effects into account as it calculates primordial chemistry. If such a photon ionizes an atom in a region where most of the gas is already ionized, the energetic photon ultimately gives its energy to heat the surrounding gas via electron-electron scatterings. In the lower right-hand panel of Figure 1, one notices that the region of lowdensity gas that is ionized has a ray-like character emanating from the central halo, evidence of the ray tracing radiation scheme.



**Figure 2.** The evolution of the CR7 model from z = 15 when the black hole is inserted down to z = 6.6, the redshift at which CR7 is observed. Plotted here are surfaces of constant density with brighter colors corresponding to higher densities. The halo is fed by the major filament in a turbulent, highly non-spherical process. The arrows indicate the central halo hosting the massive black hole.

In Figure 3, we plot the emissivity in Lyman- $\alpha$  as a function of position at redshift z = 6.6. The plot emphasizes regions of high ionization fraction and density, including the central CR7 halo and major filament. We point out, however, that if one could visit the early universe and photograph this halo with a Lyman- $\alpha$  "camera," this is not what one would "see." Because Lyman- $\alpha$  scatters so frequently, one would likely actually observe only a single Lyman- $\alpha$  blob near the center of the frame. The figure is instructive for illustrating the geometry of the halo and its cosmological environment. The apparent pixilation near the edge of the image is an artifact the adaptive mesh refinement; the simulation adds more resolution to regions of interest, i.e. with strong gradients in temperature or density.

The total luminosity in Lyman- $\alpha$  is a very luminous 5.0e44 erg s<sup>-1</sup>, well in excess of the lower bound set by observations (>  $8.0e43 \text{ erg s}^{-1}$ ). Even in environments free from heavy elements, it is believed that only half of this luminosity would escape in settings dominated by a black hole (e.g., Hartwig et al. 2015). A substantial portion of this emission is not only from recombination of hydrogen ionized by the central black hole, but from collisional excitations of Lyman- $\alpha$  by very hot gas. The HeII 1640 Angstrom luminosity is 2.0e43 erg s<sup>-1</sup>, in agreement with observations. Our Monte Carlo Lyman-α transfer calculation predicts a line width a bit in excess of CR7, ~280 km s<sup>-1</sup>. Our line profile varies as a function of viewing angle (as is expected for strongly asymmetric systems), but always consists of two major peaks (Figure 4). A doublepeaked line profile is expected for most Lyman- $\alpha$  emitters as photons must scatter from the line center to escape. A single, off-center peak is observed in CR7, which may be due to extinction of the bluer peak by the intergalactic medium (see Dijkstra et al. 2016).



**Figure 3.** Logarithm of Lyman- $\alpha$  emissivity from recombination as a function of position for our model of CR7 at redshift z = 6.6. Color intensity highlights regions of high ionization fraction and gas density. The CR7 object is located in the center of the domain on a major filament in our cosmological calculation. The filaments glow as light from the black hole ionizes the gas and electrons recombine with the hydrogen and cascade down its energy levels.

If one ignores the bluer peak, the red peak is indeed offset from 0 km s<sup>-1</sup>, although our offset is significantly larger than that observed in CR7. For much of the black hole's growth down to z = 6.6, the black hole accretes at 80–90% Eddington, which would give rise to a black hole luminosity of **L**<sub>BH</sub>  $\cong$  **2.43**  $\times$  **10**<sup>46</sup> erg s<sup>-1</sup>. The Lyman- $\alpha$  luminosity we calculate represents ~25% of the of the black hole luminosity, which is in agreement with estimates that up to 30% of the black hole luminosity can be converted to Lyman- $\alpha$  with the help of secondary ionizations (Baek & Ferrara 2013).

Our work demonstrates that a massive black hole such as one caused via direct collapse is likely capable of producing the emission in CR7. If CR7 strongly resembles the model we have constructed here, we can make some predictions for follow-up observations. Black holes are hard x-ray emitters, and a substantial amount of x-rays leave the system without ionizing or heating the gas. Observing CR7 in the x-ray might conclusively diagnose CR7's power source. CR7 may further be a luminous radio emitter if it is powered by a black hole. Our black hole is accreting in so-called "quasar mode" (e.g., Dubois et al. 2011) at very large fractions of Eddington. Because we do not model all scales of the accretion flow (including the formal accretion disk), we cannot self-consistently make a radio prediction. However, objects accreting at near-Eddington rates can possess thick accretion disks that are supported by internal radiation pressure. These are predicted to launch powerful jets, which could be visible in the radio.



Figure 4. Lyman- $\alpha$  line profiles as computed by our Monte Carlo transfer code. Line profiles from three different orientations are displayed in various colors. All spectra are double-peaked with the relative strength of the peaks depending on viewing orientation.

Using scaling relations in Meier (2001) that loosely relate black mass, accretion rate, and radio power, we predict radio luminosities to be anywhere from ~1.0e40 erg s<sup>-1</sup> to ~1.0e45 erg s<sup>-1</sup>, depending on choices of black hole spin and whether the accretion disk is thick or thin. The radio luminosities resulting from thick accretion disks and resulting jets would be visible to the JVLA and may already appear in archival data from the COSMOS survey. Luminous radio emission would likely conclusively identify the power source of CR7 to be a massive black hole, as luminous synchrotron emission would likely only otherwise be expected from supernova remnants that likely do not exist in CR7 from the lack of observed heavy elements.

One may wonder how Lyman- $\alpha$  and HeII 1640 Angstrom emission change when the spectrum of the black hole or mass of the initial black hole particle is varied. We ran a grid of 12 models to explore sensitivity to these conditions. We find that luminosities in both lines scale with black hole masses up to approximately seed sizes of 10,000 solar masses. Beyond this mass, the black hole is liable to have blown out so much gas and dust by z = 6.6 through radiation pressure that emissivities drop accordingly: There is simply no gas left to shine. In general, we also find that a harder black hole spectrum yields stronger line emissivities. There are likely two effects at play here. The more energetic the photon, the more likely this photon is to escape to larger radii. These energetic photons will ionize atoms far from the black hole in largely neutral gas, which increases the efficiency of secondary ioni-

zations. We have only assessed these trends for a single snapshot at z = 6.6, so we cannot yet interpret them in the context of the general evolution of the halo. Understanding the time-dependence of the emission as a function of black hole mass and emission spectrum will allow for more robust interpretation of emission from CR7-like objects. Such a detailed study is currently in preparation.

## **4** Conclusions

CR7 offers our first direct observation of a population III structure and may provide critical insight into the nature of very young galaxies prior to enrichment by heavy elements. This intriguing object may be the birthplace of a young direct collapse black hole, and our work provides proof of principle that a massive black hole could be powering Lyman- $\alpha$  and HeII 1640 Angstrom emission in CR7. Follow-up in the x-ray and radio are required to conclusively diagnose the power source for this object, but the COSMOS observations may be offering the astronomical community a tantalizing glimpse of the exotic beginnings of the largest black holes in the universe.

# References

Baek, S., Ferrara, A., 2013, MNRAS, 432, L6

Barnes, L., 2009, PhD. Diss. Cambridge

Begelman, M. C., 2010, MNRAS, 402, 673

Begelman, M. C., Volonteri M., Rees, M. J., 2006, MNRAS, 370, 289

Bryan, G. L, et al. 2013, ApJS, 211, 219

Chen, K.-J., Heger, A., Woosley, S., Almren, A., Whalen, D. J., Johnson, J. L., 2014, ApJ, 790, 162

DeBuhr, J., Quataert, E., Ma, C.-P., Hopkins, P., 2010, MNRAS, 406, L55

Dijkstra, M., 2014, Pub. Astron. Soc. Australia, 31, e040, doi:10.1017/pasa.2014.33

Dijkstra, M., Gronke, M., Sobral, D., 2016, arXiv e-prints:1602.07695

Dubois, Y., Devriendt, J., Slyz, A., Teyssier, T., 2011, arXiv:1109.1457

Einenstein, D. J., Loeb, A., 1995, ApJ, 443, 11

Ferrara, A., Haardt F., Salvaterra, R., 2013, MNRAS, 434, 2600

Ferrara, A., Salvadori, S., Yue, B., Schliecher D., 2014, MNRAS, 443, 2410

Hahn, O., Abel, T., 2011, MNRAS, 415, 2101

Harrington, J. P. 1973, MNRAS, 162, 43

Hartwig T., Grover, S. C. O., Klessen, R. S., Latif, M. A., Volonteri, M., 2015, MNRAS, 452, 1233

Hirano, S., Hosokawa, T., Yoshida, N. et al., 2014, ApJ, 781, 60

Jiang Y.-F., Stone, J.M., Davis, S., 2014, ApJ, 796, 106

Johnson, J. L., Whalen, D. J., Li, H., Hol, D. E., 2013, ApJ, 771, 116

Latif, M. A., Ferrara, A., 2016, arXiv preprint, arXiv:1605.07391

Meier, D. L., 2001, ApJ, 548, L9

Mortlock, D. J., Warren, S. J., Venemans, B. P. et al. 2011, Nature, 474, 616

Rees, M. J., 1984, Ann. Rev. Astron. Astrophys, 22, 471

Regan, J. A., Haehnelt, M. G., 2009, MNRAS, 393, 858

Smidt, J., Wiggins, B. K., Johnson, J. L., 2016, arXiv e-prints: 1603.00888

Smith, A., Bromm, V., Loeb, A., 2016, arXiv e-prints:1602.07639

Sobral, D., et al. 2015, ApJ, 808, 139

Whalen, D., van Veelen, B., O'Shea, B. W., Norman, M. L., 2008, ApJ, 682, 49

Wiggins, B. K., Smidt, J. M., Whalen, D. J., Even, W. P., Fryer C. L., Migenes, V., 2015, J Utah Acad Sci Arts Lett, 797, 97

Wise, J., Abel, T., 2011, MNRAS, 414, 3458

Wu, X.-B., Wang, F., Xiaohui, F., et al. 2015, Nature, 518, 512

Yang, Y., et al. 2006, ApJ, 640, 539

Zheng, Z., Miralda-Escude, J., 2002, ApJ, 578, 33

# Appendix: Lyman-a Transfer

Our method in transferring Lyman- $\alpha$  radiation through the simulation domain is identical to the method outlined in explicit detail in Barnes (2009). For an excellent review on Lyman- $\alpha$  transfer applied to galaxies in early universe, see Dijkstra (2014). In this section, we briefly summarize the basics of Lyman- $\alpha$  transfer or orient the reader.

Monte Carlo radiative transfer simulates a radiation field by breaking into particles or photon "packets." The calculation then follows the subsequent scattering of the particles through the medium, keeping track of changes in photon packet frequency. Because we are simulating steady-state emission from a single data "snapshot" of our CR7 model at redshift 6.6, a few simplifying assumptions can be made: (1) we can neglect light travel times of photon packets and worry only about the random walk of the photon through the cloud, and (2) photon packets need only be launched once from emitting regions once and not continuously throughout the calculation. In our work, all photons packets have equal "weight," meaning that the radiation field is divided equally between the particles. This also means that when the photon packets arrive at the virtual camera at the edge of the simulation domain, each photon packet contributes equally to the spectrum.

In Lyman- $\alpha$  transfer calculations, frequency v is represented in Doppler units *x*, i.e.,

$$x = \frac{v - v_0}{v_D}$$

for  $\mathbf{v}_0 = 2.466 \times 10^{15}$  Hz and  $\mathbf{v}_D = \frac{\mathbf{v}_t \mathbf{v}_0}{c}$  with  $\mathbf{v}_t = \sqrt{\frac{2\mathbf{k}_B T}{\mathbf{m}_p}}$ . The optical depth of a Lyman- $\alpha$  photon is

#### $\tau = \int \sigma(s) n_{\rm H}(s) \, \mathrm{d}s,$

where  $n_{\rm H}$  is the local number density of neutral hydrogen, *s* is arclength, and  $\sigma = f_{12}\pi cr_{\circ} \Phi(x)$  for oscillator strength  $f_{12} \approx 0.42$ , electron radius  $r_{\circ}$  and line profile  $\Phi(x)$ , which we take to be a piecewise-defined Voigt profile with a Gaussian core and Lorentz wings, as is the norm in the literature. The transfer calculation consists of 5 basic steps.

1. *Randomly determine how far the photon packet will travel before it is absorbed.* In reality, we are drawing the optical depth that the photon will penetrate before its next scattering. Optical depth is summed along the trajectory of the photon until this optical depth

is achieved. As the photon passes through each cell, its frequency is boosted into the rest frame of the fluid.

- 2. Determine a random direction for re-emission. At this stage, the photon has been absorbed by an atom. Atoms remain in 2p state for a very short period of time  $(A^{-1} \sim 10^{-9} s)$  before decaying again and emitting a photon of the same frequency in a random direction.
- Determine the velocity of the scattering atom. Atoms in a gas 3. possess random motions characterized by a Maxwell-Boltzmann distribution. This random motion serves to Doppler shift or change the frequency of the "scattered" photon. The frequency of the atom can also be changed by a recoil effect: The photon deposits momentum onto the scattering atom, but this has been shown to be unimportant at Lyman- $\alpha$  frequencies (Zheng & Miralda-Escude 2002). A component perpendicular to the photon's trajectory is chosen from a Gaussian distribution, and the component of the atom's velocity parallel to the photon's trajectory is chosen from a distribution that characterizes the probability of an atom having a particular parallel velocity component given that it successful scattered the photon. A detailed description of this function and the rejection algorithm that was used to select parallel velocity components from it can be found in Zheng & Miralda-Escude (2002) in their Appendix. We use the Ho definition in Barnes (2009) to complete this scheme.
- 4. Scatter the atom in frequency space. The combination of the photons change in direction and the velocity components of the scattering atom determine the exiting photon's frequency. This function, called the redistribution function, takes the form  $R(x_{in}) = x_{in} + \frac{\vec{v}_a \cdot (\vec{n}_{out} \vec{n}_{in})}{v_t}$ where  $\vec{v}_{a'} \vec{n}_{out'} \vec{n}_{in}$  are vectors,

 $\vec{v}_a$  being the velocity of the atom, and  $\vec{n}_{outv}\vec{n}_{bet}$  being the outbound and inbound directions of the scattered photons, respectively. We next deboost into the laboratory frame and go back to step 1 until the desired fraction of photons have left the simulation domain.

To determine the Lyman- $\alpha$  spectrum or line profile, we use a peeloff method. At every scattering, a "fake" or "tracer" photon is scattered back to the camera at the edge of the domain. Tracer photons are accumulated into a binned spectrum, but their contribution to the spectrum is weighted by their probability of making it to the camera, which is situated at the edge of the simulation domain unabsorbed. This probability is just **P** = **e**<sup>-1</sup> where  $\tau$  is the optical depth from the photon's
location of scattering to the edge of the simulation box where the camera is predetermined to be. To generate the spectra in this paper, we set up a three cameras with views down the *x*-, *y*-, and *z*-axes.

We tested our code by plotting the redistribution function for a number of incident photon frequencies in Doppler units x. This appears in Figure 5. Our results compare well with, e.g., Behrens & Niemeyer (2013), who carried out a similar test. This test is important as it demonstrates the functionality of the somewhat sophisticated rejection method required to draw valid parallel atom velocities.



**Figure 5.** Redistribution function test for temperatures of 10 K using the rejection method in Zheng & Miralda-Escude (2002) with Barnes's (2009) definition of Ho and neglecting recoil. The agreement with a similar test carried out by Behrens & Niemeyer (2013), their Figure B.1, is quite good.

We also carried out a static slab test. In this test, photons are placed on the x,y plane in the line center  $v_0 = 2.466 \times 10^{18}$  Hz. Photon packets are given a randomly oriented direction in three dimensions. The domain is held at 10 K with no bulk fluid velocity. We set the density of the gas  $n_{\rm H}$  such that the optical depth at line center  $\tau_0 = \sigma_0 n_{\rm H} \Delta x = \sigma(x = 0, 10 \text{ K}) n_{\rm H} \Delta x = 10^4, 10^8 10^6$  where  $\Delta x$  is the half-thickness of the slab. This problem enjoys an analytic solution, which was calculated by Harrington (1973) to be

$$J(x) \propto \frac{x^2}{1 + \cosh \left[\left(\frac{\sqrt{\pi^2}}{54}|x|^2}{a \tau_0}\right]\right]}.$$

This solution and the result of our code is shown in Figure 6. The agreement is acceptable with deviations appear at greater optical depths.



**Figure 6.** Static slab test. Our transfer code's solution appears in solid while the analytic solution is superimposed as a dashed curve. Innermost peaks correspond to a line-centered optical depth  $T_0$  of  $10^4$ , the middle peaks to  $10^5$ , and the outmost to  $10^6$ .

# Positive Psychology Holistic Determinants, Testosterone Treatment, and Veteran Happiness

Hannah-Lee Brau

Brigham Young University

#### Abstract

In this study, I design a survey instrument and construct a data panel from the responses of a sample of U.S. veterans. As part of the survey, I estimate the level of happiness each veteran exhibits using the Oxford Happiness Questionnaire. The Oxford scale consists of 29 questions and uses a Likert scale that ranges from a low of one to a high of six. The average is typically around 4.3. The average happiness score for my sample of 76 veterans is 3.73. After measuring the happiness level, I ask 30 additional questions driven by the literature to determine the factors of veteran happiness. Next, I conduct Spearman Correlation tests, t-tests for equality of sample mean divided on the median of the happiness score, and a multivariate ordinary least squares model with all of the explanatory factors. I find significance for four holistic happiness variables (faking happiness (+), spending money on loved ones (-), listening to music often (+), and using technology often (-)); one demographic variable (age (+)); six military-related variables (activeduty service (+), years of service (-), months deployed (-), service in Iraq (+), Afghanistan (+), Korea (+)); and two intervention variables (psychotherapy (-), exercise (+)). Testosterone treatment is not statistically significant.

#### **INTRODUCTION**

The field of positive psychology is a fairly new field. Extant research (e.g., Lopez et al., 2015; Morgenegg, 2014; Lyubomirsky, 2008) has indicated that various happiness determinants may improve depression and increase happiness levels. A somewhat controversial factor I also explore in the measurement of veteran happiness is the effect of exogenous testosterone therapy. A recent Wall Street Journal article about exogenous testosterone and potential health issues reports, "The cardiovascular concerns flared in November 2013 when a widely publicized study in the Journal of the American Medical Association (JAMA) of nearly 9,000 veterans being evaluated for coronary artery disease found that those who used testosterone had a 29% higher risk of heart attack and stroke than those who didn't. (Beck, 2014)" However, one of the results of having low testosterone in men, particularly as they age, is depression. I measure the impact of testosterone medication on depression levels, an extension of the Beck (2014) study cited above.

In the following literature review, happiness and major depressive disorder (MDD) are used somewhat interchangeably. The reason for this is because MDD is carefully defined and studied in the academic literature and can be viewed as the antithesis of happiness. In the subsequent hypotheses, each hypothesis applied to happiness could also be thought of as having the opposite predicted sign on MDD.

The remainder of the paper proceeds as follows. First, I review the literature of holistic happiness activities and testosterone treatment. Next, I review the data, which is followed by the empirical methods. Then, I discuss the results from the univariate tests and the multivariate regression model. The last section summarizes and concludes.

#### LITERATURE REVIEW

According to Barton (2015), "In 2013, the Department of Veterans Affairs (VA) estimated that about 1.5 million veterans required mental health care, including services for Major Depressive Disorder." Many veterans go through bouts of depression (Barton, 2015). Research is being conducted to discover the cause and consequences of depression. Factors may be thought of more holistically, such as listening to music or meditating, or along the physical dimension of testosterone levels. Holistic interventions such as exercise, good nutrition, and positive social networks have all been argued to improve happiness and MDD conditions (Lyubomirsky, 2008). Conversely, too much testosterone can cause significant behavior problems, including aggression, hysteria, and disinhibitory sensation seeking, while not enough testosterone can cause depression, substance abuse, and irritability (Dabbs et al. 1990).

This literature review accumulates the extant research from academic articles and authors that address the existence, causation, languages, and value of the topic used in this field of study. The remainder of this review is as follows. First, the existence and correlation of several emotional problems associated with happiness interventions and low testosterone levels is discussed. Next, the key definitions of depression (i.e., MDD), happiness interventions, and testosterone replacement theory are identified. Because the research done in this field focuses more on MDD than happiness levels, as stated above, this segment of the literature will be referring to MDD. In the third section, the frequency and degree of the issue is discussed in terms of everyone in general, and veterans specifically. The last section discusses actions to be taken and concludes the literature review.

#### Holistic Happiness Interventions and Happiness Levels

The three sources I primarily used for the holistic happiness variables are Lopez et al., 2015; Morgenegg, 2014; and Lyubomirsky, 2008. In Lyubomirsky's book, *The How of Happiness*, she motivates through academic research a list of "happiness activities." These activities include activities such as "Practicing Religion and Spirituality," "Taking Care of Your Body (Meditation)," and "Taking Care of your Body (Physical Activity)."

Looking outside of purely academic research, in an article titled, "The Surprising Science Behind 'Supremely Happy' People," Morgenegg reports that the number one contributor to being happy is a person's social circle and how good it is. Morgenegg discusses the following list:

Happy people surround themselves with other happy people.

Happy people try to be happy.

Happy people spend money more on others than they spend on themselves.

Happy people have deep in-person conversations.

Happy people use laughter as a medicine.

Happy people use the power of music.

Happy people exercise and eat a healthy diet. Happy people take the time to unplug and go outside. Happy people get enough sleep. Happy people are spiritual.

The list above and the Happiness Activities from Lyubomirsky (2008) are very highly correlated. Based on these two secondary sources, and the primary sources they cite, I ask the following holistic happiness questions on the questionnaire:

I meditate regularly. I surround myself with happy people often. I often fake being happy when I feel down. I spend more money on myself than others. I have meaningful/deep conversations with friends often. I seldom laugh. I listen to music often. I exercise (days per week). I eat a healthy, balanced diet. I frequently use technology (phone, TV, computer, etc.). I find it difficult to fall asleep. I feel I get enough sleep. I am a spiritual person.

# *Existence and Correlation of Problems Relevant to Low Testosterone Levels*

To understand the existence and correlation of problems relevant to low testosterone levels, it is instructive to review previous research studies done in this field of study. According to Giltay et al. (2011) major depression has been associated with low testosterone levels. Another study done by Almeida et al. (2008) concluded older men with low levels of testosterone are more likely to experience depression than men who have higher levels of testosterone. According to Almeida et al. (2008), men with testosterone in the lowest quintile were almost three times more likely to experience depression than men with normal levels of testosterone.

Not only is the existence of a problem related to low testosterone relevant to men, according to Schutter and Honk (2009), but low testosterone levels also play a role in the mental and emotional health of women. Consideration of women in the study of this field is being carefully studied because women have been found to be more than two times more likely to experience mood disorders, including MDD, and anxiety disorders (Giltay et al. 2011). Women are also more likely to experience physical pain associated with mental disorders including back, head, and muscle aches (Runnals et al. 2013). It was found that depressive disorder and generalized anxiety disorder, along with other anxiety disorders, are associated with lower testosterone levels in women. Based on this review of testosterone as a factor in happiness levels and MDD, I ask the following 2 questions on the questionnaire:

I take prescribed testosterone.

If you have taken prescribed testosterone, how long have you used it?

#### The Specific Case of Veterans

Mckenzie et al. (2009) found that veterans who served in lowerranking statuses as soldiers were more likely to become depressed and turn to alcohol to self-medicate. Veterans who have served on deployment terms and who have seen the traumas of war are more susceptible to emotional problems including MDD (Kearney et al. 2012). According to Kearney et al. (2012), among the soldiers who were deployed and served in Iraq, 12 months after deployment 20.7% to 30.5% of soldiers were diagnosed with various emotional problems including MDD.

To sum up the existence and correlation of the problem of the impact of low testosterone on military veterans, low testosterone affects men and women alike with its negative impact on happiness levels (i.e., increased depression). Veterans are susceptible to emotional and mental trauma due to the nature of fighting in a war.

To understand the information presented in this study, a solid understanding of the terms and language used in this field of study is important. According to the Mayo Clinic definition:

> Depression is a mood disorder that causes a persistent feeling of sadness and loss of interest. Also called major depression, major depressive disorder or clinical depression, it affects how you feel, think and behave and can lead to a variety of emotional and physical problems. You may have trouble doing normal day-to-day activities, and depression may make you feel as if life isn't worth living (http://www.mayoclinic.org/ diseases-conditions/depression/basics/definition/con-20032977).

Barton (2015) also defined MDD as a mental illness that lessens one's quality of life and also puts a person at a higher risk for suicide. As previously discussed, veterans experience MDD often, which makes this mental disorder relevant to this study. To combat MDD, testosterone replacement therapy has been shown to have some positive effects (Seidman, 1998). According to WebMD (2015), testosterone replacement therapy can be administered via intramuscular injections (2–10 weeks apart), testosterone patches, testosterone gels, mucoadhesive material applied on the gums twice a day, and oral pills. The Merriam-Webster definition for prescribed testosterone is "a synthetically administered hormone that is responsible for developing and maintaining male secondary sex characteristics." Although it is more prominent in men, women also have testosterone, just at lower levels (Schutter and Honk, 2009).

Specific definitions of MDD and testosterone replacement therapy are both important to understand when studying how the two correlate. "[Mental health issues] occur in alarming proportion among both civilians and soldiers exposed to war or violence" (Kearney et al. 2012). As noted above, Barton (2015) concluded that 1.5 million veterans needed professional mental health care in 2014. Mental health has been a concern for veterans since there were wars to fight (Kearny et al. 2012). According to Kearny and his research team and as previously discussed, 20.7% to 30.5% of Iraqi War veterans sought help for mental health-related issues 3 to 12 months after deployment. According to the same article, the statistics among Vietnam-era veterans were very close to those of the Iraqi war. Mental health problems are a concern not only for deployed veterans, but also for soldiers with histories of military sexual trauma. For those soldiers with a history of military sexual trauma, 29% to 60% deal with mental health issues including MDD.

Along with the above stated problems, MDD in veterans can also put veterans at a greater risk for comorbidity (Runnals et al. 2013). According to the same study, veterans who experience MDD are at a higher risk to develop physical pain and other anxiety disorders such as post-traumatic stress disorder (PTSD).

Veterans have been exposed to many situations in which emotional and mental trauma may occur. The frequency and degree of this issue is outlined in the above section, including statistics from multiple studies.

Based on the research on veterans, I ask the following questions in my questionnaire:

I am glad I served in the armed forces.

Branch of service.

I was \_\_\_\_\_ during my service (Active, National Guard, Reserves). Number of years of Service (enter a number).

Total months deployed overseas (enter a number).

Area I was/am deployed to (Choose all that apply): Iraq, Afghanistan, Korea, Vietnam, WWII, Other

I was actively involved in combat.

#### Demographic and Other Control Variables

In this section, I discuss seven more variables that may impact happiness and MDD. The first four are demographic in nature:

I am (gender). I am (age). My annual income level. Marital status.

These variables are designed to control for relevant factors that may impact happiness. For example, younger people may be happier than older people if they enjoy better health and more vibrant energy.

The second class of variables deals with either mood medication or therapy. As this paper is about happiness/MDD, it follows that people who self-medicate with alcohol or who medicate with a physician's prescription should be controlled for (Barton, 2015). In addition, psychological counseling therapy may impact happiness and depression levels (Stewart et al. 2014). As such, I ask the following three questions:

I drink alcohol (days per week). I am taking medication to help with mood. I receive psychological counseling therapy.

### DATA

I created a 59-question survey made partially from the Oxford Happiness Level survey (first 29 questions), partially of the questions motivated above (questions 30–42), and partially of a series of demographic questions (questions 43–59). The purpose of the survey was to identify U.S. veterans' happiness levels. I performed a beta test by having five veterans take the survey and let me know of any errors/clarification issues that needed editing before I distributed the survey to the rest of the veterans. After making sure that the survey was good, I sent it out to as many veterans as I could reach via my contacts' listserves. There are 76 completed surveys from veterans on which the data and results are based.

The sample deserves some discussion as it is a sample of convenience. I used five main contacts: an uncle who served 30 years in the Navy, an uncle who served 28 years in the Army, an uncle who served 6 years in the Army (including combat in the first Gulf War), a local veteran's support group, and a colonel who is on his 26<sup>th</sup> year of service. Each of these contacts sent out emails and Facebook posts to their military contacts requesting them to complete the survey. As the survey was anonymous, it is not possible to determine which main contacts generated the most completed surveys. The five contacts are geographically dispersed throughout the country, and their contacts are scattered all over the country and world. A potential limitation of this study is this survey of convenience. While there is no reason to believe the sample is not representative of the population of veterans, there is also not direct evidence that it is generalizable.

Table 1 lists and defines the variables and the abbreviations of the variables that are used when discussing the data and results of the survey. For instance, *Surround* is short for, "I surround myself with happy people often." The remainder of the list in Table 1 follows this same pattern directly from the survey instrument with the exception of the last row. *HAPPYSCORE* is the score from the Oxford Happiness Questionnaire and is computed by averaging the first 29 questions of my survey (after controlling for reverse-coded questions).

Table 2 lists the variables and their corresponding sample size, mean, median, minimum, and maximum. To demonstrate how to read the table, 76 people answered the question about the variable *Surround*, with the average person marking 4.53 on the scale. The median number was 5 and the minimum was 1 with the maximum of 6. Table 2 shows that the *HAPPYSCORE* average is 3.73, which is below the national average of around 4.3 (Hills and Argyle, 2002). In fact, the maximum score is only a fourth of a point higher than the national average. Although outside of the scope of this study, it is interesting that my sample has an average lower than that of Hills and Argyle (2002).

### **EMPIRICAL METHODS**

The initial empirical method used is the Spearman Correlation to determine how two variables vary with each other (Spearman, 1904). I use the Spearman measure because most of the variables are categorical in nature (i.e., Likert scale of one through six). The Spearman correlation coefficient ranges from one (perfectly positively correlated) to negative one (perfectly negatively correlated). The statistic itself is computed as:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)}.$$
(1)

Table 1. Varia	able Definitions
Variable	Survey Question
Surround	I surround myself with happy people often.
Fake	I often fake being happy when I feel down.
Spend	I spend more money on myself than others.
Convo	I have meaningful/deep conversations with friends often.
NoLaugh	I seldom laugh.
Music	l listen to music often.
EatHealthy	I eat a healthy balanced diet.
Meditate	I meditate regularly.
GladServed	I am glad I served in the armed forces.
Tech	I frequently use technology (phone, TV, computer, etc.).
DifSleep	I find it difficult to fall asleep.
EnuffSleep	I feel I get enough sleep.
Spiritual	l am a spiritual person.
Alcohol	I drink alcohol (days per week).
Exercise	I exercise (days per week).
Gender	I am (gender).
Age	I am (age).
Branch	Branch of service.
ActiveDuty	I was during my service (Active, National Guard, Reserves).
YearSvc	Number of years of Service (enter a number).
MonthsDepl	Total months deployed overseas (enter a number).
Iraq	Area I was/am deployed to (Choose all that Apply)Iraq
Afghan	Area I was/am deployed to (Choose all that Apply)Afghanistan
Korea	Area I was/am deployed to (Choose all that Apply)Korea
Vietnam	Area I was/am deployed to (Choose all that Apply)Vietnam
wwii	Area I was/am deployed to (Choose all that Apply)WWII
OtherDep	Area I was/am deployed to (Choose all that Apply)Other
Combat	I was actively involved in combat.
Testost	I take prescribed testosterone.
TestLong	If you have taken prescribed testosterone, how long have you used i
MoodMed	I am taking medication to help with mood.
PsychTherapy	I receive psychological counseling therapy.
Income	My annual income level.
Marital	Marital Status.
Нарру	I consider myself a happy person.
HAPPYSCORE	Score from Oxford Happines Questionnaire.

where n is the sample size and  $di = x_i - y_i$  is the difference in the ranks of the two variables in question. To test for statistical significance,  $\sigma$ , the standard error of the estimated coefficient is computed as:

$$\sigma = \frac{0.6325}{\sqrt{n-1}}.$$

Table 2. Descriptive Statistics						
Variable	Ν	Mean	Median	Minimum	Maximum	
Surround	76	4.53	5	1	6	
Fake	76	2.83	3	1	6	
Spend	76	2.17	2	1	6	
Convo	76	3.54	4	1	6	
NoLaugh	76	1.91	2	1	6	
Music	76	4.50	5	1	6	
EatHealthy	76	4.09	4	1	6	
Meditate	76	2.46	2	1	6	
GladServed	76	5.89	6	4	6	
Tech	76	5.74	6	1	6	
DifSleep	76	2.62	2	1	6	
EnuffSleep	76	3.54	3	1	6	
Spiritual	76	4.80	5	1	6	
Alcohol	75	2.17	2	1	5	
Exercise	75	2.85	3	1	5	
Gender	75	0.85	1	0	1	
Age	70	49.51	52	0	81	
Branch	75	1.31	1	1	4	
ActiveDuty	74	1.54	1	1	3	
YearSvc	74	17.02	14.5	5	35	
MonthsDepl	72	22.00	15	0	87	
Iraq	76	0.47	0	0	1	
Afghan	76	0.21	0	0	1	
Korea	76	0.16	0	0	1	
Vietnam	76	0.01	0	0	1	
wwii	76	0.00	0	0	0	
OtherDep	76	0.45	0	0	1	
Combat	74	0.43	0	0	1	
Testost	75	0.08	0	0	1	
TestLong	69	1.23	1	1	4	
MoodMed	75	0.08	0	0	1	
PsychTherapy	74	0.12	0	0	1	
Income	75	6.15	7	1	7	
Marital	75	0.84	1	0	1	
Нарру	75	0.89	1	0	1	
HAPPYSCORE	76	3.73	3.74	2.69	4.55	

The next test is the difference in means t-test (Rice, 2006). Based on a test for equal variances, either the equal variance or unequal variance specification of the test is used. T-tests are based on the Happiness Score, and the class is constructed by dividing each variable at the median ( $X_1$  and  $X_2$  in the below statistic). For variables in which equality of variance cannot be rejected, I use the following statistic:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{X_1 X_2} \cdot \sqrt{\frac{1}{n}}}$$
(3)

where

$$s_{X_1X_2} = \sqrt{(s_{X_1}^2 + s_{X_2}^2)} \tag{4}$$

and n is sample size, s is standard error.

For variables with unequal variance, I use:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{X_1 X_2} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$
(5)

where

$$s_{X_1X_2} = \sqrt{\frac{(n_1 - 1)s_{X_1}^2 + (n_2 - 1)s_{X_2}^2}{n_1 + n_2 - 2}}.$$
 (6)

The third statistical test is an ordinary least square (OLS) regression to test the various variables in a multivariate setting. The base model is:

$$y_i = \alpha + \beta x_i + \varepsilon_i. \tag{7}$$

where y is the dependent variable and equals the Happiness Scale score for each veteran (i) in the sample. The intercept is denoted as  $\sigma$ . X is the vector containing all of the independent variables listed in Table 1. Finally,  $\varepsilon$  is the error term of the regression.

#### Results

Looking at Table 3, we can see the Spearman Correlation results. The results for *Surround* ( $\pm 0.27$ ), *Convo* ( $\pm 0.28$ ), *Music* ( $\pm 0.07$ ), *Eat*-*Healthy* ( $\pm 0.04$ ), *Meditate* ( $\pm 0.01$ )), *Age* ( $\pm 0.06$ ), *Afghan* ( $\pm 0.01$ ), *Combat* ( $\pm 0.29$ ), and *PsychTherapy* ( $\pm 0.10$ ) are statistically significant. Thus, in a univariate setting, the two strongest correlates with *Happiness* are surrounding onself with positive people and having meaningful conversations. The largest negative correlate is experiencing combat. A finding that may seem surprising is that veterans who are receiving psychological therapy are negatively related to happiness levels. The most likely answer is due to a self-selection bias. That is, veterans who are depressed seek therapy; those who are not depressed do not. So by

Table 3. Spearman Correlations							
Surround	0.27299	Tech	-0.05104	ActiveDuty	-0.00619	Testost	-0.11362
	0.017**		0.6615		0.9583		0.3318
	76		76		74		75
Fake	0.09622	DifSleep	-0.06738	YearSvc	0.03036	TestLong	-0.1266
	0.4083		0.56		0.7974		0.2999
	76		76		74		69
Spend	-0.06723	EnuffSleep	0.14235	MonthsDepl	-0.02662	MoodMed	-0.13975
	0.5639		0.2199		0.8243		0.2318
	76		76		72		75
Convo	0.27623	Spiritual	0.15079	Iraq	-0.14729	PsychTherapy	-0.19375
	0.0157**		0.1935		0.2042		0.0981*
	76		76		76		74
NoLaugh	-0.12218	Alcohol	0.07241	Afghan	0.30335	Income	0.06085
	0.2931		0.5370		0.0077***		0.6040
	76		75		76		75
Music	0.20971	Exercise	0.03747	Korea	0.03375	Single	-0.07993
	0.0690*		0.7496		0.7722		0.4925
	76		75		76		76
EatHealthy	0.24031	Gender	-0.10891	Vietnam	-0.13698	Married	0.09486
	0.0365**		0.3523		0.2380		0.415
	76		75		76		76
Meditate	0.31228	Age	-0.2242	OtherDep	0.0658	Divorced	-0.0896
	0.0060***		0.0621*		0.5722		0.4415
	76		70		76		76
GladServed	0.03107	Branch	0.17174	Combat	-0.29144	Нарру	0.03495
	0.7899		0.1407		0.0118**		0.7660
	76		75		74		75

definition, veterans in therapy are those who are struggling with happiness levels. In the multivariate tests, I attempt to control for this effect.

\*\*\* 0.01 significance ; \*\* 0.05 significance ; \* 0.10 significance

Table 4 reports the variables with statistical significance in their difference in means tests. To form the two groups, the *Happiness* score is divided at the median. The mean of the high score group is then sub-tracted from the mean of the low score group. Six variables have significantly different means in the two classes. Happier people tend to surround themselves with good support, have meaningful conversations, meditate, and were deployed to Afghanistan. All of these results are consistent with the Table 3 Spearman Correlation results. The two variables *Combat* and *PsychTherapy* are also consistent with Table 3 in that they are negatively related to veteran happiness.

I estimate two multivariate models in this study, one including all of the dependent variables in the survey, and the other including only those dependent variables that are significant in the first model. I estimate the second model as a robustness check in case there is a problem related to degrees of freedom with the OLS model when using all of the dependent variables. The grand model is reported in Table 5.

Table 4. T-Tests				
Variable: Surround (Surround)				
happycut	N	Mean		
0	38	4.2105		
1	38	4.8421		
Diff (1-2)		-0.6316		
p-value		0.0239		
-				
Variable: Convo (Convo)				
happycut	N	Mean		
0	38	3.2105		
1	38	3.8684		
Diff (1-2)		-0.6579		
p-value		0.0382		
Variable: Meditate (Meditate)				
happycut	N	Mean		
0	38	2.0263		
1	38	2.8947		
Diff (1-2)		-0.8684		
p-value		0.0155		
Variable: Afghan (Afghan)				
happycut	N	Mean		
0	38	0.1053		
1	38	0.3158		
Diff (1-2)		-0.2105		
p-value		0.0248		
Variable: Combat (Combat)				
happycut	N	Mean		
0	38	0.5263		
1	36	0.3333		
Diff (1-2)		0.1930		
p-value		0.0959		
Variable: PsychTherapy (Psyc				
happycut	N	Mean		
0	38	0.1842		
1	36	0.0556		
Diff (1-2)		0.1287		
p-value		0.0896		

The model itself has significant explanatory power with an F-statistic p-value of 0.0035. The adjusted R-square is 0.4753 indicating that the model as a whole explains 47.53% of the variability in the *Happiness* score of veterans in the sample.

Table 5. Ordinary Least Square Regression						
with All of the Survey Ouestions						
Analysis of Variance						
Source	DF	Sum of Mean		F Value	Pr > F	
		Squares	Square			
Model	34	4.92	0.14	2.7	0.0035	
Error	30	1.61	0.05			
Corrected Total	64	6.53				
Root MSE	0.23	R-Souare	0.75			
Dependent	3.71	Adi R-Sa	0.48			
Mean						
Coeff Var	6.24					
Variable	Parameter	Heter	oscedas	ticity	Variance	
	Estimate	Standard	t Value	Pr >  t	Inflation	
		Error				
Intercept	3.510	0.657	5.34	<.0001	0	
Surround	0.019	0.031	0.59	0.5583	3.42	
Fake	0.081	0.022	3.64	0.0010	2.49	
Spend	-0.041	0.016	-2.50	0.0180	2.04	
Convo	0.058	0.037	1.56	0.1295	4.25	
NoLaugh	0.015	0.028	0.52	0.6068	4.06	
Music	0.113	0.025	4.49	<.0001	2.51	
EatHealthy	0.020	0.028	0.73	0.4681	3.24	
Meditate	0.022	0.020	1.09	0.2853	2.54	
GladServed	-0.087	0.089	-0.98	0.3366	2.57	
Tech	-0.089	0.023	-3.80	0.0007	1.65	
DifSleep	0.029	0.021	1.43	0.1631	3.29	
EnuffSleep	-0.014	0.019	-0.75	0.4585	2.52	
Spiritual	0.016	0.035	0.47	0.6408	2.66	
Alcohol	0.034	0.024	1.40	0.1703	2.48	
Exercise	-0.053	0.046	-1.14	0.2617	5.24	
Gender	-0.043	0.088	-0.49	0.6297	2.08	
Age	0.009	0.004	2.13	0.0417	3.56	
Branch	0.032	0.046	0.69	0.4954	2.39	
ActiveDuty	0.079	0.038	2.06	0.0487	2.13	
YearSvc	-0.008	0.004	-1.92	0.0650	3.15	
MonthsDepl	-0.004	0.002	-2.19	0.0365	3.09	
Iraq	0.177	0.085	2.07	0.0467	4.18	
Afghan	0.242	0.073	3.32	0.0024	2.03	
Korea	0.161	0.062	2.61	0.0139	2.12	
Vietnam	-0.185	0.136	-1.36	0.1846	1.71	
Combat	-0.115	0.092	-1.25	0.2209	4.22	
lestost	0.123	0.096	1.28	0.2121	3.24	
IVIOOdMed	-0.014	0.105	-0.14	0.8920	3.43	
Psych I herapy	-1.221	0.148	-8.26	<.0001	11.62	
Income	-0.016	0.020	-0.80	0.4310	2.08	
	-0.063	0.106	-0.60	0.5545	5.44	
DIVORGED	-0.156	0.103	-1.52	0.1384	4.01	
	-0.007	0.027	-2.4/	0.0194	4./3	
nappy_inerapy	1.158	0.172	0.73	<.0001	9.48	

For ease of interpretation, the statistically significant variables are bolded. As these variables include the controlling effect of the other variables being tested simultaneously, they should be the most robust. Going in order of the table, the variable *Fake* lends support for the saying, "Fake it 'til you make it" in that it is significant and positive. That is, people who fake being happy, when they actually are not, score higher on the *Happiness* score. Interesting, going against predictions, people who spend money on others (*Spend*) are not significantly happier. In fact, they are sadder according to the Oxford scale. Veterans who listen to music often (*Music*) also experience higher happiness levels. The survey simply asked about frequency of listening to music and not what type of music they listen to.

Consistent with the idea that technology can add stress to life, the *Tech* variable is negative and significant. Veterans who spend a lot of time with technology are less happy—perhaps they are trying to "self-medicate" with their technology. Now that other variables are being controlled for, *Age* is positive and significant instead of negative as in the Spearman Correlations. This means in the univariate tests other effects were captured in *Age* to cause the negative correlation. Now that those other variables are controlled for, we see that the older veterans in my sample are happier, perhaps because many of them may be retired and have less financial stress.

The next group of variables captures the military variables. ActiveDuty, Iraq, Afghan, and Korea are all positive and significant. These results indicate that veterans who served active duty are happier than those who were in the National Guard or reserves. In addition, those who served in Iraq, Afghanistan, or Korea were happier than those who served in "other" deployments such as Europe, Asia, or Latin America. Interestingly, the "other" deployments constitute noncombat missions. In the regression, the variable Combat is no longer significant, although it was in the univariate tests. This result indicates that it is not the actual combat that is correlated with a lack of happiness, it is where the person served in combat or was deployed that drives the happiness level. Two service-related variables that negatively impact veteran happiness levels are the years of service of the veteran (YearSvc) and the number of months deployed (MonthsDepl). Both of these variables indicate that after controlling for other variables, happiness levels are lower with the length of overall military service and length of deployments.

A variable that is interesting because it is not significant is the *Testost* variable. The use of external testosterone treatment is fairly controversial as discussed in the literature review. In my sample, in which about 10% of respondents were on prescription testosterone, it did not significantly correlate with *Happiness* levels in either the univariate or the multivariate settings. This finding may be a limitation of the study because such a small proportion of the sample was taking

testosterone when they took the survey. With a larger proportion taking testosterone, the statistical power would increase and might reveal a correlation.

Dependent Variables from First Model						
Analysis of Variance						
Source	DF	Sum of Mean		F Value	Pr > F	
		Squares	Square			
Model	14	3.59	0.26	4.3	<.0001	
Error	51	3.01	0.06			
Corrected Total	65	6.61				
Root MSE	0.24	R-Square	0.54			
Dependent Mean	3.71	AdjR-Sq	0.42			
Coeff Var	6.55					
Variable	Parameter	Heteroscedasticity		ity	Variance	
	Estimate	Standard	t Value	Pr >  t	Inflation	
		Error				
Intercept	3.603	0.288	12.53	<.0001	0	
Fake	0.073	0.021	3.49	0.0010	1.45	
Spend	-0.032	0.019	-1.68	0.0986	1.15	
Music	0.110	0.026	4.29	<.0001	1.54	
Tech	-0.106	0.030	-3.54	0.0009	1.31	
Age	0.004	0.003	1.51	0.1382	1.49	
ActiveDuty	0.095	0.038	2.52	0.0148	1.31	
YearSvc	-0.008	0.004	-1.90	0.0624	2.18	
MonthsDepl	-0.002	0.001	-1.75	0.0869	1.62	
Iraq	0.078	0.071	1.10	0.2773	2.11	
Afghan	0.270	0.067	4.02	0.0002	1.39	
Korea	0.152	0.057	2.67	0.0102	1.46	
PsychTherapy	-1.034	0.246	-4.20	0.0001	4.95	
Happy_Ex	-0.075	0.021	-3.60	0.0007	1.48	
Happy_Therapy	0.931	0.244	3.82	0.0004	4.46	

Table 6. Ordinary Least Squar	e Regression v	with Only	Significant
Dependent Variables from Fir	st Model		

The final three significant variables are related to each other in an attempt to get a fuller understanding of their impact on happiness. As discussed above, veterans who seek psychological therapy have lower happiness levels, and this is confirmed in the regression as well. In an attempt to determine whether exercise and therapy help with happiness levels (causation) instead of simply correlations, I create the *Happy\_Ex* and *Happy\_Therapy* variables. These two are interaction terms that multiply the self-reported binary question on whether the veteran felt he or she was overall generally happy (*Happy*, the last question in the

survey) by the *Exercise* variable and the *PsychTherapy* variable, respectively. The result on *Happy\_Ex* indicates that when people judge themselves to be happy, exercise is correlated with lower levels of happiness. The result on *Happy\_Therapy* concludes when people judge themselves to be happy, it seems therapy is correlated with more happiness.

Table 6 reports the parsimonious model that includes only the significant dependent variables from the first OLS model. All of the variables retain their same sign of correlation, but two lose their significance. *Age* and *Iraq* both still have positive coefficients, but the p-values are now too high for statistical significance. Overall, the robustness the model reports in Table 6 demonstrates support for the findings in Table 5.

#### CONCLUSION

In this paper, I construct a survey and test hypotheses dealing with holistic happiness activities and testosterone treatment on happiness levels (MDD). I find over a dozen variables that impact the level of happiness in veterans. *Happiness* is measured using the Oxford Happiness Scale. Factors that are positively correlated with *Happiness* are: faking to be happy when you are not, listening to music often, age, serving on active duty, serving in Iraq, Afghanistan, or Korea, and exercise for a given level of happiness. Factors that are negatively correlated with Happiness are: spending money on friends, using a lot of technology, serving longer periods in the military, being deployed for longer periods in the military, and seeking psychotherapy either unconditionally or conditioned on happiness levels. The somewhat controversial use of testosterone treatment to help depression showed no significant impact in the tests.

My results are limited in that they show correlation and not causation. That is, I cannot discern whether people who are already happy like to meditate or if meditation helps people become happier. The result on seeking psychotherapy is an example where the causation and correlation are particularly important. Future research can explore the topic of happiness activities on happiness levels by trying to establish causality and not just correlation. Given this limitation, my study shows that happy people are significantly correlated with specific activities and experiences in their lives and suggests that perhaps seeking these happiness activities would be a good idea for people seeking higher happiness levels.

## References

Almeida, O., Yeap, B., Hankey, G., Jamrozik, K., & Flicker, L. (2008). Low free testosterone concentration as a potentially treatable cause of depressive symptoms in older men. *Archives of General Psychiatry*, 65(3), 283-283.

Barton, Casey P. (2015). *Major Depressive Disorder in Veterans: Antidepressant Use and Suicide Prevention*. Nova Science Publishers, Hauppauge, NY.

Beck, M. (Oct. 27, 2014). In men's fight against aging, how much risk to take? *Wall Street Journal*.

Dabbs, J. M., Hopper, C. H., & Jurkovic, G. J. (1990). Testosterone and personality among college students and military veterans. *Personality and Individual Differences*, *11*(12), 1263-1269. doi:10.1016/0191-8869(90)90153-I

Giltay, E., Enter, D., Zitman, F., Penninx, B., Pelt, J., Spinhoven, P., & Roelofs, K. (2011). Salivary testosterone: Associations with depression, anxiety disorders, and antidepressant use in a large cohort study. *Journal of Psychosomatic Research*, 72(3), 205-213.

Hills, P. and Argyle, M. (2002). The Oxford Happiness Questionnaire: A compact scale for the measurement of psychological well-being. *Personality and Individual Differences*, 33, 1073-1082.

Kearney, D., Mcdermott, K., Malte, C., Martinez, M., & Simpson, T. (2012). Association of participation in a mindfulness program with measures of PTSD, depression and quality of life in a veteran sample. J. Clin. Psychol. *Journal of Clinical Psychology*, 68(1), 101-116.

Lopez, S.J., J.T. Pedrotti, and Snyder, C.R. (2015). *Positive Psychology: The Scientific and Practical Explorations of Human Strengths*, 3<sup>rd</sup>. ed., Sage, Los Angeles.

Lyubomirsky, S. (2008). *The How of Happiness: A New Approach to Getting the Life You Want*. Penguin, London.

Morgenegg, R. (17 Sep 2014). The surprising science behind "supremely happy" people. July, 14, 2015. https://www.lds.org/ church/news/the-surprising-science-behind-supremely-happypeople?lang=eng&cid=facebook-shared.

Mckenzie, D., Creamer, M., Kelsall, H., Forbes, A., Ikin, J., Sim, M., & Mcfarlane, A. (2009). Temporal relationships between Gulf War deployment and subsequent psychological disorders in Royal Australian Navy Gulf War veterans. *Social Psychiatry and Psychiatric Epidemiology*, 45(9), 843-852.

Rice. J.A. (2006). *Mathematical Statistics and Data Analysis*, 3<sup>rd</sup> ed., Duxbury Advanced, Belmont, CA.

Runnals, J., Voorhees, E., Robbins, A., Brancu, M., Straits-Troster, K., Beckham, J., & Calhoun, P. (2013). Self-reported pain complaints among Afghanistan/Iraq era men and women veterans with comorbid posttraumatic stress disorder and major depressive disorder. *Pain Medicine*, 14(10), 1529-1533.

Schutter, D., & Honk, J. (2009). An endocrine perspective on the role of steroid hormones in the antidepressant treatment efficacy of transcranial magnetic stimulation. *Psychoneuroendocrinology*, 35(1), 171-178.

Seidman, S. (1998). Testosterone replacement therapy for hypogonadal men with SSRI-refractory depression. *Journal of Affective Disorders*, 48(2-3), 157-161.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, 100(3-4), 72–101.

Stewart, M., Raffa, S., Steele, J., Miller, S., Clougherty, K., Hinrichsen, G., & Karlin, B. (2014). National dissemination of interpersonal psychotherapy for depression in veterans: Therapist and patient-level outcomes. *Journal of Consulting and Clinical Psychology*, 82(6), 1201-1206.

WebMD. Is Testosterone Replacement Therapy Right for You?. Retrieved July 18, 2015. http://www.webmd.com/men/guide/ testosterone-replacement-therapy-is-it-right-for-you#1

# Abstracts

#### ARTS

#### **Dance and Class in the Great Depression**

#### Madison Cavaness

Weber State University

The American Dream is one that suggests the capability of moving between classes easily (Dickstein, 218). During the Great Depression, belief in the American Dream was still alive despite the turmoil of the time. It was an idea used in many forms of art to lift the spirits of the people. However, what if the upper class was attempting to appease the lower classes by distracting them with performance and dance? This research argues that 1930s movie musicals functioned as a means to divert the lower class's attention from legitimate concerns surrounding their situation in a way that was harmless and seemed to pose no threat to either class. Instead, the form served to maintain class standing through reinforcement of societal norms of the time. Reinforcing class divisions served as a mechanism to maintain control. Those in power did whatever they could to maintain power. Marxist Critical Theory asks whom does it benefit? Who is truly benefitting from the movie musical? This research will show it was the upper class benefitting from the films. Production companies like RKO Radio Pictures used actors like Fred Astaire and Ginger Rogers to appeal to the lower class, as if to say there was not truly a division between classes. Their 1936 movie "Swing Time" provides a glimpse into the classes of 1930s America, during the Great Depression, particularly in how dance was utilized in a period of great despair. The performances of Astaire and Rogers within the film provide an insight into how the power of the upper class was maintained over the lower class. This research will include a synthesis of written source material and critical analysis of "Swing Time" from a Marxist perspective, illustrating the differences in class and demonstrating how power was maintained during the Great Depression of 1930s America.

### ARTS

# The Road to Salvation Is Lined with Kitsch

#### Victoria Willard

Weber State University

Archaeologists have unearthed ceramic St. Menas flasks from locations all over Western Europe. These flasks were made near the burial place of St. Menas outside of Alexandria, Egypt, and sold at pilgrimage sites around the Mediterranean. These flasks are not normally studied by art historians, as they were mass-produced souvenirs and not considered unique works of fine art. However, they offer an opportunity to study mass production, pilgrimage travel, and the nature of belief in the early Christian world. Despite their rough crudely crafted appearance, these flasks were of great spiritual value to those who purchased them. The flasks themselves, through their clay, embodied the sacred location of the saint's burial site and could be used to carry water, dirt, or oil acquired at sacred places. These natural materials served as relics that offered the bearer special access to divine power by invoking the good will of the saint. The owners of these items believed that the flasks, with their sacred clay, image of the saint, and other sacred contents, had the power to ensure safe travel and heal people of their infirmities. Ultimately, many of these flasks were discovered in Christian burial sites, suggesting that the owners of these items believed that these relics could assist the owner in successfully reaching heaven. This presentation will trace the objects from their origins in Egypt through to their final resting places and offer insights into the meaning and making of these items

#### ARTS

### **Substrate Poetry**

Zach T. Power Weber State University

Poetry has a long-standing tradition with the white sheet of paper. Even as poetry has moved online, the website as a medium is nonetheless skeumorphic and referential of its traditional medium: the blank white

sheet of paper. As a result, the word processor has been the main source of creating poetry, reducing poetry (mostly) to a systems art based on the keyboard (at least in a Western sense). Substrate poetry is poetry that resists the traditional substrates of poetry (white sheet of paper, keyboard, word processor, etc.) and seeks meaning through new mediums where scripted visual language is possible. Substrate poetry does not try to just place language/poetry in new places/mediums, but seeks to integrate the meaning of the language with the meaning of the substrate. Some examples of nontraditional substrates include social networks such as Twitter, Vine, Tumblr, etc., or a return to handwritten language, or even file formats such as .GIF. Additionally, substrate poetry is aware of the physical ontology of a poem: if digital, the poem exists electromagnetically on a server or computer; if physical, it exists in visual contrast on ink and paper, etc. Ultimately, substrate poetry hopes to highlight the possibilities with the ways in which poetry exists to free it from the trappings of its tradition and free poetry to exist more extensively in response to new technology. Art always works in tandem with technology, and substrate poetry seeks to highlight and clarify that relation.

#### ARTS

# The Invisibles: The Unseen Artists Behind the Bangs Sisters' Precipitated Paintings

#### Loclyn Torres

Weber State University

Throughout the history of art, it has not been unheard of for a painting to be credited to the wrong artist, only for the true artist to be identified centuries later. However, there remains an entire movement in art built on the backs of unseen and unknown artists. During their time, they were the wizard behind the curtain, and today they have been all but forgotten. They are the invisible artists who made the Precipitated Spirit Portraits during the heyday of Spiritualism (1848–1920s). Grieving loved ones could receive full oil portraits of their deceased done by 'spirits' working through the powers of the mediums. While most mediums used cabinets or the dark to hide the painting until it was complete, there was a pair of phenomenal sister mediums, May and Lizzie Bangs, who were able to make paintings appear in the light just two feet away from the sitter's face. People traveled from all over to watch as art was created, spontaneously, without hands. But no one knows just how these portraits were actually created. To this day, there are unanswered questions: how did the Bangs Sisters turn a blank canvas into a fully rendered portrait just a couple feet from the sitter's face, how did they know what the dead looked like, how did the sisters get the finished portrait into the room undetected, what are these 'unearthly' portraits made out of, and the greatest question of them all, who actually created these portraits? By synthesizing the research and reports of both the skeptics and believers, I seek to answer these questions and finally reveal the truth behind these otherworldly portraits and turn one of the greatest art mysteries into a new page in art history.

#### ARTS

# The Invisibles: The Hope of a Decade: Fred Astaire's and Ginger Rogers's Dances of the 1930s

#### Kathie Debenham, Pat Debenham

Weber State University

This presentation examines from a Laban Movement Analysis perspective how the choreography and performance of Fred Astaire and Ginger Rogers embodied hope during dark days of the Great Depression. Astaire and Rogers's partnership on screen was a physical metaphor for opportunity and for freedom from the oppressive reality of the deep and pervasive poverty that gripped the United States. Their musicals, though by some considered light-weight in relationship to the reality of life outside the movie house, exuded 'hope' and encouraged a sense of optimism that "tomorrow would be a better day." We use the lenses of Body, Effort, Shape, and Space to examine two specific works, 'Let's Face the Music and Dance' from "Follow the Fleet" and 'Pick Yourself Up' from "Swing Time." These dances illustrate how Astaire and Rogers exhibited in physical form a hope for both the nation and the individual.

#### ARTS

### **Russia's Imitation Act Through Don Quixote**

#### **Madeline Fowers Baum**

Weber State University

Current dance history research holds that dance, as a cultural artifact, reflects culture. Noted dance anthropologist, Joann Kealiinohomoku stated "all dance forms reflect the cultural traditions in which they developed" (Kealiinohomoku 33). Through an analysis of dance, important cultural values are revealed, such as a culture's aspirations, political structure, and social expectations. Therefore, critically analyzing the ballet "Don Quixote" should provide insight into Russian culture in the late 1800s. However, further inspection of the ballet reveals several seeming contradictions challenging Kealiinohomoku's assertion. Don Ouixote was a Spanish novel that premiered as a ballet in Moscow in 1869. How does a Spanish story reflect Russian culture? To address this question, the New Historicism frame of analysis will be very enlightening. New Historicism is a critical theory that interprets history as a cultural artifact: a way to learn how time and place are linked together and reflect a specific culture (Tyson 286). Understanding the cultural atmosphere of 19th-century Russia is essential to understanding why Russian ballet displayed foreign elements. 19th-century Russia went to great lengths to emulate European culture. This emulation began in the late 1600s when Tsar Peter the Great attempted to close the perceived cultural gap between Moscow and Western Europe through a 'Westernizing project.' Peter aspired for Russia to appear as powerful, cultured, and sophisticated as France, the leading world power of the time (Homans 246). As Russia emulated European powers, particularly France, foreign cultural relics leaked into Russian ballet. "Don Quixote" is a perfect example of Russian efforts to adopt Western values in ballet. A review of written source material and a critical analysis of "Don Quixote" will demonstrate how Peter the Great's Westernizing efforts were noticeable centuries later as Western values were still emulated through ballet to make Russia appear powerful, and therefore cultured.

#### **BIOLOGICAL SCIENCES**

# Determination of Treatments to Reduce Late Gassy Defect in Cheese due to *Lactobacillus wasatchensis* WDC04 Contamination

#### **Issac Bowen, Craig Oberg**

Weber State University

Lactobacillus wasatchensis WDC04 is a newly discovered lactic acid bacterium that causes serious commercial losses from bloated cheese packages and textural defects. Experiments were performed to determine its salt tolerance at pH 5.2 and 6.5, resistance to high temperature/short time pasteurization, and pH growth range. MRS with 1.5% ribose (MRS-R) was prepared at either pH 5.2 or 6.5 with salt concentrations ranging from 0.0% to 10.0%. Two milliliters of the MRS-R test medium was added to each well in a 24-microwell plate, and a preinoculated absorbance reading was taken at 600 nm. After, 100 µL of WDC04 was inoculated into each well, and the plate was incubated at 25°C for 3 d (pH 6.5 MRS-R) or 2 d (MRS-R pH 5.2). Plates were placed in a Teacon Infinite 2000, and absorbance readings (A600) were taken every 4 h for 24 h. Results showed WDC04 grew best at 3.0% salt (pH 6.5) and 2.0% salt (pH 5.2) but showed some growth up to 6.0%. A narrower salt range (5.25%-6.75%) at pH 6.5 was used to determine whether a salt concentration used in cheese could suppress WDC04 growth. Above 6.0% salt, WDC04 was inhibited and ceased to grow. Using the same methodology, the pH range (2-8) for WDC04 was determined. WDC04 grows best at pH 5-6 (cheese pH) but not below pH 4 or above pH 7. WDC04 was heat shocked in a hot water bath at 72°C for 15 sec and plated. With an initial count of 108 CFU/ml, results showed a decrease of 105 CFU/ml in survival of WDC04, indicating WDC04 could be contaminating the cheese by surviving pasteurization. These results suggest that using a higher salt/moisture ratio in cheese and increasing pre-pasteurization sanitation to remove biofilms would decrease the likelihood of Lb. wasatchensis WDC04 in ripening cheese, thus reducing the possibility of late gassy defect.

#### **BIOLOGICAL SCIENCES**

# Rapid Floristic Analysis with Geolocated Photographs and Geographic Information System (GIS) Software

Sadie Larson, Matthew Hutchinson, James Harris

Utah Valley University

Geographic Information System (GIS) computer software has become a powerful tool in biogeographic analyses, but the technique is somewhat limited by the time and effort required to collect statistically significant amounts of distribution data through traditional methods. Geolocated photographs (photos with embedded geographic coordinates) provide a rapid, precise method of collecting vast amounts of biological distribution data that can then be incorporated into a GIS analysis of distribution. The speed and simplicity of geospatial data collection with photographs provide a practical method of producing a high-resolution characterization of biological distribution in a region of interest. The technique is particularly well suited to vegetation studies. To test the concept, we assembled over 9000 geolocated photographs of plants from Mt. Timpanogos in the Wasatch Range of Utah. Latitude, longitude, and elevation data were used in ArcMap GIS software to correlate plant distribution with a wide variety of ecological factors. Our GIS analysis allows a broad range of ecological questions about the flora of Mt. Timpanogos to be addressed. Because of the simplicity of collecting geolocated photographs of plants, our technique lends itself well to student and citizen science projects.

#### **BIOLOGICAL SCIENCES**

# Insect Origins, Key Innovations, and Trophic Succession

Robert L. Bossard Bossard Consulting

Like Aphrodite, insects appear to rise perfected from the sea. New evidence shows subterranean freshwaters should be investigated for cryptobiotic microcrustaceans closely related to insects. Key innovations that help create insect diversity include hexapody, water conservation, dicondylic jaws, foldable wings, sucking mouthparts, elytra, a single pair of wings, sociality, and symbioses with microbes, fungi, plants, and animals. I examine the trophic succession hypothesis that herbivorous insects preceded carnivorous and parasitic insects during biosphere evolution.

#### **BIOLOGICAL SCIENCES**

# The Effects of Caffeine-Free American and Mexican Coca-Cola® on Glycemic Index and Glycemic Load

#### **Spencer Rickers, Jordan Cruze, Sariah Saili, Nathan Kartchner** *Utah Valley University*

Most soft drinks sold in the United States are sweetened with highfructose corn syrup (HFCS) rather than sucrose or cane sugar (CS). Public concerns say that consuming HFCS is less healthy than consuming CS. The purpose of this study is to examine the glycemic index (GI) and glycemic load (GL) of caffeine-free Coca-Cola® variants containing CS and HFCS from a carbohydrate standpoint. Twenty nonpregnant, healthy females, between the ages of 19 and 30, with a BMI between 18 and 25 were selected to participate. Each subject consumed three 12-oz drinks containing 39 g of carbohydrate in the form of CS (Mexican Coca-Cola®), HFCS (American Coca-Cola®), and dextrose water (DW), with tap water (TW) as the control. Capillary blood glucose levels (BGLs) were measured at 15-minute intervals over four two-hour testing periods. The GL and GI were calculated for each drink by taking the area under the blood glucose/time curve. One-sample T-test on differences of GI with a Pvalue of 0.773 failed to reject the null hypothesis and concluded that the average GI for American Coca-Cola® is not significantly greater than that of Mexican Coca-Cola®. Non-parametric procedures also concluded that the median GI for American Coca-Cola® is not significantly greater than that of Mexican Coca-Cola®. Corresponding tests on GL data do not change this conclusion. We conclude that differences in GI and GL for noncaffeinated soft drinks sweetened with CS versus HFCS are not significant, as revealed by our analyses.

#### **BIOLOGICAL SCIENCES**

## Student's T-Test and Scientific Reproducibility

#### Julian Chan, Brian Knaeble, Russell Costa, Douglas Getty, Katie McLean, Riddhi Rampeearee

Weber State University

Many scientific results cannot be reproduced. Although there are many reasons for irreproducibility in science, here we focus on the influence of statistical procedures that are less than optimal. For simple inference about a population mean, the one-sided t-test is less powerful than the optimal likelihood-ratio test, assuming independent and normally distributed observations. Here we conduct simulations to determine whether this power discrepancy is of practical significance. We assess the sensitivity of our results to departures from independence and normality. We estimate false-positive rates for both tests under various assumptions for alternative parameters, and we use Bayes' rule to evaluate implications for scientific reproducibility. Results are tabulated for reference by scientists concerned about reproducible inference for a population mean.

#### **BIOLOGICAL SCIENCES**

# X-ray Analysis of Lead (II) Binding to *H. volcanii* Malate Synthase

#### Michael Adams, Bruce Howard

Southern Utah University

Elucidation of protein structures at the atomic level using X-ray crystallography is an effective technique for studying metabolic enzymes and defining binding interactions of small molecules and ions. Our research centers on the structure of the malate synthase isoform H (MSH) from *Haloferax volcanii*, which is a key enzyme in the glyoxylate pathway of cellular metabolism and allows this organism to integrate two carbon compounds for anabolic biosynthetic reactions. We have collected X-ray diffraction data from a protein crystal soaked in a solution containing lead (II) acetate. This heavy atom derivative provided SIRAS phasing to solve the structure of the native enzyme,

but this structure has not previously been analyzed and fully refined. Here we report the iterative model-building and refinement of this structure at 2.1 Å resolution to an overall R value of 0.1815 and an R free of 0.2157. This structure allows a detailed analysis of lead (II) ion binding to the protein. In addition to the displacement of the required magnesium ion and accompanying distortions in the local vicinity of the active site, we find three additional binding sites for lead ions. Strong peaks are observed at these lead-binding sites in anomalous difference Fourier maps, as well as very high electron-density peaks in the 2Fo-Fc map at these four locations. Lead binding at inter-subunit contacts may explain the increased resolution of X-ray diffraction from this derivative versus the native protein.

#### **BIOLOGICAL SCIENCES**

# Determination of Treatments to Reduce Late Gassy Defect in Cheese due to *Lactobacillus wasatchensis* Contamination

#### Craig Oberg, Marissa Walker, Michele Culumber Weber State University

Lactobacillus wasatchensis WDC04, a new nonstarter lactic acid bacteria (NSLAB), was recently isolated from "gassy" Cheddar cheese and may be an important cause of late gassy defect in aged cheese. One way to control WDC04 may be to incorporate other NSLAB strains into cheese that inhibit its growth. Experiments were performed to determine whether inhibition occurs between common NSLABs and WDC04 utilizing the agar-flip method. A lawn of WDC04 was swabbed on MRS agar with 1.5% ribose (MRS-R) and incubated anaerobically at 25°C for 2 d or 4 d. Agar was then aseptically flipped over, and individual NSLAB challenge cultures were swabbed on the exposed surface. Plates were incubated anaerobically at 30°C or 37°C for 5 d. Growth of NSLAB cultures was compared with their growth on MRS-R plates without a WDC04 lawn (controls). In a second experiment, the media contained 4% NaCl and pH 5.2 to mimic the cheese environment. In a third experiment, MRS-R (4% NaCl, pH 5.2) was also used but the NSLAB cultures were initially swabbed as the lawn and incubated, and then WDC04 was streaked on the opposite

side of the agar. In the first and second experiments, *Lb. curvatus* WSU1 showed the greatest inhibition by WDC04 while *Lb. paracasei* Lila and *Lb. rhamnosus* 7469 were the least inhibited. All challenge NSLAB strains showed decreased levels of growth compared with control plates. In both experiments, most NSLAB cultures showed more inhibition by WDC04 grown for 4 d compared with 2 d before the culture challenge. Results from the third experiment suggest some NSLAB strains can affect growth of WCD04 under cheese-like conditions with *Lb. casei* F19, *Lb. paracasei* Lila, and *Lb. rhamnosus* 7469 exhibiting inhibition. Since there was no direct contact between WDC04 and NSLAB strains, any inhibition was due to secretion of inhibitory compounds.

#### **BIOLOGICAL SCIENCES**

# Study of an Outbreak of Skin Infections among a University Football Team in Utah

#### Marissa Walker, Craig Oberg, Joel Bass, Karen Nakaoka Weber State University

An outbreak of soft tissue infections occurred during the 2015 football season, spreading among eight football players at a university in Utah. During this two-week outbreak, eight student athletes (ATH) developed skin infections of the axilla, elbows, iliac crest, and lower leg. A roommate of an infected athlete was also infected even though the person did not play football. Aggressive wound care and empirical antibiotic treatment resolved the infections in these nine individuals. Team managers implemented mandatory showers after each game or practice and enhanced environmental cleaning, especially the locker rooms. Antibiotic treatment and the other interventions were successful in stopping the outbreak shortly after two weeks. Right and left nares of ATH were swabbed and plated onto media that was selective for the bacteria Staphylococcus. Thirty-five (31%) of 114 ATH were positive for Staphylococcus aureus (SA) in at least one nares. Thirteen (68%) of 19 roommates of the nine infected students were either infected with or carried nasal SA. Five (36%) of 14 locker mates of the eight infected ATH carried nasal SA. These data indicate that close contact may facilitate transfer of SA, resulting in infection or nasal carriage. Ten (29%) of the 35 SA isolates were resistant to at least one of 6

antibiotics tested, but only one was methicillin-resistant SA (MRSA). Eight of these 10 isolates were resistant to erythromycin and 6 of these had inducible clindamycin resistance. Several of the SA isolates had large hemolytic zones when plated on SBA plates, indicating higher virulence. These antibiotic resistance and virulence factors may have implications relating to a patient's clinical course. This study indicates the importance of hygienic vigilance and aggressive medical intervention in ending community-acquired SA epidemics. It also indicates the potential benefit of characterizing these isolates using factors not normally considered in community outbreaks of SA.

#### **BIOLOGICAL SCIENCES**

# Impact of the Integration of Produced Water on Microalgae Productivity

Brian McNeil, Jason C. Quinn

Utah State University

Microalgae represents a promising biofuel feedstock as it can be cultivated on low-quality land and can be integrated with a variety of waste streams. Produced water is the largest waste stream generated in the oil and gas industry. A variety of economic and life-cycle studies assume the seamless integration of algal cultivation systems with produced water. The integration of the two systems is synergistic as algal cultivation requires large volumes of water and there is a need for remediation of produced water. Produced water contains inorganic and organic contaminants that could inhibit growth and lipid production of microalgae directly impacting the yield of the system. In this study, produced water from the Uintah Basin, Utah, oil and gas industry was used as the primary growth media for microalgae, Nannochloropsis salina. Experiments included the evaluation of microalgae productivity cultivated on different dilutions: 0% (control), 25%, 50%, 75%, and 100% produced water. Results show that produced water severely inhibits growth of microalgae, with no growth shown in the 100% group. Results integrated into techno-economic and life-cycle modeling work to illustrate the impact of the reduced productivity on economic viability and environmental impact of a microalgae biorefinery.

#### **BIOLOGICAL SCIENCES**

# A Novel Method for Detecting Chytrid Fungus in Wild Populations of the Canyon Tree Frog

Kevin Rorie, Tucker Marsong, Curt Walker Dixie State University

We noted that the mouthparts of canyon tree frog tadpoles in Zion National Park looked abnormal in certain populations known to harbor chytrid fungus, Batrachochytrium dendrobatidis. Because reports existed in the literature that cartilage growth was adversely affected by the fungus, particularly in tadpole mouthparts, we reasoned that this might constitute a method for detecting the presence of the fungus in wild populations quickly and easily. To test our hypothesis, we swabbed tadpole mouthparts according to standard protocol for chytrid fungus detection on amphibian skin. DNA was then extracted from the swabs, and a PCR analysis was performed to check for the presence of chytrid. Results showed that none of the tadpoles with normal mouthparts tested positive for chytrid; most of those with abnormally shaped mouthparts tested positive. We therefore conclude that we have found a reliable way to quickly and easily test for the presence of chytrid fungus in wild populations of Hyla arenicolor tadpoles in the field and that infection rates can also be guickly estimated with this method.

#### **BIOLOGICAL SCIENCES**

# Evaluation of Study Skills and Lifestyle Factors Effect on Performance in Organic Chemistry: Evidence Supporting the Need for Active Learning Methods in Student Comprehension

**Don R. Davies, Heather Root, Valerie Herzog** *Weber State University* 

Student performance in sophomore Organic Chemistry courses was measured against lifestyle and study skill factors at a large, openenrollment, public university in the Mountain West region. The survey

items were first evaluated individually for correlation to student actual performance on a midterm examination. Items requiring significant student output (active learning methods), such as verbally explaining concepts, participating in study groups, and working practice exams, showed highest correlation to actual performance. Student chemical foundation, as measured by their 2<sup>nd</sup>-semester General Chemistry grade, also was a significant contributor to student success in Organic Chemistry. A factor analysis grouped the items into 4 general categories including sleep patterns, active learning activities requiring significant student contribution (study groups, verbally explaining principles), passive learning activities based mostly on student reception of information (traditional learning methods such as attending lecture, reading the textbook, and being tutored), and foundation and attitude, which included prior performance, level of anxiety experienced while taking exams, and how well they liked Organic Chemistry. The factor that correlated most strongly to actual performance was foundation and attitude, but was closely followed by active learning activities.

### **BIOLOGICAL SCIENCES**

# Analysis of Urinary Biomarkers in the Diagnosis of Depression

#### **Daniel E. Hayward, Jeremy F. Garcia, Grant E. Jones** *Weber State University*

Depression is a common psychological disorder and is frequently misdiagnosed. Currently, there is not an objective diagnostic tool for depression; clinicians must rely on symptoms of their patients to make a diagnosis. This study aims to provide clinical values of key biomarkers that will objectively differentiate between depressed and healthy patients. This study will analyze levels of brain-derived neurotropic factor (BDNF), norepinephrine (NE), and cortisol found in participant urine. Adult men and women, currently living in the northern Utah area, will comprise the sample of 100 participants. Each participant will submit a first-morning-void urine sample and respond to the standardized demographic questionnaire and depression inventory. The Beck Depression Inventory will be used to classify participants as depressed or healthy control. All analytes will be
measured using enzyme-linked immunosorbent assays (ELISA). Multivariate linear regression analysis will be used to determine how well this panel of analytes can be used to predict depression. Results will find whether patients who identify as depressed will have decreased BDNF and NE levels and increased cortisol levels. These analyte concentrations will also decrease and increase respectively based on severity of depression in the participant. The researchers hope to show that clinical psychologists and psychiatrists can accurately diagnose depression with the aid of an objective assay of BDNF, NE, and cortisol rather than solely interpreting the subjective feelings of their patients.

### **BIOLOGICAL SCIENCES**

## **Classification of Bacteriophage Isolated from the Great Salt Lake using Electron Microscopy**

#### Matthew J. Domek, Brent D. Nelson, David M. Belnap Weber State University

Bacteriophages lyse bacteria and play a crucial role in the recycling of nutrients in a halophilic environment such as the Great Salt Lake (GSL). A previous study showed that the bacteriophage CW02, isolated from the GSL, was a dsDNA bacteriophage with an icosahedral head and short noncontractile tail and belonged to the bacteriophage family Podoviridae. CW02 was also shown to share a conservative protein fold in a capsid protein originally identified in bacteriophage HK97. Very few bacteriophage isolated from the GSL have been assigned within the bacteriophage classification scheme. In this study, we attempted to classify recently isolated bacteriophage from the GSL based on morphology using transmission electron microscopy (TEM) and molecular techniques. Bacteriophages were isolated from water and soil in or near the GSL. Bacterial lysate containing bacteriophage were centrifuged and filtered to remove bacterial debris. The sample was concentrated using 100,000 molecular weight cut-off filters. Samples were further purified by CsCl density gradient ultracentrifugation. Six bacteriophages were imaged using TEM. TEM showed all bacteriophages infecting Salinivibrio costicola bacterium SA-39 were icosahedral with no detectable tail while bacteriophage infecting S. costicola bacterium SA-40 had a circular head with a long tail. This

suggests that the structure and shape of the bacteriophage capsid play an important role in the specificity of the bacteriophage to host. Based on shapes found using TEM, the bacteriophage infecting SA-39 likely belongs to the group Podoviridae while bacteriophage infecting SA-40 possibly belong to either long-tailed bacteriophage families Myoviridae or Siphoviridae.

## **BIOLOGICAL SCIENCES**

## **Chronic Effects of Intermittent Sound Disturbance of Buenos Aires Tetra**

Jeremy Arnt Weber State University

Hearing-specialist fish are more susceptible to noise than generalists. The hearing-specialist Hyphessobrycon anisitsi was subjected to two, five-week trials. One trial was subjected to intermittent anthropogenic boating noise (135±3 dB rel:  $V/\mu Pa$ ) twice per day for one hour and observed. The ambient noise trial (85±3 dB rel: V/µPa) was observed during the same time periods, and differences of growth and behavior were assessed. Growth between trials of length, weight, and Fulton's condition factor and the frequency of five behavioral categories were assessed to determine stress and habituation. We used ANOVA with repeated measure to determine difference between growth and two-tail T-test to determine differences between behaviors. In addition, simple linear regressions were performed to determine habituation/ sensitization over the trials and within one-hour noise periods. We found no growth consequences between trials. However, differences between three of the five behaviors were observed, showing sensitization for the duration of the one-hour periods. Also, sensitization was observed in three of five behaviors over the course of the five-week trials. Finally, specimens showed habituation to the initial stimulus in two of the five behaviors. These results suggest that H. anisitsi may only reach the lower stress level of the general adaptation syndrome, which would indicate slight hormonal changes leading to the behavioral change but no change in growth.

#### BUSINESS

## The White Collar Crime Registry: A Utah Original

### Chelsea Dye, Ronald M. Mano

Westminster College

White collar crime is growing across the nation and is particularly growing in Utah. For a number of years, there have been many areas of Utah with extremely low violent crime rates but increasingly high rates of white collar crime schemes. Utah has long been a breeding ground for such schemes in part because of a culture of close personal relationships in communities and a general feeling of trust for fellow neighbors and community members. In an effort to arm citizens with a resource for information. Utah has become the first state in the nation to create a white collar crime registry. HB 378 (becoming Utah Code Ann. §77-42-101, et al) passed the Utah House 65 to 7 and unanimously passed the Utah Senate during the 2015 Legislative Session, creating the "Utah White Collar Crime Registry." Much like a sex offender registry, the white collar crime registry allows users to search for offenders. The system includes the registrant's name, aliases, date of birth, height, weight, eye color, hair color, a current photo of the registrant, and a list of crimes for which he/she was convicted. The law aims to have a place where a potential investor can verify whether the person they are investing has a past history of white collar offenses. The law requires registration of a person that has been convicted of one or more of seven specified crimes: securities fraud, theft by deception, unlawful dealing of property by a fiduciary, fraudulent insurance, mortgage fraud, communications fraud, and money laundering. The offender must register for 10 years for a first offense, and additional 10 years for a second offense, and for life on a third offense. Because restitution is the part of a sentence often ignored by those convicted of white collar offenses, the law also includes a provision that allows offenders to petition for removal from the registry after 5 years if they have met certain conditions including restitution. They must also have completed any treatment ordered by the court or the Board of Pardons, not be convicted of any other crimes, had notice provided to victims about the petition, and not been found liable in any case that involves fraud, deceit, breach of fiduciary duty, or misappropriation of funds as an element. The law requires all individuals convicted of the specified

crimes after December 31, 2005, to register. However, those convicted prior to the enactment of the law can avoid registration if they have complied with all court orders, have paid all restitution, and have not been convicted of any other offenses for which registration would be required. This presentation will include a discussion of the requirements of the law as well as case studies supporting why Utah's Attorney General Sean Reyes proposed that the Utah Legislature create such a registry. The presentation will also include a discussion of the arguments advocated by opponents of the law.

### BUSINESS

## **Examining Leadership of the Global Millenial Generation in the Workplace**

#### Sean Costello, Jonathan Westover

Utah Valley University

With the young age cohort, commonly referred to as the "Millennial Generation," growing in labor force participation and prevalence in the work place, increasing interest and focus has been placed on how to best lead and manage this unique rising generation. Furthermore, major media outlets, industry publications, and academic inquiry into the topic of managing and leading the millennial generation is on the rise. However, despite the high level of attention being placed on this topic, there is a major gap emerging-while popular media and industry experts continue to insist on a large generational gap in the workplace, academic research suggests there is no huge comparative difference from one generation to the next in how individuals prefer to be led or to lead. Continued academic research in this area is needed as the Millennial Generation increasingly takes its place in the work force and there is sufficient lag time for there to be more thorough and rigorous longitudinal studies. This project seeks to build off of the existing industry research and growing body of academic studies to explore if there are any significant differences and similarities between major generational cohorts (Baby Boomers, GenX, and Millennials) with regards to how they like to lead and how they like to be led. In addition, this project will explore best leadership practices for the upcoming Millennial Generation in the work place and how to successfully apply those leadership practices in the work place. Once this is established,

the project will then focus the examination on the comparative international generational differences in the U.S. and abroad, particularly within the high-tech sector. This project will utilize secondary analysis of existing social science data sets, as well as primary research in the form of surveys and personal interviews (pending UVU IRB approval).

#### BUSINESS

## The Effects of Gender on the College Major Choice Model in Utah: What Prevents Female Students from Selecting Business as Their College Major

#### S. Paige Gardiner

Utah Valley University

Nationally, 48% of undergraduate business degrees were awarded to female students. In Utah, the number of female students selecting business as their college major was smaller. The University of Utah enrolled 28.8% females; Utah Valley University enrolled 23.3% females; Southern Utah University enrolled 35.9% females; Dixie enrolled 32.9% females: and Brigham Young University enrolled 21.7% females. The purpose of the research was to understand what personal characteristics Utah female students considered when selecting their college major. Survey research measured female students' perceptions and attitudes about possible careers in business, assessing business careers for enjoyment and flexibility. This study is significant because as the state of Utah has already realized-to be competitive in a national economy-it must increase its female representation in the workforce. This research is also significant to females in Utah and business school administrators who need to understand the college major choice model and how it differs for female and male students.

### BUSINESS

## Western States Work to Corral Fraudulent Tax Return Hackers

#### Jill Jasperson

Utah Valley University

Identity theft is a national problem, especially in the state tax submission arena. Millions of dollars are being siphoned to fraudsters that file fake tax returns. I will go through the eight western states that collect income tax (three states do not have income tax—Washington, Nevada, and Wyoming) and discuss how each state (Oregon, California, Idaho, Montana, Utah, Colorado, Arizona, and New Mexico) are handling this identity theft crisis. This information is gleaned from their state tax commission websites and press releases. I will compare and contrast the efforts of each state: Are they accentuating the fraud issues or not really notifying the public of the problem? Is the taxpaying public being informed of the risks they are taking by filing electronically? Are the states creating measures to allay public fears? Is there plenty of information on how to file a fake return? These and other issues will be discussed.

### BUSINESS

## **Conditions of Shareholder Wealth Maximization Ethicality under Classic Philosophical Paradigms**

**Rebekah Inez Brau, James C. Brau** Brigham Young University

In many business textbooks and classes, students continue to be taught that the goal of the firm is to "maximize shareholder wealth." The shareholder wealth maximization (SWM) goal is also the utility function in most finance and economics academic articles. On its surface, this goal may seem a bit callous, or perhaps even unethical. The goal may even contribute to some negative stereotypes of capitalism or Wall Street. In this paper, we analyze not if, but when SWM can be an ethical goal of the firm. Beginning with the seminal work of Berle (1931), Manne (1959), and Friedman (1962), we present a discussion of the goal of shareholder maximization within the theoretical construct of four main traditions of ethics thought: Deontology, Justice/Fairness, Utilitarianism/Consequentialism, and Virtue Ethics. We provide necessary and sufficient conditions under each ethics paradigm to determine whether SWM is ethical or not. Under several of the ethical camps, the conditions are very narrow for SWM to qualify as an ethical goal.

#### BUSINESS

## The Determinants of Achievement in an Introductory Marketing Class

## Stephen Owen, James C. Brau, Mike Swenson

Brigham Young University

We begin this paper with a careful review of previous literature on the determinants of academic performance. We perform extensive analysis on past literature including that of Harris (1940), which reviews relevant literature from 1930 through 1937. In his study of 328 articles on potential determinants of academic performance, he develops categories such as intelligence, high school grades, study habits, teaching methods and conditions, incentives and direct motivation, amount of course work taken, and extra-curricular factors. Gender, attendance, and course interest are among other significant variables found within more recent research. Subsequently, we perform a survey of 835 undergraduate students at the end of a semester of an "introduction to marketing" class at a large private university. With the help of prior literature, other faculty members, and students, we created a survey of approximately 40 questions that could be potential factors of course performance. The factors addressed include the aforementioned as well as other program-specific factors. The primary research objective is to provide students and instructors information on factors that are most significant to their learning and course outcomes. This will allow both students and instructors to give focused efforts on the variables that provide the greatest marginal benefit for student learning and outcome in an introduction to marketing class.

### BUSINESS

# Artificial Intelligence, Business Law and Mr. Batman

#### Shadman Bashir

Dixie State University

Law can be dry and uninteresting for non-lawyers. A system of laws are the platform on which a society stands, but the myth and reality of legal lingo and jargon often keeps it and its practitioners separated from social and non-legal academic thought process. This paper is a comparative analysis of everyday legal hypothetical questions and their possible answers. This includes, but is not limited to, some of the most popular super heroes, and the legal scenarios in which they get involved in context of the Bill of Rights. Could it be possible for the courts to admit legal evidence collected by Batman? Could someone be tried for attempted murder of someone who can't die? Within the legal arena, we are used to dealing with a natural and the legal person. Are we prepared for an artificial person? In other words, what will happen when an illegal act is committed by an artificial intelligence machine? Can a machine that kills claim an insanity defense and get away with murder?

### BUSINESS

# The Dilemma of Maintaining Relevance of the CPA Designation

#### Jennifer Harrison, Ronald M. Mano

Westminster College

The post-nominal letters CPA compose the mostly widely recognized symbol of the accounting profession within the United States—that of the Certified Public Accountant. Among the myriad of accounting certificates, it is singular in that it is the only accounting license. While each state is the regulating body behind licensure and has a state-run professional accounting association, the American Institute of Certified Public Accountants (AICPA) is the national accounting professional association. The AICPA represents the accounting profession through

advocacy and education of the public, legislative entities, and accounting professionals. While the AICPA's influence is considerable, that influence has diminished since its founding in 1887. As of the 1930s, the Securities and Exchange Commission (SEC) has authority to make accounting standards for publically traded companies-with responsibility granted to the Financial Accounting Standards Board (FASB). Additionally, with the passage of the Sarbanes-Oxley Act early in the 2000s, the Public Company Accounting Oversight Board now is the authoritative body with oversight responsibilities over most aspects of an accountant's professional practice. Finally, with the importance of access to global capital, and thus globally accepted reporting standards, the AICPA must find its standing among other national accounting-professional organizations. Membership in the AICPA, which until 2015 was comprised of only CPAs, is voluntary. Like any association, the AICPA must provide compelling reason for those who are qualified to join to willingly do so. The AICPA must prove its relevance to potential members. As such, the AICPA's current mission statement is: Powering the success of global business, CPAs, CGMAs, and specialty credentials by providing the most relevant knowledge, resources, and advocacy, and protecting the evolving public interest. To accomplish its mission, the AICPA is involved in many activities-some old, some new. Of interest to us is the inclusion of "CGMAs and specialty credentials" in attracting membership and maintaining relevance. In this paper, we will examine the proliferation of accounting certifications. More specifically, we will focus on those certifications sponsored by the AICPA and discuss whether the introduction of new accounting certificates dilutes or augments the value of the CPA designation. In an effort to maintain global relevance, is the AICPA deploying a brilliant strategy or becoming its own worst enemy?

#### BUSINESS

## Another Piece of the Corporate Payout Puzzle

#### Lauren Lo Re, Rob Patterson, Mahfuz Raihan Westminster College

Corporate payouts are economically significant and have fluctuated dramatically for decades. Several researchers have documented these

trends and waves in payout activity; however there has been little explanation for why these patterns occur. Existing literature focuses on examining cross-sectional variation in firm-level characteristics as the basis for explaining payout policy, yet questions remain. In this paper, we examine the role of macroeconomic factors in the payout equation, while controlling for firm-level factors, and find that payout activity fluctuates with changes in the business cycle. We find that IPI is a significant factor in explaining both aggregate dividend and repurchase activity. Specifically, IPI, as a measure of aggregate output, operates within a feedback loop in which economic activity drives investment and investment drives economic activity. In periods of increased investment, aggregate payout declines. Short-term interest rates also significantly influence aggregate dividend and repurchase activity. The theoretical framework suggests that as costs of external finance increase, earnings decrease, and payouts decrease. We find evidence in support of this and find a significant, negative relationship between short-term rates and payout activity. Given that economic growth relates to each of these factors, our results indicate that payout activity and form fluctuate with changes in the business cycle indicating that changes in the business cycle contribute to the explanation of why many firms initiate payout activity, and select the same form of payout, at the same time. Our results indicate that waves of payout activity result from responses to the broad economic environment as well as firm-level characteristics. We find that changes in the macroeconomic factors contribute to an explanation of aggregate payout trends in terms of both level and form of payout.

### BUSINESS

# **Student Golf Tournaments: "Speed Dating on the Golf Course"**

#### Michael Lauret, Charlie Leveroni, Ronald M. Mano Westminster College

This article discusses the development of student golf tournaments at two separate institutions. It discusses why they are valuable as well as a particularly unique setup that accomplishes the golf tournament goals better than any other format that these authors have ever heard of before. Also discussed are the side benefits that have been experienced through the golf tournaments.

#### BUSINESS

## **Exploring Shifting Global Labor Management Practices and Comparative Job and Life Satisfaction**

#### Jonathan Westover, Joe Light, Kaitlin Carlisle, Bergen Eski Utah Valley University

According to recent studies, Denmark consistently enjoys greatly higher levels of life satisfaction than other countries throughout the world. Despite geographical proximity, even neighboring countries do not achieve the same level of satisfaction as the Danish. Additionally, a growing body of research has shown that shifting global labor management practices have a large influence on changing worker attitudes and values. More specifically, various intrinsic and extrinsic motivators in the workplace environment have been shown to improve both worker job satisfaction, as well as their global life satisfaction. This research utilizes comparative data from multiple waves of the World Value Survey to explore and examine the possible reason for these country differences and to parse out the many variables that contribute to varying job satisfaction and life satisfaction levels across nations, with a specific comparative focus on Denmark, Belarus, and the United States.

### **EDUCATION**

## The Effects of Dual Credit Enrollment on Higher Education and Labor Market Outcomes

### Richard Haskel, Peter Seppi, David Tille Westminster College

This study considers the effects of Dual-Credit Enrollment and Early College High School (ECHS) programs on higher education and labor

market outcomes for Utah's 2008 and 2009 public high graduation cohorts via an examination of the Utah Data Alliance longitudinal public education data set, one of 46 state longitudinal data systems (SLDS) under development in the U.S., including a consideration of the issues involved in initiating similar SLDS studies outside of Utah. The study assesses high school graduation rates, dual course credits earned, higher education enrollment, time-to-completion, graduation, and early labor market outcomes with a focus towards how these accelerated learning programs affect the student household and state, specifically, and public education finance, generally. As participation in dual-credit programs is voluntary and by self-selection, the study employs propensity score matching (PSM) method as a quasi-experimental design methodology in an effort to limit the endogeneity bias present in such nonexperimental data. Although PSM offers many advantages, its strength as an estimator is dependent on the existence of complete and quality matching variables. To assure accurate model specifications given the available data, receiving operator characteristic (ROC) analysis is applied to variations on the PSM models. Estimated outcomes reflect positive effects for the examined student populations differentiated by program participation, with the strongest outcomes arising from ECHS participation. The economic effects of accumulating higher education course credits and decreases in higher education time-to-completion may yield the most interesting outcomes, enjoy the strongest causal claims, and result in measurable householdand state-level savings. These outcomes may also reveal potential weakness in the structure of higher education course and major programming and the difficulty presented as high school students make higher education decisions. Certain challenges involved with using other SLDS data sets to form similar research studies are presented with possible solutions considered.

### **EDUCATION**

# Why We Need Woodshop Again: A Look into the Purpose of School

#### **Dorothy Eilertsen** *Weber State University*

Despite government-mandated programs and strategies in pursuit of a stronger school curriculum, more competitive test scores, and high success rates among a diverse range of students, it is what we aren't teaching that is failing our upcoming generation. Even after programs such as the Goals 2000: Educate America Act, Approving America's Schools Act, and the No Child Left Behind Act, our country's students fail to receive an education befitting a leading nation of the world. In fact, the purpose of school among higher grades has been lost in the myriad efforts to enhance it. In this paper, I will explore the educational shortcomings of our upper grade school programs and the necessary changes that could facilitate growth and reshape the education of the growing generation of 21st-century Americans. Specifically, I argue that school has become an institute of correct answers, rather than a refuge for innovation. I propose the need for home economics, woodshop, and basic money management, and other such classes where knowledge of the answers is applied to produce relevant skills. This paper will draw upon ideas and attitudes shared in the 2015 Sundance Film Festival official selection titled "Most Likely to Succeed," a documentary by Ted Dintersmith about what and how students learn and the increasing need to rethink the purpose of school. Furthermore, the film suggests that important life skills are not necessarily learned by memorizing information in the classroom as much as by getting up and doing something that can be transferred into the professional field. If 53% of recent college graduates are jobless or underemployed, as the film indicates, then there must indeed be a change in the dynamic of education.

## **EDUCATION**

## Benefits of Computer-Mediated, Student-Centered, Collaborative Learning in the Human Development Classroom

Dana Erskine Utah Vallev University

A paradigm shift is taking hold in American higher education. Colleges are moving away from being institutions that exist to provide instruction toward being institutions that exist to produce learning. In an effort to contribute to this shift, an in-class research study was conducted that paired student-centered, computer-mediated, and collaborative learning techniques in the Human Development classroom. This type of instruction shares power with the students by allowing them to pose and seek out answers to their own questions (student-centered) through the analysis and evaluation of information found in professional journals (computer-mediated) while working as teams (collaborative learning) in the classroom. Many significant results were found using these techniques, including enhanced student ability to re-evaluate prior subject matter knowledge and improved ability to translate empirical information into their own words.

## **EDUCATION**

## **Grading for Presence: Participation Grades in Post-Secondary Institutions**

Erin M. Walton Salt Lake Community College

The focus of this research is the effect of a participation grade upon students in a post-secondary learning environment. The study evaluates whether requiring students to be in class assists in achieving higher final and cumulative scoring. This research was performed on two sections (one with the participation grade and one without) of the same course taught over a five-week period and included 32 participants. It was discovered, using both narrative and survey data, that the section containing the added grading aspect performed significantly better than the one without. It can be concluded that the addition of the participation grade has the potential to increase student learning and retention. It was also noted that students who were required to attend class had a higher rate of course satisfaction on their final course evaluations. Finally, upon presentation to colleagues, the research was found to be relevant in their respective classroom settings. This suggests a wider reaching potential for instructors seeking to increase student learning outcomes.

### **EDUCATION**

## Combining Classics with Contemporary: Incorporating Young Adult Literature in the Secondary Classroom

Alyssa Devey

Brigham Young University

One of the best ways to help students enjoy reading and become better readers is to incorporate young adult literature (YAL) in the classroom. Unfortunately, many teachers struggle to utilize YAL. Minimal access to YAL, time constraints, common core testing, and the sometimes difficult content of young adult novels create barriers to teaching YAL. Several articles point out the need to teach YAL, but as Edward Sullivan points out, most of the articles are written by "university professors teaching adolescent literature" instead of by secondary teachers who could provide practical means of incorporating YAL into the classroom (Sullivan et al. 11). Although these articles do make important arguments to convince those who are not yet willing to incorporate YAL in the classroom, this article focuses on helping the already-willing teacher to overcome the hurdles of incorporating YAL. For this reason, I discuss a list of practical suggestions for how secondary teachers can incorporate more YAL with what they are already doing with classic texts.

## **EDUCATION**

## Evaluation of Study Skills and Lifestyle Factors Effect on Performance in Organic Chemistry: Evidence Supporting the Need for Active Learning Methods in Student Comprehension

**Don R. Davies, Heather Root, Valerie Herzog** *Weber State University* 

Student performance in Sophomore Organic Chemistry courses was measured against lifestyle and study skill factors at a large, openenrollment, public university in the Mountain West region. The survey items were first evaluated individually for correlation to student actual performance on a midterm examination. Items requiring significant student output (active learning methods), such as verbally explaining concepts, participating in study groups, and working practice examinations, showed the highest correlation to actual performance. Student chemical foundation, as measured by their 2<sup>nd</sup>-semester General Chemistry grade, also was a significant contributor to student success in Organic Chemistry. A factor analysis grouped the items into 4 general categories, including sleep patterns, active learning activities requiring significant student contribution (study groups, verbally explaining principles), passive learning activities based mostly on student reception of information (traditional learning methods such as attending lecture, reading the textbook, and being tutored), and foundation and attitude, which included prior performance, level of anxiety experienced while taking examinations, and how well they liked Organic Chemistry. The factor that correlated most strongly to actual performance was foundation and attitude, but it was closely followed by active learning activities.

### ENGINEERING

# Throttled Launch-Assist Hybrid Rocket Motor for an Airborne NanoSat Launch Platform

## Zachary Spurrier, Stephen Whitmore, Sean Walker, Stephen Merkley

Utah State University

This document details the design, integration, and testing of a throttled launch-assist hybrid rocket motor for an airborne nano-launch platform. Gaseous oxygen and additively manufactured Acrylonitrile Butadiene Styrene (ABS) are used as the propellants. This study establishes the requirements for this launch-assist propulsion system, develops the system design features, develops a closed-loop proportional throttle control law. The detailed end-to-end system design is presented. Initial static tests were performed with a cylindrical fuel port to verify system functionality and establish a baseline for the propellant regression rate and optimal oxidizer-to-fuel (O/F) ratio. Subsequent tests are performed using a helical fuel port to increase the volumetric efficiency of the system and allow operation near the optimal O/F condition. Multiple restarts of each system configuration are demonstrated. Results of both open- and closed-loop throttle tests are presented.

### ENGINEERING

## A New Perspective to Antenna Design for Reconfigurable Wireless Networks

Mehedi Hasan, Israfil Bahceci, Bedri A. Cetiner

Utah State University

In recent years, wireless communication technologies have been undergoing a period of unprecedented growth leading to a dramatic increase in data traffic. Network concept like internet of things, build on cloud computing and networks of data-gathering sensors, indicates an exponential increase of information exchange. This calls for reconfigurable wireless networks where each layer of the network shall be reconfigurable to account for a diverse set of operational

requirements. Subject to these diverse operational requirements, an efficient reconfigurable physical layer shall consist of reconfigurable antennas. In Utah State University, a new class of antennas called multifunctional reconfigurable antennas (MRAs) has been developed. Parasitic tuning based reactive surfaces on top/around driven antenna surface give MRAs the capability to dynamically modify resonant frequencies, polarizations, and radiations patterns. Until now, while designing an MRA, the considered parameters did not take into account spatial and temporal statistics of multipath propagation the environment, e.g., angular spread, path correlations, coherence time, and bandwidth of the underlying channels. Link level analysis shows that while the traditional design parameters are important, the statistics of radio channel cannot be overlooked. For example, channel correlations can play vital role to increase spectral efficiency and improve the error performance. This observation gives a new perspective to antenna design. Simulation results indicate considering statistics of propagation environment along with traditional parameters while designing antennas improves overall system performance.

## ENGINEERING

## Fabrication of Dye-Sensitized Solar Cells using Different Nanocrystals in Ferritin as the Dye

### Alessandro Perego

Brigham Young University

Solar energy is frequently lauded as a potential game changer in the energy landscape but, unfortunately, commercially available photovoltaic technologies are based on inorganic materials (mainly silicon), which require high costs and highly energy consuming preparation methods. Dye-sensitized solar cells (DSSCs) present a valuable and sustainable alternative to silicon solar cells. DSSCs have been studied for the past 20 years because of their simplicity in fabrication, but little progress has been made concerning their energy efficiency conversion (best reported efficiency is  $11.9\pm0.4\%$ , compared with 25.6% for silicon solar cells). Finding ways to improve the light harvesting and electron transfer reactions of the dye is the key to improving the energy conversion of DSSCs. Ferritin is a 12-nm-diameter spherical protein with an 8-nm hollow interior, which

naturally contains iron oxide nanocrystals. The natural core of ferritin can be removed and other metal oxide nanoparticles can be synthesized inside the empty ferritin. The choice of metal used in the growth of the nanoparticles determines the wavelengths of light that can be absorbed. The Watt and Colton laboratories at Brigham Young University have been studying these ferritin nano-architectures to quantify and characterize the light harvesting and oxidative charge separation reactions of ferritin to tune the wavelength of light that is harvested by each nanocrystal. Theoretical efficiencies have been calculated and they can reach up to 44.9% of solar energy conversion. Additionally, ferritin possesses the ability to prevent photocorrosion in metal oxide semiconductors, and it is also thermo-stable up to 80°C. These unique properties make the ferritin nano-architecture an intriguing photocatalyst for DSSC cells. Engineering a working device that combines the technology of DSSCs with the proprieties of ferritin has the potential to revolutionize the market of solar energy.

### ENGINEERING

# Security Imaging Using Wi-Fi-based Channel State Information

#### Khem Narayan Poudel, David Schurig, Neal Patwari University of Utah

This paper presents a novel Wi-Fi-based investigation system that uses a wireless channel state information (CSI) about recent past activities for security imaging. The proposed system explores a physical layer channel state information using 30 subcarriers for different position, pose, size, and speed of violated object. This system uses both magnitude and phase values with frequency and spatial diversity–based Multiple Input Multiple Output (MIMO) system to specify unique activity. We propose that if a crime is committed, investigators download channel state information records from nearby receivers and reconstruct the image of recent state on similar boundary condition and multipath propagation model using computational imaging techniques. We conducted the experiment using off-the-shelf IEEE 802.11 devices and analyzed using full-wave simulation in CST microwave studio. Our results from extensive experiments and simulation demonstrate the overall investigation can be improved compared with existing approaches and achieve better identification, which is important for many application in security, search and rescue.

### ENGINEERING

# FDTD: A Powerful Tool in Computational Electromagnetics

## Khem Narayan Poudel, Santosh Pokhrel

University of Utah

The Finite-Difference Time-Domain (FDTD) method is a powerful tool for modeling any types of electromagnetic (EM) applications. It employs second-order finite centered approximation to both temporal and spatial derivatives in solving the Maxwell's equations in time domain. These equations incorporate the compatible parameters (permittivity, permeability, and conductivity) to solve EM problems in any media such as homogeneous-inhomogeneous, dispersivenondispersive, and isotropic-anisotropic. The approximated solution obtained thereafter must be stable in both space and time to avoid attenuation and undesirable growth of wave. This can be achieved through the proper selection of Courant stability factor (S). This factor is defined as the ratio of product of wave velocity and time step to spatial grid length. The progress in FDTD highlights the year 1966 when Yee proposed a space and time staggering grid (Yee cell). Then in 1994, Berenger's split field formulation of Maxwell's equations laid the foundation for perfectly matched layer (PML). This novel technique helps to solve the problem in unbounded regions by creating absorbing boundary layers. These boundaries are especially designed to absorb plane waves of any polarization, frequency, and angle of incidence completely without reflection. The FDTD is applied to problems across the EM spectrum varying from very low frequency (VLF) to the optical applications frequency range. The most common include communication, ionospheric remote sensing, earth modeling, global positioning system, radar, hypothesized earthquake precursors, and space weather effects on the near-Earth environment.

#### ENGINEERING

## A Comparison of Solar Tracking Algorithms Utilized by the SAM Instrumentation

#### Zakk Rhodes, Gene Ware

Utah State University

The Sun and Aureole Measurements (SAM) tracker is a ground-based measurement system for atmospheric particles that tracks the solar disk and records data about the sun halo, aureole, and upper-level atmosphere. To take accurate images from which to extract the atmospheric and solar data, SAM must track the Sun within a low margin of error. This can be accomplished through implementing a few different methods of analyzing the image data to find the solar disk and correcting the tracker mount angles to more precisely align SAM with the solar disk. The comparison of three possible algorithms for locating the solar disk is discussed: the pixel-average method, Canny-edge detection with a Hough parametric transform analysis, and the current process used by Visidyne, Inc.

#### ENGINEERING

## The Impact of Circumsolar Radiation on CSP Renewable Energy Harvesting

## Gene Ware, Doran Baker, Zakk Rhodes, Mark Norman, Alireza Ghasempaur

Utah State University

A significant amount of the solar radiation harvested by concentrated solar power (CSP) plants occurs in the circumsolar region around the solar disk. The energy in this region appears to be primarily due to the forward scattering of the energy from the solar disk. Scattering mechanisms are reviewed and applied to CSP power generation. Circumsolar measurement using the Sun and Aureole Measurements (SAM) instrument is discussed.

# LETTERS—FOREIGN LANGUAGE, HUMANITIES, PHILOSOPHY

### Where has Nabonidus gone?

Spencer C. Woolley

University of Utah

"Nabonidus, the great king, the strong king, the king of Babylon, the king of the four corners" eludes easy comprehension. According to both his own words and the denunciations of his enemies, he discarded the traditions of his kingly forebears and did his own thing. But what his own thing was, and where he did it, and why he departed from precedent remains a puzzle. Nabonidus slips into and out of Mesopotamian history like a Euphrates carp, wriggly and difficult to pin down. This paper may then be viewed as a fishing expedition-one with pronged spears to skewer the recalcitrant creature from foamy rapids, rather than a net cast over still waters. The ancient documentary sources will be examined: Nabonidus's inscription at Harran, the Babylonian Chronicle, and the Cyrus Cylinder. Given the paucity of ancient sources, an analysis will also be made of the various excavations at Tayma. This paper argues that while the archaeological and documentary evidence do not wholly support the standard reconstruction of Nabonidus's reign, that standard interpretation cannot be cast away; Nabonidus does not dwell at Babylon, but the evidence remains weak that he dwelt at Tayma. If he does reside at Tayma for the decade that the sources state, he does not reside in the usual fashion of a Mesopotamian potentate.

# LETTERS—FOREIGN LANGUAGE, HUMANITIES, PHILOSOPHY

## That's Rich: Survival in the Tudor Court

#### Joshua C. Wiggins

Salt Lake Community College

It has been well documented how merciless the Tudor court could be. Men and women could garner incredible power and wealth so long as they were in the monarch's good graces, but could be swiftly destroyed when they (almost inevitably) fell out of favor. Thomas Wolsey, Anne Boleyn, and Thomas Cromwell all fell victim to this cycle. One man spectacularly broke the trend by becoming more corrupt than the system he served and living to have it documented. This man died peacefully of old age, surrounded by the enormous fortunes he had gathered. This was Sir Richard Rich. This paper will examine some significant moments in Sir Rich's life through primary sources and historical analysis to document the motivation for everything he did: survival at court.

## LETTERS—FOREIGN LANGUAGE, HUMANITIES, PHILOSOPHY

## **Russian California, Outpost of Empire**

#### **Brian Simons**

Utah State University

In the first half of the 19<sup>th</sup> Century, the Russian Empire expanded as far eastward as Northern California. Expansion into California was done to procure furs for a high-demand Chinese market and was a link in a global trade network spanning the North and South Pacific. Many different native groups and colonial powers played important roles. The Russian treatment of Native Californians is highly romanticized by scholars as an example of a positive form of colonialism, one in which natives were treated humanely and both sides gained mutual benefit. This is in contrast with the brutality experienced by native peoples elsewhere in the Russian Empire, as well as that committed by the neighboring Spanish in California. A close examination of newly accessible sources, in conjunction with dismissal of many less-veritable documents, sheds new light on interpretations that have been replicated for decades. A reading of Russian- and Spanish-language sources demonstrates that the Russian experience in California is directly in line with Russian expansion in other regions. Despite open criticism of Spanish brutality toward natives, the California Russian settlement directly supported and traded with the California Missions to supply the Alaskan colonies with grain. Elements of positive Russian-native interaction were largely pragmatic, were dictated by a number of outside factors, and in no way constituted a new set of higher colonial ideals. Russian California was governed and guided by the same principles of colonialism that governed both prior and simultaneous Russian expansion elsewhere in the Tsarist Empire.

## LETTERS—FOREIGN LANGUAGE, HUMANITIES, PHILOSOPHY

# No Longer Either French or Anything: The Situation of Jewish Women in North Africa during World War II

#### Avenel Rolfsen

Westminster College

The Holocaust has become one of the quintessential discussions of history, and as such, much has been written about the catastrophic inhumane actions the Nazis levied against the Jews of Europe during World War II (WWII). The vast majority of this writing focuses on the plight of European Jews during this time. What many people are inclined to forget is that the Jews of Europe were not the only ones affected by this ghastly event. Undoubtedly, the Jews of Europe received the short end of the stick; however, WWII also affected North African Jews. In Tunisia, Algeria, and Morocco, the racial laws of Vichy France greatly affected the lives of 400,000 Jews. Literature discussing the Jewish situation within North Africa is minimal and limited in scope. Most of this literature is marginalized within larger surveys of North African history. Very little focuses on North African Jews during WWII; even less focuses on the lives of Jewish women in the region at this time. In contrast to Jewish women's protests and rebellion in Vichy France, North Africa's Jewish women are strangely quiet. This paper examines Jewish women's lives in North Africa during WWII and argues that they actively mitigated their oppression in many ways. Jewish women created makeshift schools, took new jobs, engaged in prostitution, and protested food rations and the absences of their husbands and sons. This research fills a large gap in North African and women's history; it also adds to the expanding literature on Holocaust and women's history. It presents a world in which Jews and women are not simply oppressed but one in which they work to diminish their oppression.

# LETTERS—FOREIGN LANGUAGE, HUMANITIES, PHILOSOPHY

### The Mechanics of Scientific Belief

#### **Michael Warren Cook**

Westminster College

Does it matter that scientists operate with faith in unfalsifiable theories? If a scientist practices science in the same manner that a cleric practices theology, then what separates the two? If a scientific community practices a level of belief, or even faith, does it lose its secular credibility for explaining reality? In making his criterion for scientific theory, Karl Popper transforms much of modern science into a faith-based system. However, at the same time modern scientific communities may use unfalsifiable beliefs to explain reality, the difference between clerics and scientists lies in the formation and function of their respective unfalsifiable theories. Thus, remarkably, acknowledging the belief at the heart of science actually delineates the practice of unfalsifiable science from the practice of theistic religion— and it illuminates a paradoxical path toward faith-based secular discovery.

## LETTERS—FOREIGN LANGUAGE, HUMANITIES, PHILOSOPHY

# **Resettlement Experiences of Children who Entered the United States as Refugees**

#### Elizabeth Katherine Gamarra

University of Utah

According to the United Nations High Commissioner for Refugees in 2014, approximately 16.7 million refugees exist globally and nearly half are children. As families become acclimated and accustomed to the U.S culture, they face numerous challenges. In fact, according to this research study, children face a significant number of stressors during resettlement, impacting them within their family structure, among their peers, and in different social interactions. This qualitative study sought

to answer the question: What are the core issues confronting children ages 8–14 with a refugee background as they resettle in the U.S.? This study identifies and explores six core themes inclusive to school, emotional health, cultural identity, social interactions, laws and safety, and changed family dynamics. The development of effective interventions remains a neglected research area. Literature is also limited in terms of this age group-latency-age children-and therefore, this study seeks to fill this gap by adding to a curriculum in development. The themes identified above indicate that although there are positive aspects of resettling for families to experience, more needs to be done to support parents and children in their adaptation and transition to the U.S. Findings indicated that unattended needs concentrated more in social, emotional, and cultural areas as during resettlement they focus on meeting basic needs of survival. Findings therefore suggest that communities, social agencies, and schools should devote additional resources and time to facilitate transition for children and their parents.

# LETTERS—FOREIGN LANGUAGE, HUMANITIES, PHILOSOPHY

# A Philosophical Critique of St. Thomas Aquinas's Theory of Reason and Revelation

#### Hannah Lee Brau

Brigham Young University

St. Thomas Aquinas, the Italian philosopher, believed that there were two sources of knowledge: reason and revelation. This theory holds significant implications for the academic disciplines of both philosophy and theology. The concepts have been discussed intensely since the thoughts of St. Thomas Aquinas, and there are still differing opinions about his theory. I plan to write a critical paper about St. Thomas Aquinas and how he sees the relation between reason and revelation. I will include a critical response to Aquinas and what I think about this subject. Sourcing examples in historical texts, such as the story of Adam and Eve from the book of Genesis, I will give a synopsis of popular western understanding of the topic. Next, I will address St. Thomas Aquinas's view on the issue, citing his own texts and other texts written by philosophers around his time. Last, I will write a critical response to Aquinas and his work, referencing sources that have led to my beliefs on the topic. These sources will include scriptural references, as well as philosophical works. I will analyze and explain why I feel the way I do, taking into consideration the thoughts of philosophers and theologians who have addressed reason and revelation.

## LETTERS—FOREIGN LANGUAGE, HUMANITIES, PHILOSOPHY

## An Academic Librarian's Interpretation and Brief Synopsis of a Historical Text of Scroll 8 HEV1 from the Dead Sea Scrolls

#### Peter L. Kraus

University of Utah

In the fall of 2013, an exhibit of the Dead Sea Scrolls was on loan to the Leonardo Museum in Salt Lake City, Utah. For an academic librarian, this was a once-in-a-lifetime opportunity to examine parchments that were created between 408 BCE to 318 CE not as religious texts, but as physical objects. The duplication of a number of these objects relates to modern library and archival best practices as defined by the practice of LOCKSS (Lots of Copies Keep Stuff Safe).

## LETTERS—FOREIGN LANGUAGE, HUMANITIES, PHILOSOPHY

## A Series of Glimpses: Witnessing to World War II Through Memory and History

#### Adriana Pinegar

Westminster College

The relationship between history and memory is long and complex. While some theorists argue that they are at odds with one another, this presentation will explore the necessary relationship between the two. Using Michael Frayn's 1998 play, Copenhagen, and the scrapbook of a Danish police officer and resistance fighter during World War II, the presenter will posit the central role of uncertainty in the negotiation of individual memory and history. The position of the observer or witness to history affects the way the past is remembered and recorded. Individual witnesses, even and perhaps especially where they stray from the accepted historical narrative, testify to something that would otherwise be lost: the nature of the event. The observer therefore plays an important role in interpreting the testimony according to its place in the flow of time.

# LETTERS—FOREIGN LANGUAGE, HUMANITIES, PHILOSOPHY

# The Modern Hollywood Sequel: An Analysis of Theme and Story

#### James Nick Reddoch

Salt Lake Community College

Since "The Godfather Part II" won critical acclaim as the most successful sequel, Hollywood has attempted to bank off the movie sequel. This paper explores the similarities in theme and story in successful mainstream American films, tracing a lineage from "Part II" and "The Empire Strikes Back" to the recent sequels in other genres. Other works explored include "The Dark Knight," other recent superhero films, police action films from the 1980s, and comedy. Using those two films as a foundation, other works are examined to find a pattern of plot and theme that is recycled to reach the most emotional connection to an audience. A related pattern of similar themes is reviewed that shows that Hollywood writers rely on the same tropes, metaphors, and sometimes even settings to make a sequel feel more personal than the original film. One major example seen repeatedly is the use of returning to a time or setting before both films, which is one way a film adds a tone of nostalgia. A common occurrence seen repeatedly throughout sequels is an attack or invasion inside the home or personal space, such as the assassination attempt in "Part II" or the invasion of the rebel base in "Empire." A feeling of loss is present, then a return or retreat where the protagonists must recoup and learn new skills or knowledge. The effect of these common storylines and themes are designed to make the audience feel closer to the central character, and with that closeness feel a more personal film than even the original, but also giving the audience the need to explore more with new locations and information. Avoiding the pitfall of giving more and exploiting what worked in the previous film, successful sequels tend to rein it in for a more personal film.

### LETTERS—LITERATURE

## "Thy Placeless Power": Herman Melville, Mobility, and the Poetics of Placelessness

#### **Todd Goddard**

Utah Valley University

In this paper, I investigate Melville's unease with the erosion and absence of abiding places, which in turn are linked to proliferating spatial mobilities in the first half of the nineteenth century. While literary critics have pointed to Melville's celebration of travel and the freedoms of movement, few have commented on his concern for the consequences of mobility on the integrity of place and the implications of placelessness for identity and authorship alike. What we see in Melville's works, I argue, particularly in Moby-Dick (1851) and culminating in The Confidence-Man (1857), is a place-anxiety or placepanic that derives in part from the accelerating velocities of modernity and a sense of the increasing loss of stable, bordered, and bounded places. In extreme forms, such placelessness figures forth in his works as the void, a boundless and eternal "Nothing" akin to infinity and zero wherein all distinction of identity and place disappear. At other times, Melville registers other permutations in the increasing loss of distinct and delineated places, in natural erosion and ecological decay, in the rapid building up and wearing out of place, and in the transportation technologies that underwrote many of these phenomena. In attempting to articulate these subtle (and not so subtle) transformations-many on the edge of semantic availability-that characterize the period known as "the great acceleration," Melville adopts an aesthetic and style that punctuate, even as they mimic, this placeless mobility and instability.

### LETTERS—LITERATURE

## Temporarily Embarrassed Millionaires: F. Scott Fitzgerald's "Winter Dreams"

#### **Elizabeth Robison**

Weber State University

American novelist John Steinbeck said: "Socialism never took root in America because the poor see themselves not as an exploited proletariat but as temporarily embarrassed millionaires." Self-perpetuating and potentially dangerous, the American dream builds upon itself, and like the capitalist agenda it has been created to support, it has an uncanny ability to absorb any kind of contradiction. It propagates this elusive dream by offering hope of a better life. Maggie Combs says: "...the American dream is a veneer that keeps people from seeing the reality of social classes." One of the best examples of this American dream "veneer" is seen in F. Scott Fitzgerald's short story "Winter Dreams." We see Fitzgerald's main character, Dexter, chasing this dream wholeheartedly through his commodification and the obtainment of the ultimate status marker, the fabulous Judy Jones. Driven by the false consciousness that is the American dream, Dexter is left with nothing but broken promises and personal devastation. Fitzgerald shows that Dexter's pursuit of the American dream truly is a "winter dream," cold, empty, and unfulfilling. Fitzgerald demonstrates the oppression of the ideology of the American dream by showing how it has served no one in his story, not even the little guy, Dexter, who experiences success and some upward mobility only to be left broken in the end. This alienating effect leaves the reader asking questions about his own winter dreams and Judy Joneses. The American dream, like Judy, is not truly desirable. Like Dexter, it leads us on for a lifetime with no ability to fulfill its promises of unending happiness through wealth and success. The commodification, status markers, and false consciousness of the American dream serve no one in Fitzgerald's work, and while promising happiness, they offer nothing but long-term dissatisfaction.

#### LETTERS—LITERATURE

### The Rhetorical Education of Matteo Ricci, S.J.

#### **Roberto Leon**

Brigham Young University

Matteo Ricci, S.J. is world-renowned for his accommodating style to preaching Christianity in sixteenth-century China, in which he excelled at studying classical Chinese as well as made significant contributions to the first Latin translation of the Four Books of Confucianism. Many historians focus on the trial and error associated with his appreciation for Chinese culture, without considering the effect of his unique education on his choices. Ricci was taught with the innovative curriculum of the Jesuits. Generally, scholars have been quick to label sixteenth-century Jesuit rhetoric as being strictly in line with the Scholastic tradition, on account of its reliance on Aristotle's Organon. A closer look at their curriculum, however, reveals that this is an incomplete picture of Jesuit rhetorical education. In fact, the curriculum has a decidedly rhetorical form specific to the Renaissance, with emphases on the development of eloquentia perfecta and regard for the human. Through a comparison of the De Arte Rhetorica of Cipriano Soarez with other sixteenth-century rhetorics, as well as attention to other aspects of the early Jesuit curriculum including St. Ignatius de Loyola's theorhetoric Spiritual Exercises, this paper aims to situate Jesuit rhetoric in the Renaissance and lay the foundation for understanding the mind of Matteo Ricci.

### LETTERS—LITERATURE

# Leaking Paine: Richard Carlile's Illegal Publication of *The Age of Reason*

Andrew Doub

Brigham Young University

This paper examines the concept of the "leak" of censored texts in Romantic-era print culture and its means of production before the

digital age. Between 1818 and 1824, radical printer and publisher Richard Carlile made a determined effort to disseminate copies of Thomas Paine's banned text The Age of Reason in England. Despite strict censorship laws and harsh legal penalties used to curtail previous publishers of this title, Carlile employed a number of creative techniques that kept Paine's deistic writings in print and in circulation during the Regency period. These included republishing public domain court documents and reading The Age of Reason in its entirety into testimony during his trial for seditious libel, making it part of the public record. Carlile managed to provide widescale access to a work that had been successfully suppressed by the British government since its original publication in 1794. My paper argues that Carlile's approach to subverting censorship of The Age of Reason should be viewed as an early instance of what is today known as the mass "leaking" of a restricted document. His unconventional publishing methods, network of collaborators, and use of technology to distribute this work place Carlile at the vanguard of document leaking, making his contribution an important but often overlooked episode in Regency print culture. My study places Carlile's work in a new context and contributes to recent scholarship, including work put forth by Andrew Franta and Russ Castronovo, that seeks to identify parallels between modern surveillance state controversies and similar events from the Romantic period. Like the massive document leaks of our day, Carlile's publication set a new standard for law-breaking innovation in the pirate publication industry.

## LETTERS—LITERATURE

# G.K. Chesterton: Teaching a Religious Writer in the Secular Classroom

#### John Schwiebert

Weber State University

As a writer of nonfiction, G.K. Chesterton is best known for his occasional journalism and biographies and two book-length defenses of orthodox Christianity, *Orthodoxy* (1908) and *The Everlasting Man* (1925). An instructor who assigns Chesterton in a nonreligious-affiliated college or university faces two special challenges. The first is

Chesterton's stylistic eccentricity and penchant for paradox, which can frustrate students and has even driven some religiously sympathetic readers, like T.S. Eliot, to distraction. The second is his dogmatic Christianity. In this paper, I will address the first challenge and concentrate on the second, advancing three major points: (1) Chesterton's great common sense outweighs the challenges and makes his work both assignable and uniquely worth reading. (2) Chesterton is best introduced in essays and prose excerpts that advance his religious sensibility but not, explicitly, his dogma. This sensibility is distilled in his 1936 Autobiography: "I had in childhood, and have partly preserved out of childhood, a certain romance of receptiveness, which has not been killed by sin or even by sorrow. ... Existence is still a strange thing to me...." I will discuss three particular selections and how and why they compel student interest. (3) Chesterton's vision, with its emphasis on such religious virtues as humility, wonder, and regard for the sacredness of the individual (what he termed "the democratic emotion") runs against the cultural grain and is therefore worth airing. Reading Chesterton enables students of all religious persuasionstheistic, agnostic, atheistic, and other-to discuss basic questions of meaning and purpose they are unlikely to be invited to discuss in any other academic context.

### LETTERS—LITERATURE

## Digital Literacy as Evaluative Pedagogy: Assessing the Digital Revolution from Social Networking to Cyber Warfare

#### James Young

Weber State University

In this paper, I will present my findings on the impact of digital technologies on American culture after teaching a freshman English composition course on the Digital Revolution. As a "Digital Immigrant" rather than a "Digital Native," I was curious about how digital technologies are influencing the life of young college students. I wanted to know why they are constantly on their iPhones and iPads and what effect that is having on their social and intellectual lives. I began the semester with an overview of the many ways digital technology is

having an impact on education, business, politics, medicine, psychology, and sociology by reading a group of essays by Sherry Turkle, Nicholas Carr, and John Palfrey in Mark Bauerlein's The Digital Divide. After explaining to these Weber State University sophomores the intricacies freshmen and of research and documentation. I asked them to write three short research papers on the impact of modern technology on American culture. The first paper asked them to describe and evaluate the usefulness of a digital device or process (iPhones to Google Glasses); the second asked students to investigate a new phenomenon in the Digital World (cyberbullying to NSA spying on American citizens); and the third paper asked them to propose a solution to a technological problem (fighting viruses to the international theft of American digital products). When I finish this course, I will report to the Academy what I discovered about my students and the digital worlds that are now a part of all of our lives.

## PHYSICAL SCIENCES

# How Much Less Is More? The Predictive Consequences of Overfitting

#### **Bill Bynum, Brian Knaeble, Gano Hasanbegovic, Garret Wilcox** *Westminster College*

The law of parsimony, or Occam's Razor, states that given competing hypotheses that equally explain the phenomena at hand, one should choose the hypothesis with the fewest assumptions. Statistician R. A. Fisher reminds scientists each data set contains a natural amount of information that cannot be increased through ingenious statistical models. Mathematician John von Neumann claims that with four parameters he can fit an elephant, and with five he can wiggle its trunk. In statistics, a model with more explanatory variables than justified by the sample size is said to be overfit, violating the law of parsimony. Overfit models can predict poorly because they have modeled noise rather than underlying relationships between variables. To avoid overfitting, there is a rule of thumb for regression stating that the number of observations must be at least ten times the number of explanatory variables. To evaluate this claim with respect to prediction accuracy as measured with mean square error (MSE), we have conducted statistical simulations. The simulations have been carried out under a variety of assumptions and the results are tabulated for easy reference.

#### PHYSICAL SCIENCES

### Schrödinger equations with 1-D potential wells

#### Chin-yah Yeh

Salt Lake Community College

Wave mechanics has been proved to work in general, but the hydrogen atom is about the only case where the result of wave mechanics has an exact match with real-world phenomena. Here we raise a few basic questions about why wave mechanics works. Why do we have dimension 3 in our physical space? Why do we end up using the Hamiltonian as the operator in the Schrödinger equation? Hence, the Schrödinger equations of 1-D particles with the following Hamiltonians shall be tested: a)  $H(p,q) = p^{2}/(2m) + k |q|$ ; b)  $H(p,q) = p^{2}/(2m) + k |ln q|$ ; c)  $H(p,q) = p^{2}/(2m) k(exp (-|q|/a))/|q|$ .

#### PHYSICAL SCIENCES

# Experimentally determined optical constants for yttrium oxide in the extreme ultraviolet

## Margaret Miles, David D. Allred, R. Steven Turley, Benjamin D. Smith, Joseph B. Muhlestein, Stephanie Thomas

Brigham Young University

A need for accurate, experimentally determined optical properties of yttrium oxide has grown since alternating layers of aluminum and yttrium oxide have been computed to be a good candidate for high reflectance at 30.4 nm (the wavelength produced by the 2p to 1s transition in He+). Specifically, a knowledge of yttrium oxide's optical properties in the extreme ultraviolet could benefit fields where use of extreme ultraviolet radiation has intensified, such as production of more powerful computer chips, astrophysics, and imaging of protein

structures. We have determined the index of refraction of yttrium oxide using reflectance measurements at the Advanced Light Source. We have compared those measurements with an index of refraction calculated using atomic scattering factors for the component elements. From analysis of our measurements, it appears that the independent atom approximation breaks down near 30.4 nm. In comparing our data with previous measurements for single-crystal yttrium oxide, it also appears that physical characteristics of the yttrium oxide significantly affect the index of refraction.

### PHYSICAL SCIENCES

### **Basic Statistical Adjustment**

#### Brian Knaeble, Jingyi Huang, Thomas Vitti

Westminster College

Causal effects can be estimated from observational data using Bayesian networks, which are directed acyclic graphs representing conditional dependencies between random variables. Here we consider simple networks each made from three dichotomous variables. For each network, we condition on the third variable to produce an adjusted estimate for the effect of the first variable on the second variable. We conduct statistical simulations to demonstrate unbiased estimation and bias amplification, and we assess the sensitivity of bias amplification to network parameters. Results are organized for easy reference and improved intuition about appropriate and inappropriate adjustment.

### PHYSICAL SCIENCES

## Density Functional Theory Investigation of Polycyclical Peroxide Stability

#### **Charles J. Simon, Don R. Davies, H. Laine Berghout** *Weber State University*

Polycyclic peroxide compounds have been of interest recently for their antimalarial activity. The synthesis of stable peroxide compounds can
be challenging, thus making computational determination of the stability of promising compounds advisable. We use density functional theory to gauge the stability of one such peroxide, 2,3,10-trioxabicyclo[5.2.1]decan-4-ol. An intramolecular hydrogen bond between the two rings of this bridged bicyclic compound is anticipated to contribute to the stability of the molecule. Based on coordinate scans of the hydroxyl dihedral bond angle at the B3LYP/6-311+G(2d,p) level of theory, we estimate the strength of this intramolecular hydrogen bonding interaction at 8.6 kcal/mol, considerably above the 5-kcal/mol typical for R-O-H---O=C-R'. Decomposition of similar peroxide compounds proceeds via homolytic bond cleavage of the O-O bond. We will share our continuing work on this system including studies of the activated complex for hemolytic peroxide-bond cleavage and the unusually large intramolecular hydrogen-bond energy.

#### PHYSICAL SCIENCES

# First step towards allowing aluminum's far UV reflectance to be accessible for broadband VUV for space-based instruments

# David D. Allred, R. Steven Turley, Spencer Willett, Stephanie Thomas, Michael Greenburg

Brigham Young University

In four years, NASA will be in the midst of its decadal review, establishing priorities for the 2020s. Very likely one of the chief astrophysical missions will contain a LUVOIR (large UV optical IR) telescope. This space-based observatory will likely contain the largest mirrors ever flown and will probe the cosmos seeking to address key questions of the origin, current status, and evolution of our universe. These investigations will profit from a truly broadband mirror. Thus, the reflective coating will almost certainly be aluminum. To be viable, the top surface of such a space-mirror needs to be bare without the tarnish layers that naturally form. We will discuss our research into protecting as-deposited aluminum mirrors before atmosphere exposure with a robust, protective layer, or layers, that potentially can be easily, and cleanly, removed once the mirror system is in space without marring the mirror surface nor redepositing material removed from the

protective layer on the mirror or other spacecraft components. This could open up the 11-15eV band for space-based astrophysics without sacrificing IR, visible, and UV reflectance. We will report on two systems: First, protective polymer films that can be readily vacuum deposited and later be completely removed with a hydrogen plasma; and second, inorganic films that can be evaporated to coat the aluminum immediately after its deposition, before it comes in contact with air, and which can be expected to reevaporate in space, when heated mildly.

### PHYSICAL SCIENCES

# Emission from Newborn Supermassive Black Holes in the Early Universe

#### Brandon K. Wiggins, Joseph M. Smidt, Jarrett L. Johnson

Southern Utah University

The COSMOS survey recently discovered what is now the brightest known Lyman alpha emitter to date. This source is a young galaxy in the early universe that appears to be devoid of spectroscopic evidence of any elements other than hydrogen and helium. While some believe we may have observed the first galaxy hosting rare first-generation stars, evidence is mounting that the source represents instead the powerful glow of a newborn supermassive black hole. In this paper, we summarize efforts to model the creation of this galaxy ab initio and follow the collapse of a large, atomically cooled halo with radiation hydrodynamical cosmological simulations. These state-of-the-art calculations include chemistry and x-ray feedback allowing us to model the growth of a large black hole with greater fidelity than previous works. We post-process the simulation to obtain emission from the black hole and glowing cosmic web in an attempt to reproduce observed signals and comment on the viability of interpreting this new COSMOS survey galaxy as the first ever observed direct collapse black hole. Our discussion includes a brief introduction to supermassive black hole growth and large-scale structure formation in the early universe for audiences beyond our field.

#### POSTERS

### Intrinsic factor inhibition due to the use of isotretinoin

#### Maikie Sengdeng, Bahar Alimadadi, Ashley Rogers

Weber State University

Acne is a common skin condition that is described as pimples, blackheads, whiteheads, and oily skin typically of the face. Isotretinoin (commonly known as Accutane) is a therapeutic drug used for the treatment and prevention of severe acne by influencing skin cell-cycle progression, cellular differentiation, and cell survival. This drug is a derivative of vitamin A known for its ability to treat acne that has not responded to antibiotics. Isotretinoin is known to express both mild and very serious side effects, such as nosebleeds or spontaneous abortions. A previous study done on the effects of isotretinoin concluded that patients prescribed with isotretinoin showed a significantly prolonged partial thromboplastin time (aPTT). aPTT is a timed, clinical test to detect abnormalities in the coagulation cascade, specifically the intrinsic pathway. The intrinsic pathway is comprised of four coagulation factors: XII, XI, IX, and VIII. Once these factors are activated, they play a major role in hemostasis and the coagulation cascade to form a hemostatic plug that prevents internal bleeding. An irregular or prolonged aPTT indicates that one or more factors of the intrinsic pathway may be inhibited, thus hypothesizing that isotretinoin may be the cause of this delay. The objective of this study is to investigate which factor(s) of the intrinsic pathway is being inhibited. To determine which factor(s) are affected, normal patient plasma will be collected, treated with a standardized concentration of isotretinoin, and undergo a factor assay. The factor assay is comprised of four different factor-deficient plasmas. Each factor-deficient plasma will include the treated patient and an aPTT will be measured. The expected results will show a delay in one or more of the deficient plasmas and will indicate what factor(s) are inhibited in the intrinsic coagulation pathway.

#### POSTERS

# French and American Police Organizational Leadership Training

#### Sarah Gordon, Malcom C. Collier

Utah State University

This study takes a comparative approach to organizational leadership theory and practices in French and American police organizations. We explore the concepts of adaptive leadership and situational leadership in specific examples, including the training programs of the Los Angeles Police Department and the French Gendarmerie. We also compare the use of hands-on activities, simulations, and role play in the two models. The poster provides translations of French materials.

#### POSTERS

# **Religious Freedom: A Global Concern**

#### **Jill Jasperson**

Utah Valley University

World religious freedom is tenuous. American citizens cherish freedom of religion as identified in the First Amendment of the Constitution. However, there are religious/belief atrocities committed across the globe without regard to those freedoms. Because American citizens are concerned about these atrocities, the International Religious Freedom Act (IRFA) was signed into law. This act created a special office in the State Department to defend religious freedom abroad: the United States Commission on International Religious Freedom (USCIRF). This commission creates a lengthy annual report, and has identified the world's top religious freedom abusers in map form. See http://www.uscirf.gov/reportsbriefs/annual-report. The author created this poster identifying these abusers with an interesting twist. The juxtaposition of the countries is a telling story of the sorrow and heartbreak created through lack of this important right.

#### POSTERS

# **Measuring Learning Outcome in Chemistry Outreach**

#### Saúl Quintero, Alan Doong, Ron Valcarce

Salt Lake Community College

Elemental Expeditions is a chemistry outreach program that was created at Salt Lake Community College to provide hands-on chemistry learning to underserved schools in the Salt Lake City area. Through the use of written and pictorial assessment, we measured the change in science-based informational knowledge and perception in K-6 elementary school students. Our assessment tool was administered pre and post classroom presentation and the change in responses was measured.

#### POSTERS

# A Scientific Study Determining Variables Affecting Metal Oxide Oxidation States Coloring High Fire Ceramic Art Glazes

# Victoria Willard, Wendy Schatzberg, Shane Christensen, Heath Papa

Dixie State University

In this study, we have integrated the fields of art and science, more specifically the artistic medium of ceramics with the science of glaze chemistry. High-temperature silica-alumina-based glaze colors are manipulated using appropriate transition metal oxides as colorants. There is a range of variables that all contribute to determine the final transition metal oxidation state used in glazes. This final oxidation state affects the shape of the electron cloud around the metal ions that are left suspended in the cooled glaze. We explored a series of controlled tests and evaluated which glazing variables had the largest impact on the final colors produced. We tested different ways to manipulate the final glaze colors and investigated any impact resulting from including iron oxides into the glaze body. We chose iron oxide because it is capable of producing extensive range of color. Iron oxides distinctively respond to both oxidation and reduction kiln atmospheres, which was an important variable. We ran tests on the iron oxides and varied them by iron types, amount, purity, initial oxidation state, and particle sizes. The particle sizes were varied from a large mesh down to a nano-scale sizes. We tested how the glaze colors produced by these iron oxide were affected by the presence of other transition metals using a bi-axial blend method. Lastly, we fired the glazes on a porcelain clay body in both reduction and oxidation atmospheres. We built a gas test kiln specifically for these experiments that has capabilities reaching and holding temperatures as high as 2200 degrees Fahrenheit.

#### POSTERS

# Flow and Fluidity: Perception and Experience in Drawing Landscapes and Riverscapes for Landscape Architecture

#### **Caroline Lavoie**

Utah State University

This poster includes a theoretical and practical discussion of how drawing enriches and informs the field of Landscape Architecture, focusing particularly on the unique challenges of large-scale landscapes and riverscapes. The project includes large-scale pen and ink drawings and site descriptions as well as reflections on the function and value of drawing. It investigates perceptions, experiences, and representations of large-scale landscapes and bodies of water. Drawing enables a better understanding and internalization of the flow of water as it meanders through space, time, and human experience. The paper/poster explains how drawing enables the artist, and the viewing public, to experience vast watersheds, powerful rivers, trickling brooks, historic canals and bridges, restored streams, and immense canyon landscapes carved by water. It explores drawing water as a creative and communicative tool in stream restoration for hydrologists and stream ecologists. With fluid lines, the example drawings traverse the horizons of expansive landscapes, retracing the path of water. A relationship with the landscape and with the movement of bodies of water is revealed in the flowing lines of the drawings. All sketches were drawn on site, and the project shows how such on-site, large-scale drawings help us to

perceive, interpret, internalize, and interact with the flow of rivers, streams, waterfalls, and bays. Finally, the author explains how drawings such as those exhibited are useful as they focus on different aspects of the landscape. They explore the layers that form the visual landscape (the foreground, middle ground, and background) sometimes blurring their boundaries, while capturing the fluidity and changing nature of the landscape and riverscape.

#### POSTERS

# **Geodata Sets for Emergency Management**

#### **Brian Stearmer**

Weber State University

The project scope had two components: (1) collaborate with the appropriate teams and individuals within the Utah National Guard and Emergency Operation Center to identify important Intelligence Requirements (IR) and (2) identify, organize, and create key Geographic Information System (GIS) data sets that could answer IRs in Defense Support of Civil Authorities operations across the state of Utah. Each feature on the surface of the Earth has a different energy reflectance that can be identified by multispectral imagery. A statewide multispectral baseline image would enable us to understand the components of a disaster and better tailor response efforts, but yet there wasn't one available within any GIS data holding. The key in largescale emergency management scenarios is applying available resources to where they are needed. However, communication lines are also degraded and/or overflowing, identifying where resources are needed is the most obscure and often costly decision leaders have to make. GIS analysts can use remote sensing multispectral imagery to quantify the nature of various disasters without having to take time and resources for manual reconnaissance. For example, solid and broken asphalt have a unique energy reflectance. Using classification software, the analyst can identify damaged roads without having to drive across the entire state. It took several thousand hours to identify, adjust, and compile Utah's 17 different Landsat scenes. I narrowed the timeframe to September, early fall had the least amount of cloud cover while still a good vegetation reflectance. There are automated tools, but manual

color adjustment yielded the best result. Because of the accuracy of the adjustment, seamlines between images are nearly undetectable. The end result was a 99.99% cloud-free mosaic, with an average date of September 4, 2015. Looking to the future, continued work will involve creating seasonal baselines and integrating airborne platforms that can assess post-catastrophe differences.

### POSTERS

# Increasing Rates of Homicide in the U.S.: The "Ferguson Effect"?

#### John Hill, Brianne Hill

Salt Lake Community College

Over 30 major American cities are experiencing a recent increase in homicide rates. The researchers analyze the "Ferguson Effect" as a causal relation or correlation to this increase. The "Ferguson Effect" is identified as overly passive and ultra-reactive policing created after the backlash resulting from the Michael Brown shooting (in Ferguson, Missouri, 2014). The consequence is alleged to be an empowered criminal element proliferation concurrent to a discouraged police force disengaging from enforcement activities.

#### POSTERS

# The Relationship Between Various Forms of Hypoxia and Postpartum Depression

#### Mallory M. Rogers, Perry F. Renshaw, Rebekah Huber

University of Utah

Studies have shown a relationship between hypoxia and depressive symptoms. Hypoxia is defined by Merriam-Webster as a "deficiency in the amount of oxygen reaching the tissues." Hypoxia can be caused by various factors including environmental conditions, medical conditions, or health behaviors. The proposed study will examine the relationship between various causes of hypoxia and the occur rence of

postpartum depressive symptoms. The proposed study will define hypoxic conditions as: residence at high altitude (≥2000 ft.), moderate to heavy smoking (as defined by PRAMS data), asthma, and hypertension. Subjects will be divided into five categories: 1-high altitude, 2-smokers, 3-hypertensive, 4-asthmatic, and 5-nonhypoxic. Using a multivariable analysis in SPSS, a logistic regression will be used to examine the relationship between postpartum depressive symptoms and each form of hypoxia. We hypothesize that women in a state of hypoxia will be more likely to experience postpartum depressive symptoms. This hypothesis is supported by literature on hypoxia and mood disorders altitude and suicide, smoking and depression, and asthma and depression. There is currently very little research on postpartum depression or the variables associated with it. As postpartum depression affects 10-15% of American mothers, an increased understanding of these factors may serve to further advance prevention and treatment of postpartum depression.

#### POSTERS

# Oil in the Watershed

#### Abby Jorgensen, Robyn Hyde

Westminster College

The goal of this ongoing research is to develop a QuEChERS (quick easy cheap effective rugged safe) extraction method to determine the level of contamination of oil in the snow in the Salt Lake City watershed. This requires the extraction of nonpolar organics from polar water. Extracts are analyzed by synchronous scan fluorescence spectroscopy (SSFS). Sample preparation requires only 4 mL of hexane and 0.2 g of salt. Total preparation analysis time has been shortened to approximately 1 hour. Reference materials have been created by spiking distilled water with various motor oils and gasolines. Findings show there is linearity between concentration and fluorescence intensity. Calibration is possible, and the limit of detection is at or below 100 ppm. Semi-quantitative data will be presented showing levels of contamination on snowmobile trails, ski resort parking lots, and mountain lakes with high traffic. This research could help inform

policies that are meant to protect water quality in the Salt Lake City watershed.

# SOCIAL SCIENCE

# **Indirect Health Effects of War**

#### **Daniel Poole**

Salt Lake Community College

This study examines the impact of armed conflict on female and male adult cardiovascular disease mortality. Indirect health consequences of war have not been given enough attention in social science research. The depletion of resources, access to health care, and general disruption to everyday life during times of armed conflict create excess stress and burdens that increase deaths caused by cardiovascular disease. I use a variety of data to measure demographic, developmental, and conflictrelated outcomes spanning a forty-year period from 1960 to 2000 in more than one hundred countries. I find that all types of armed conflict increase cardiovascular disease mortality rates among females and males across countries and over time, with the effect being greater on females.

#### SOCIAL SCIENCE

# **Quality of Life and Direct Democracy**

#### Josh Smith, Ryan Yonk

Utah State University

We explore the relationship between quality of life and direct democracy. We measure quality of life by creating our own index that includes indicators of public safety, health, economic development, infrastructure, and education. Although direct democracy in the United States is limited by the growing influence of interest groups and political parties, ballot measures remain one indication of citizen preferences through direct democracy. Thus, we use ballot measures as an indicator of direct democracy. To explore the relationship between quality of life and direct democracy, we employ simple statistical tests. We find that citizens with a higher quality of life tend to prefer the status quo, rather than voting for a change to the existing policy structure. Citizens in states with higher quality of life tend to vote against ballot initiatives and fewer ballot initiatives ultimately become law in high quality of life areas. These findings imply that, in the case of ballot initiatives, quality of life has a significant impact on voter turnout. Our research reveals interesting patterns surrounding how and why citizens engage in the democratic process.

#### SOCIAL SCIENCE

# Effects of Roman Imperialism on Central European Populations

#### L. Brock James

University of Utah

Using paleodemographic data gathered from European skeletal assemblages dating from between the Iron Age and the Early Medieval periods, we attempt to draw conclusions about whether the benefit of Roman Imperialism, primarily public works such as sanitation and military protection, outweighed the epidemiological impact of an increased susceptible population size due to increased trade and freedom of travel for Roman citizens. Although basic epidemiological theory would indicate that the sudden increase in susceptible population size should have been followed by outbreaks of epidemic disease, we propose that this impact was softened by the effective infrastructure of the Roman state. To test this conclusion, our data set has been subdivided into three temporal groups: Iron Age (pre-Roman), Roman, and Early Medieval (post-Roman). Demographic data will be used to create a hazard analysis of death during these periods in an attempt to see changes in the overall "shape" of death, using skeletal pathologies as a proxy for quality of health, in an attempt to judge the quality of public health before, during, and after Roman occupation.

# SOCIAL SCIENCE

# Depleted Uranium Munitions and the Mind of the Killer Robot

#### Shadman Bashir

Dixie State University

Depleted uranium munitions are the tip of a tank-busting spear. They are great for penetration of heavy armor, and they are excellent kinetic energy weapons. Are they dangerous to the enemy? Yes. Are they dangerous to all of us? This paper is a brief study of the myth and reality of the depleted uranium munitions and their use in military and civilian environments to find the answer to this question. The second part of the paper extends the issue of depleted uranium munitions to killer robots and the use of such weapons by enhanced artificial intelligence within the conventional and unconventional theaters of war and conflict.

#### SOCIAL SCIENCE

#### The Death of Hetero/Homo

#### alithia zamantakis

University of Utah

This paper is a theoretical examination of the ways in which sexuality, love, and desire are not merely abstract, innate concepts but have very real consequences as weapons in the process of abjection, particularly of trans, nonbinary, and gender-nonconforming individuals. As gender is illegible and fluidly defined, it becomes impossible to dictate that one's attraction is oriented toward men, women, masculinity, femininity, and/or androgyneity. Ultimately, individuals are attracted or not attracted to particular body parts, personality traits, values, and life goals. None of these can be attached to a particular (a)gender as anyone of any (a)gender can have any number of these qualities. What then does it mean to have a discourse, as well as a movement, around conceptions of heterosexual and homosexual identities that deny such a reality? The answer to this question—the erasure and eradication of particular bodies—calls for the death of hetero and homo, allowing the

rethinking of these terms and ideas. This shifts from a place where certain genders are abjected and silenced to a place where all genders are equally affirmed, no one identity superimposed over another. This is not an argument for attraction to all bodies and all beings. However, it is an examination of a discourse of sexuality that is posited on fixed and concrete notions of gender and the ways in which this discourse is used to abject trans, nonbinary, and gender-nonconforming bodies.

#### SOCIAL SCIENCE

# Survival: The Evolution of Jewish Identity in the 20<sup>th</sup> Century

#### Nathan G. Caplin

Snow College

While reflecting on his civil rights leadership, Martin Luther King, Jr. observed "the art of alliance politics is more complex and more intricate" than most of its participants understand. The principle Dr. King observed was true for the Jewish and African American alliance from post-WII through the mid-1960s. An often-overlooked cause for the dissolution of Black–Jewish alliance is the transformation of Jewish Americans from ethnic minority to nearly assimilated "white" during the 20<sup>th</sup> Century. This paper explores the evolution of Jewish ethnic identity in the context of Black and Jewish cooperation and divergence in their quest for equal rights.

#### SOCIAL SCIENCE

# Survival: Exploration of African American Health Priorities

#### Tashelle B. Wright Westminster College

Health disparities among the African American population are greater than those of whites and other ethnic groups. African Americans live on average 8 years less than non-Hispanic White Americans. The

African American population in Utah is 1.3%, the smallest minority group in the state, approximately 40,000 residents. In Utah, we have the added challenges of lack of cohesion and self-identity, which other states with larger African American populations do not experience. With a lack of African American networks, outside of churches, health disparities are increasingly difficult to address. To address the health priorities of Utah's African American population, we need to identify what the African American community view as priorities. The present study used an exploratory approach and sought to establish a community taskforce to network and advocate on behalf of the African Americans in Utah. This study also determined the self-identified concerns of African American community members. Working in conjunction with the Utah Office of Health Disparities, I identified and contacted African American community leaders. We were able to establish a taskforce of 22 people. I developed a survey to distribute in person at the Juneteenth festival in Utah. The survey asked African American community members to self-identify their top three priorities. There were 66 survey participants (N=66). The top three African American priority themes selected were mental health, healthy eating choices, and obtaining an education. These surveys were further analyzed to determine the relationships between demographic variables. There was no statistically significant difference between genders. Age only played a role in the category of criminal justice. The older generation are more likely to select criminal justice as a priority (p=0.0501). With this information, we can now take action in the areas viewed most important to the African American community in Utah.

# SOCIAL SCIENCE

# Becoming Psychologically Literate: Responding to Those with Symptoms of Depression

**Amy Blommer Sophia Garcia, Claire Short, Eric Amsel** *Westminster College* 

Over 1000 participants were electronically presented with the following open-ended scenario: "Your friend or close family member shares with you that he or she has been feeling down lately. The person reports a loss of interest in activities he or she used to enjoy and instead just

stays at home, often sleeping for long periods of time. The person has seen a general practitioner to get a checkup, and the person received a clean bill of health. But the person still wants your advice. What advice would you give this person? Please be detailed." Participant responses were coded for the level of psychological literacy (PL) expressed on a 0-2 point scale. No PL responses showed no awareness of the possibility of a mental health disorder, recommending instead selfmedicating activities. Partial PL responses recognize the potential of a mental illness by recommending the person seek a second professional opinion, but continue to offer suggestions for self-medicating activities. Complete PL again involves the recognition of a potential mental illness but without any self-medicating advice. PL scenario responses were scored reliably (inter-rater reliability = .95). The assessment was valid demonstrated by higher PL scenario scores being correlated with participants' a) ratings conceptualizing depression as a mental illness (as opposed to a mental weakness), b) skills to identify depressed individuals, and c) greater general background in psychology. PL scenario scores were higher for students with Psychology majors than for those with Psychology minors and for those with minors than for those who were neither majors nor minors. Scores were also higher for Sophomores/Juniors and Seniors than for Freshmen. Finally, PL scenario scores were higher for Psychology majors than for a group of non-college-educated adults. Results suggest that psychological literacy can be trained, and its development in psychology students may be important for community mental health.

#### SOCIAL SCIENCE

# Finding Meaning in Serving Others: Factors Predicting College Students' Self-Perceptions of Being an Adult

#### Hui-Tzu Grace Chou, Ron Hammond

Utah Valley University

This study examines the impact of finding meaning in serving others on self-perceptions of being an adult among college students between 18 and 25 years old. Using probability sampling, current and former students of a state university in Utah were selected to participate in an online survey in the summer of 2014. A total of 572 respondents met

the age requirement and completed all the survey questions. The results of a multiple regression indicate that those who are depended upon by others and those who find meaning in serving others are more likely to perceive themselves as adults than their counterparts. This study also found that the self-perception of being an adult is related to whether respondents had moved out of their parents' home, had gotten married, and were financially independent.

#### SOCIAL SCIENCE

# Positive Psychology Holistic Determinants, Testosterone Treatment, and Veteran Happiness

#### James C Brau, Hannah L. Brau

Brigham Young University

In this study, we designed a survey instrument and constructed a data panel from the responses of a sample of U.S. veterans. As part of the survey, we estimated the level of happiness each veteran exhibits using the Oxford Happiness Questionnaire. The Oxford scale consists of 29 questions and uses a Likert scale that ranges from a low of one to a high of six. The average is typically around 4.3. The average happiness score for this sample of 76 veterans was 3.73. After measuring the happiness level, we asked 30 additional questions driven by the literature to determine the factors of veteran happiness. Next, we conducted Spearman Correlation tests, t-tests for equality of sample mean divided on the median of the happiness score, and a multivariate ordinary least squares model with all of the explanatory factors. We found significance for: four holistic happiness variables (faking happiness (+), spending money on loved ones (-), listening to music often (+), and using technology often (-)); one demographic variable (age (+)); six military-related variables (active duty service (+), years of service (-), months deployed (-), service in Iraq (+), Afghanistan (+), Korea (+)); and two intervention variables (psychotherapy (-), exercise (+)). Testosterone treatment was not statistically significant.

# SOCIAL SCIENCE

# Behavioral Economics and the Value of a Statistical Life

#### Alecia Hunter, Ryan Bosworth, and Ahsan Ulkibria

Utah State University

Behavioral economics suggests that the majority of the people underestimate large risks and overestimate small risks. The U.S. Environmental Protection Agency's (EPA) uses the Value of a Statistical Life (VSL) estimate in a cost-benefit analysis for regulating different environmental hazards. The EPA defines the VSL as "how much people are willing to pay for small reductions in their risks of dving from adverse health conditions that may be caused by environmental pollution." The EPA's VSL is currently calculated at about \$8.7 million using stated preferences and revealed preferences studies. In this paper, we find evidence that the VSL may be overestimated due to behavioral biases (mainly misperceptions of risks) found in the VSL's preference studies. A second finding is that policymakers are also vulnerable to behavioral biases, which is often overlooked before policy is drafted. Using behavioral economic theories, including prospect theory and framing effects, we support our hypothesis in the example of the EPA's VSL estimate: People often misperceive risks, and policymakers are also prone to similar behavioral biases

# SOCIAL SCIENCE

# **Intelligence-Led Policing**

#### John Hill

Salt Lake Community College

Intelligence-led policing (ILP) is one of the newest proactive strategies in contemporary law enforcement. It is a research-based policing model built around the assessment and management of risk. ILP serves as guide to operations, rather than operations guiding intelligence. It is believed that as agencies collect, evaluate, and disseminate information, a comprehensive "network of intelligence" is created, which proponents hope, when used effectively, will cause a significant decrease in crime. This researcher analyzes ILP with a special emphasis on its role in tactical and strategic planning for police patrol operations. The researcher theorizes that ILP can be utilized to assuage recently increasing tensions between the police and the community, as intellect and information initiate police activity in a manner preferable to more discordant tactics, such as "stop and frisk."

#### SOCIAL SCIENCE

# Framed Couplets: Agenda Setting, the Bumper Sticker Effect, and How Candidate Name Pairings Predict Winners and Losers in Presidential Elections, 1840– 2012

#### Thomas C. Terry

Utah State University

Vice President John Nance "Cactus Jack" Garner colorfully quipped, "The vice-presidency isn't worth a pitcher of warm piss." while journalist Bill Vaughan added, "The Vice-Presidency is sort of like the last cookie on the plate. Everybody insists he won't take it, but somebody always does." Whether urine or cookie, a presidential candidate has to offer the job to someone. Traditional wisdom suggests a presidential candidate should select a vice president to shore up perceived weaknesses in his/her credentials, such as in foreign policy: to balance (or counterbalance) issues of age, religion, or geography; to heal intraparty rifts; or to deliver electoral vote-rich states. However, the choice of a vice presidential running mate is far simpler than any of the complex host of metrics and rationales that often goes into a vice presidential decision. In fact, the answer is so straightforward it fits onto a bumper sticker. Picking the proper pairing of names, more than any other vice presidential variable, may be the key to victory. What this research study terms the Bumper Sticker Effect model has predicted the winning candidate in 41 of the last 44 presidential elections (93.2 percent). It is based on the differential in letter length between the presidential and vice presidential candidates' last names. To be successful, the presidential winner's name must either exceed that of his/her running mate or be no more than two letters shorter. Cognitively, voters appear to perceive that the presidential candidate is

somehow diminished in stature when the VP's name dominates that of the prospective president. The losing Gore/Lieberman (4 letters vs. 9 letters) and winning Eisenhower/Nixon (10-5) pairings are prominent examples of this phenomenon.

#### SOCIAL SCIENCE

# The New Normal: The Chinese Response to Terrorist Threat in Xinjiang China

#### Huiying W. Hill

Weber State University

The purpose of this paper is to explore how the Chinese government (Xinjiang government) deals with security issues in Xinjiang Uighur Autonomous Region after numerous attacks, killings, and bombings instigated by some extreme Uighur Muslim groups that were inspired and trained by the Taliban and other extreme Islamic groups in Central Asia. The author went to Urumqi, the capital city of Xinjiang, and another city, Korla, to conduct this research in May 2015. The preliminary findings from this study are multi-fold: 1) There has been a significant policy change/shift after the new governor, Zhang Chunxian, replaced Wang Lequan for his inefficient dealings with the 7.5 riots in Xinjiang in 2009. 2) There are also heavy presentations of armed forces everywhere in Xinjiang. 3) Metal detectors are placed at entrances of all the public arenas, such as in hotels, restaurants, department stores, malls, even public squares. The author conducted some interviews of the citizens in both cities as well. The purpose was to find out how ordinary people react to these inconvenient devices and how they cope with all the measures of security in their daily lives. Some governmental officials and public security/legal personnel were also unofficially interviewed. In the United States, we also face more and more threats from extreme Islamic groups and ISIS members. This study might shine some light on this issue, and Americans might learn what we might face in the future and how our daily lives might be altered in facing those threats and preventing terrorist acts from happening. The major methods used in this study were personal interviews, observations, and secondary sources.

# SOCIAL SCIENCE

# Water Issues in the State of Utah

#### Daniel Poole, Heather Aagard-Jimenez, Ryan Butler, Steven Graham, Zachary Hansen, Gabriela Martinez, Kathryn Thompson Salt Lake Community College

Students in the Doing Sociology course at Salt Lake Community College are helping to collect data for a collaborative iUtah project. This research project is an National Science Foundation-funded project led by Professor Doug Jackson-Smith at Utah State University. Researchers and students across most Utah institutions of higher education have been involved with various phases of the project. Salt Lake Community College students have collected survey data regarding public perceptions of water and other environmental issues in the state of Utah. The two-minute surveys are conducted using iPads to sample shoppers entering area grocery stores. Data collected from multiple phases of the study will be analyzed to explore demographic information, attitudes regarding water needs, costs, quality, and uses, and also levels of concern surrounding other environmental and recreational issues. Descriptive statistics and bivariate analyses will be used to explore the relationship as well as the perception that Utahns have regarding water and other resources.

# SOCIAL SCIENCE

# Uranium Legacy of Southern Utah

#### Gulsumkhanum Bayazitova

Westminster College

The research focuses on uranium legacy in Southern Utah, left after the uranium boom of the 1950–1970s, in the period of Cold War. After examining cases of uranium mines in Southern Utah, it was discovered that poor maintenance of the sites of former uranium mines imposes threat to the local environment as well as to people residing nearby. The research includes an investigation of the loopholes in the enforcement of corporate responsibility, as well as the current situation/ status of the uranium sites. This study involves analysis of so-called

"stand-by" uranium mines in San Juan County (Southern Utah), which have not been functioning for years. Although they are not officially shut down, neither had they gone through remediation process. This research also includes the case of Uranium Mill Tailings Remedial Action (UMTRA), which is a clean-up of uranium mills tailings located on the West bank of the Colorado River, near the town of Moab. This research is a part of a broader history of the nuclear arms race and uranium legacy of the Cold War, as well as current maintenance of that legacy in Southern Utah. The research focuses on the corporate responsibility of the uranium mine and mill owners, potential environmental threats of those uranium sites, within a framework of environmental justice towards the residents of San Juan County. Recommendations in regulations are also included in this research, which might ensure more effective corporate responsibility for maintenance and remediation processes of uranium mines.

### SOCIAL SCIENCE

# Federal Relations in the American West

#### Jamie Nelson

Salt Lake Community College

After the United States acquired what would eventually be known as the American West, groups of American and European immigrants made their way across the mountains and deserts to build settlements that had been denied to many of them elsewhere in the United States. The relationship these immigrants forged with the U.S. federal government based on their reputation as "outsiders" would set the stage for a history of anti-western sentiments and actions held and taken by the federal government. These actions continue to affect the lives of those in the rural American West today. This paper will illustrate how the federal government has shown discriminatory practices against the people of the American West through three events: The settlement of the American West, nuclear testing in the United States, and the historical and current events surrounding the debate over federal ownership of land. Through the use of historical records, existing legislation, and reports of current events, this paper will provide evidence that anti-western biases exercised by the United States government have produced negative effects on the people in the rural

American West. (\*The use of west or western in the context of this paper refers to states west of Kansas belonging to the Continental United States.)

### SOCIAL SCIENCE

# Information Technology or Demographic Transition? Testing Competing Hypotheses of Retention and Disaffiliation among Utah Mormons

#### Rick Phillips University of North Florida

Utah provides a natural laboratory for investigating the effects of demographic transition and technological change on religious identity and activity. Recent survey and polling data have revealed a rising trend in defection from religion in the United States. About one in five Americans now claim to have no denominational preference or formal church affiliation. Like most churches, The Church of Jesus Christ of Latter-day Saints (the LDS, or Mormon Church) has been affected by this trend. Studies show that between one fourth and one third of those who say they were raised LDS no longer consider themselves members of the church. This is up from around 10 percent a generation ago. Moreover, studies show that rates of defection from Mormonism are rising faster in Utah and the Intermountain West than in other parts of the United States. Two hypotheses have been put forward to explain this trend. The first is that new information technologies have lured certain members away from the church with attacks on Mormon history and theology. The second is that changing demographics in Utah have transformed the state's religious subculture, making it easier for less committed Mormons to leave the faith. This paper evaluates these hypotheses and reinterprets of the meaning of data that seem to show that patterns of retention and disaffiliation among Utah Mormons are changing.

#### SOCIAL SCIENCE

# Childhood Pornography Exposure: An Exploration and Comparison of the Female Experience and Effect

#### Lacy A. Bentley,<sup>a</sup> Cameron John,<sup>a</sup> Kent D. Hinkson, Jr.<sup>b</sup>

<sup>a</sup>Utah Valley University, <sup>b</sup>University of Utah

Numerous conversations are happening around pornography usage and addiction. Most previous studies focus on men, whether they are about usage, addiction, initial exposure or subsequent effect. This study investigates a broader context of "pornography" and seeks to understand the female dynamic of not only exposure, but subsequent usage and current belief systems. Striking differences between what women use after initial exposure and remembered feelings after their first experience are found in this study. It is our hope this research will broaden the definition and dialogues around pornography. Participants in this study report exposure at lower mean age than previously reported. We therefore propose a re-evaluation of mean age at time of exposure and follow up studies addressing the unique dynamics of the female experience.

#### SOCIAL SCIENCE

# Durkheim's Social Solidarity in the Modern Life: Experiences of Mormon Women in Transition

#### **Bethany Gull**

Utah Valley University

In *The Division of Labor in Society*, Émile Durkheim wrote extensively about the social forces that work to bind societies together or move them towards division and disintegration. Many of his ideas still have relevance today, as societies, and groups within those societies, are continuing to shift and change throughout the world. This paper will explore Durkheim's typologies of mechanical and organic social solidarity in relation to the process of transitioning out of a highdemand religion, the Church of Jesus Christ of Latter-day Saints, or Mormon Church. The Mormon Church functions as a comprehensive religious and social community for its members, much in the ways Durkheim described mechanically organized societies. Specifically addressed in this paper will be how the women's religious disaffiliation disrupted their relationships with family, friends, and other social groups and their journeys to rebuild these relationships as well as build new communities for acceptance and support. Through intensive interviews and the content analysis of a Facebook support group for transitioning Mormon women, Durkheim's ideas of social solidarity will be used to explore the shift in the women's understanding of their place in the world from one organized in a tight-knit, mechanical way to one based on broader, organic ties and the challenges that accompany a change of that magnitude.

### SOCIAL SCIENCE

# Clothing the Angel in the Home: Constructing Emphasized Feminity in the Church of Jesus Christ of Latter-Day Saints through Modesty Rhetoric, Practices, and Policing

#### CoCo James

University of Utah

In this article, I examine the official modesty rhetoric of The Church of Jesus Christ of Latter-Day Saints (Mormon) for males and females across the life course. I performed a content analysis on all references to modesty contained in the three official church magazines between 2000 and 2015, General Conference talks during that same time period, and several key documents within the Mormon canon. This content analysis revealed modesty in both attitude and dress as a deeply held belief in Mormonism with personal and spiritual implications beyond mere physical presentation. Despite those implications, Mormon modesty rhetoric, particularly that directed toward young men and women, emphasizes modest prescriptions of dress that reveal a significant cultural investment in gender essentialism and controlling female bodies. In that age group, policing of female bodies is encouraged and taught through official church communication and activities. Ultimately, this serves to reify the patriarchal system Mormons believe to be divine, with its attendant hegemonic

masculinity and emphasized femininity. In addition to this, all Mormons are taught to be aware of and responsible to multiple constructions of common gaze, particularly a weighty construction of male gaze toward females. Finally, there is a front stage/back stage difference in gendered modesty rhetoric between officially correlated church materials and talks given in General Conference or contributions to church magazines by lay members.